Adaptive Behaviour of Metaheuristic Algorithms in Solar Radiation Prediction: Influence of Input Variable Informativeness

UDC:551.521.1:004.942
Original scientific paper https://doi.org/10.46793/aeletters.2026.11.2.2

Ali Khazzar[1*], Djelloul Benatiallah[1], Kada Bouchouicha[2]

[1]LDDI Laboratory, Department of Material Sciences, Faculty of Material Sciences, Mathematics and Computer Science, University Ahmed Draia, Adrar, Algeria
[2]Center for Renewable Energy Development (CDER), Bouzareah, Algiers, Algeria

Abstract:
This study focuses on short-term (10-minute) forecasting of Global Horizontal Irradiance (GHI) using artificial neural networks (ANNs) enhanced by three metaheuristic optimization algorithms: the FireFly Algorithm (FFA), Particle Swarm Optimization (PSO), and the White Whale Optimization Algorithm (WWOA). The models were trained using meteorological and astronomical data collected from a monitoring station in Khenchela, Algeria. To identify the solar radiation component most strongly correlated with GHI, one additional radiometric input, selected from the available components, was introduced in separate experiments. Model performance was assessed using standard statistical metrics: relative Root Mean Squared Error (rRMSE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). Results indicate that meteorological and astronomical variables alone are insufficient for highly accurate GHI prediction. However, when Global Tilted Irradiance (GTI) was incorporated as an additional input, all three hybrid models exhibited significantly improved accuracy. The ANN-FFA model achieved the best performance, with rRMSE=3.73%, MAPE=6.07%, and R^2=0.9962. These findings demonstrate that GTI is the solar radiation component most closely related to GHI under the studied conditions. Furthermore, the study confirms the effectiveness of FFA, PSO, and WWOA in optimizing ANN hyperparameters for solar irradiance forecasting, with FFA yielding the most robust results.

ARTICLE HISTORY
Received: 4 June 2025
Revised: 20 March 2026
Accepted: 4 April 2026
Published: 29 June 2026

KEYWORDS
Renewable energy, Solar energy system, GHI, ANN, Metaheuristic algorithms, Predictions

1. INTRODUCTION

The continuous increase in energy demand necessitates the exploitation of all available energy sources on our planet [1], including solar energy, which is considered clean, renewable, and accessible in most countries worldwide [2]. The effective deployment of solar technologies, particularly photovoltaic (PV) systems, depends critically on precise knowledge of local solar radiation [3,4]. Key irradiance components, such as GHI and Plane-of-Array (POA) irradiance, directly influence critical aspects like system sizing, performance prediction, and grid integration [5,6]. Without accurate, high-resolution data, even well-designed PV installations risk underperformance and introduce significant uncertainty into energy planning and investment [7,8]. Therefore, accurate solar data are essential, motivating the establishment of solar radiation monitoring stations in numerous regions [9]. However, this approach remains challenging due to the high costs of installation and maintenance [10]. Consequently, researchers in this field have focused on predicting solar radiation using artificial intelligence and machine learning techniques in areas with similar

*CONTACT: Ali Khazzar, e-mail: khazzarali@univ-adrar.edu.dz

climatic conditions [11]. Researchers have developed various models, such as artificial neural networks, adaptive neuro-fuzzy inference systems, random forests, support vector machines, and deep learning models, that can forecast solar radiation using historical data from specific regions and have achieved promising results [12]. To address this topic, the following section reviews select prior studies on solar radiation prediction using diverse techniques across various global regions.

In 2021, Ağbulut et al. [13] investigated the prediction of daily global solar radiation in four Turkish cities with varying characteristics using the following machine learning models: ANN, support vector machine (SVM), deep learning, and k-nearest neighbour (k-NN) algorithms. The results demonstrated the superiority of the ANN, which achieved an R^2 of 0.932, whereas the k-NN model performed poorly. To predict surface longwave radiation (SLR) using ANN, Al-rubaye et al. [11] studied two distinct models: the first relied on data measured in the Al-Qadisiyah region of Iraq, while the second utilized data obtained from the National Aeronautics and Space Administration (NASA) for the same location during 2022. The results of the study [11] indicated that the model that relied on data from NASA outperformed the other model, with an R^2 of 0.823. In Brazil, Solano and Affonso [14] predicted solar radiation for various time horizons (one hour, two hours, three hours) using several machine learning models: Voting Average (VOA), SVR, XGBoost and CatBoost, leveraging real data from the Salvador city station. Model comparison revealed that the VOA model, referring to three approaches of voting average, delivered the best performance across all time periods. Using deep neural networks, Marinho et al. [15] conducted short-term predictions on the west coast of the United States over time horizons ranging from 5 to 30 minutes. A comparison of results revealed that architectures employing 1D-CNNs and isolated Long Short-Term Memory (LSTM) layers generally performed best over a 10-minute timeframe. Over the long term, Vanlalchhuanawmi et al. [16] compared machine learning models with deep learning models across various time periods in the United States. The results concluded that the Recurrent Neural Network (RNN) model outperforms in short-term predictions, whereas the Gradient Boosting Regression (GBR) model excels in long-term predictions. Five advanced machine learning models (Gradient Boosting) were developed by Zerouali et al. [17] for hourly solar radiation forecasting using data from three Australian stations. The results indicated that GHI was the most influential predictor, while temperature had a slight positive effect and Relative Humidity showed a negligible impact.

In Africa, numerous researchers have shown increasing interest in solar radiation forecasting, particularly in North Africa, which is recognized as one of the regions with the highest sunshine duration worldwide. In five distinct regions of Morocco, Bounoua et al. [18] compared empirical and machine-learning models. Comparing the results, the machine learning-based models showed clear superiority, with the random forest model achieving an R-value of 87.53%.

The field of solar radiation forecasting is witnessing growing interest, particularly in countries with high solar potential. In Algeria, this topic is attracting the attention of many researchers across the country, given the high hours of sunshine enjoyed by most regions. Among the most prominent studies on solar radiation forecasting in western Algeria, Soukeur et al. [19] also developed a neural network model for solar radiation forecasting. Error analysis results demonstrated the model's effectiveness, suggesting that the neural network model is suitable for such regions. Southern Algeria, characterised by abundant sunlight yet challenging atmospheric variability, has become a key area for developing hybrid AI models in solar radiation forecasting. Benatiallah et al. [20] examined three desert cities and found that while random forest performed well (R =0.9255), a hybrid ensemble with XGBoost achieved even higher accuracy (R=0.9640). Similarly, Halima et al. [21] compared a standard ANN, a GA-enhanced ANN, and an ANFIS. Their results confirmed the superior performance of the GA-ANN hybrid (R=0.9005, MAPE=8.40%).

Previous studies have addressed numerous models for estimating solar radiation, including the widely utilized ANN model. However, most investigations have focused on developing or comparing various predictive models, while the application of metaheuristic optimization algorithms to enhance ANN models has remained limited or partial. Such applications often confined themselves to a single algorithm, without comprehensive or systematic comparative analyses of different optimization strategies. Furthermore, a substantial number of studies have not thoroughly examined the additional impact of solar radiation components when incorporated as inputs to predictive models; instead, they typically employ a fixed set of variables without evaluating the

contribution of each radiation component to improving prediction accuracy. This limitation hampers the scientific understanding of the roles played by distinct physical components of solar radiation, leading to models that may be numerically effective but unoptimized in terms of input selection.

This paper presents an ANN model optimized using three metaheuristic optimization algorithms (FFA, PSO and WWOA) to determine the optimal model structure, along with an analysis of the additional effect of solar radiation components, which showed a clear contribution to improving the predictive model's performance. The main contributions of this work are threefold: (1) development of an ANN-based predictive model optimized via three metaheuristic inference algorithms, and evaluating the effectiveness of the algorithms in finding the optimal model architecture; (2) Studying the impact of solar radiation components when integrated as model inputs, while evaluating the contribution of each component to improving prediction accuracy; and (3) comprehensive statistical evaluation of the tested models' performance using established evaluation metrics, facilitating an objective comparison between the basic model and the models improved by optimization algorithms and additional solar radiation components.

2. MATERIALS AND METHODS

2.1 Study Areas and Data Collection

Measured solar radiation and corresponding meteorological data from the city of Khenchela (Table 1) were used in this study. The dataset spans the period from January 12, 2015, to April 30, 2016, and comprises 8,209 data points recorded at 10-minute intervals, including night-time periods when solar radiation is absent. Khenchela is located in northeastern Algeria (Fig. 1), at a latitude of 35.43°N, a longitude of 7.14°E, and an elevation of 1,128 m above sea level. The average annual sunshine duration in northern Algeria is approximately 2,000 hours, and the mean daily solar energy potential on a horizontal surface is about 5 kWh/m² [22].

Solar radiation measurements at the Khenchela station were carried out using a Kipp & Zonen CMP6 pyranometer operated by Shariket Kahraba wa Taket Moutadjadida (SKTM), formerly known as SKTM. To ensure data quality and enhance the reliability of the results, all missing values and outliers were systematically identified and removed.

Table 1. Meteorological and astronomical input parameters

Parameters	Abbreviation (Unit)	Category
The year	Y	Meteorological
Day of the year	D	Meteorological
Temperature	T_{avg} (K)	Meteorological
Relative humidity	RH (%)	Meteorological
Atmospheric pressure	BP (hPa)	Meteorological
Wind speed	WS (m/s)	Meteorological
Declination	DE (°)	Astronomical
Hour Angle	HA (°)	Astronomical
Extraterrestrial Irradiation	H_0 (Wh/m^2)	Astronomical

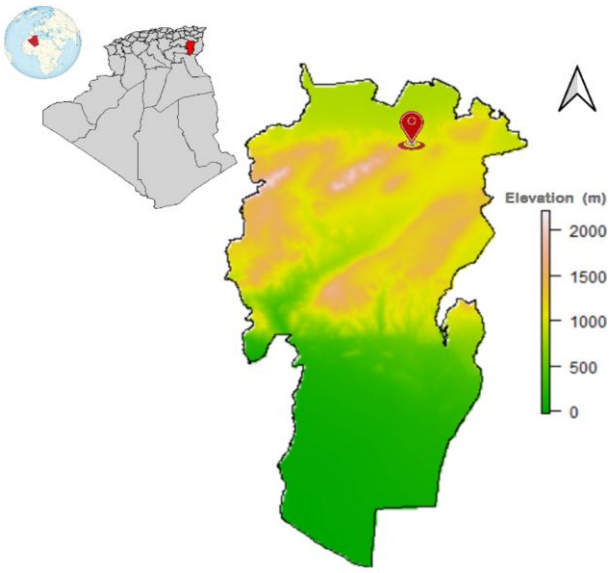

Fig. 1. Map of Algeria with the Studied Locations

To design more efficient solar systems, we need to understand the relationship between the components of solar radiation, Direct Normal Irradiance (DNI), Diffuse Horizontal Irradiance (DHI), GTI, and their direct relationship to the Global Horizontal Irradiance [23]. Fig. 2 shows a detailed diagram of the solar radiation incident on the Earth. DNI refers to the amount of direct solar radiation from the solar disk incident on a surface perpendicular to the sun's rays, while DHI denotes the quantity of solar radiation scattered within the atmosphere.

The GHI is defined as the total solar radiation received on a horizontal surface and is expressed as [24]:

$$GHI = DHI + DNI \times cos(\theta_z) \quad (1)$$

where is: θ_z - is solar zenith angle.

The GTI represents the total solar radiation incident on a surface tilted at an angle β and is calculated using the isotropic sky model as:

$$GTI = DNI \times cos(\theta_i) + DHI \times F_{sky} + \rho \times GHI \times F_{ground} \quad (2)$$

where are: θ_i - is the angle of incidence on the tilted surface, F_{sky} - is the sky view factor between the collector and the visible part of the sky, F_{ground} - is the ground view factor between the collector and the visible part of the foreground surface. The mathematical expressions for F_{sky} and F_{ground} are given as follows:

$$F_{sky} = \frac{1 + cos\,\beta}{2}, \quad F_{ground} = \frac{1 - cos\,\beta}{2} \quad (3)$$

where is: β - is Surface inclination angle relative to the horizontal plane.

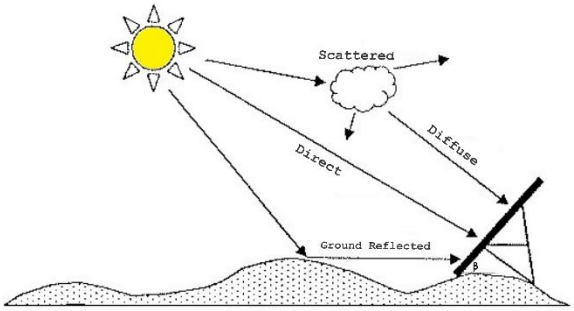

Fig. 2. Components of solar radiation incident on an inclined surface

2.2 Modeling

2.2.1 Artificial Neural Network

ANNs are inspired by the human brain's neuron network and are typically organized into layers, which are made up of interconnected nodes of input, output and hidden layers [25]. The basic processing unit in a neural network is called a neuron, and it performs two functions: collecting inputs and producing outputs. Mathematically, it is represented by a linear algebraic function, which is then taken as an argument to a transfer function f to form the output y, represented by:

$$y = f(x_1 x_2, , , , , x_n; w_{i1} w_{i2} w_{in}) \quad (4)$$

where: x_i - are inputs, w_{ij} - are parameters of the weight of layer *i* and y is the output.

Information and signals are transmitted between neurons via weighted connections, where the input x_j passes through a link, scaled by its weight w_{ij} to yield the product $x_j\, w_{ij}$, plus the neuron's bias b_i. This result serves as the argument for the activation or transfer function f. s_i is the net input to the neuron. which in turn gives the output y_i, represented by the following Eq.5:

$$y_i = f(s_i) \quad (5)$$

where:

$$s_i = \sum_{j=1}^{n} W_{ij} \times x_j + b_i \quad (6)$$

The multi-layer feed-forward neural network (MLF) using the backpropagation (BP) algorithm has always been the most frequently used ANN approach for predicting solar radiation [26].

Although a three-layer Feed-forward Neural Network (FFNN) is often sufficient for many prediction tasks [27], the architecture implemented in this study consists of: the input layer incorporates the meteorological and astronomical variables outlined in Table 1, supplemented by the solar radiation component as an additional input; 2 hidden layers; One output layer representing the predicted GHI.

The hidden layers employ the Scaled Exponential Linear Unit (SELU) activation function to enhance training stability and convergence performance. The output layer, in turn, utilizes a linear activation suited to continuous regression problems.

A Python-generated script was used to estimate global solar radiation in the city of Khenchela. The Fig. 3 shows the general structure of the model.

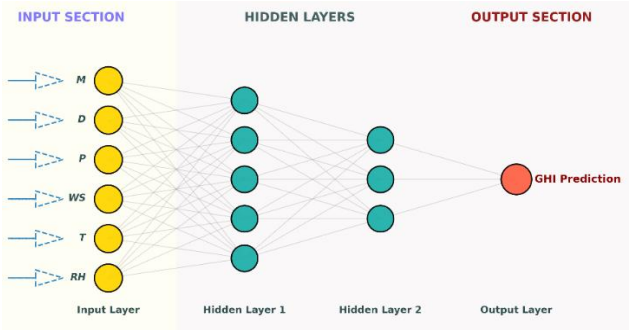

Fig. 3. ANN structure

2.2.2 Firefly Algorithm

Firefly algorithm is a metaheuristic algorithm based on population, inspired by the ideal behaviour of the flashing properties of fireflies, where the flashing of fireflies acts as a signalling system to attract other fireflies. Each possible solution is represented as a set of superior coefficients, and poorly performing solutions are gradually attracted towards better ones. The quality of each solution is evaluated based on an objective function built on the prediction error. The

coefficient of attraction between solutions can be modeled using the following relationship [28]:

$$B(r) = \beta_0 e^{-\gamma r^2} \quad (7)$$

where: β_0 - is the firefly attractiveness value at r = 0, γ - is the media light absorption coefficient, and r - is the distance between two solutions.

The solution positions (i.e., the model parameters) are updated iteratively according to the attraction mechanism, with the addition of a random element to ensure exploration of the search space:

$$x_i^{t+1} = x_i^t + \beta_0 e^{-\gamma r_{ij}^2}\left(x_j^t - x_i^t\right) + \alpha_t \epsilon_i^t \quad (8)$$

where: x_i^{t+1} and x_i^t is the current and updated positions of individual i, respectively. The term $(\beta_0 e^{-\gamma r_{ij}^2}(x_j^t - x_i^t))$ represents the attraction factor, α_t denotes the randomness coefficient, and $\boldsymbol{\epsilon_i^t}$ is a vector of random numbers drawn from a Gaussian distribution or a uniform distribution at time t. When β_0= 0, it reduces to a simple random walk process [28].

2.2.2.1. ANN–FFA Model

In this study, the Firefly Algorithm serves as the framework for optimizing the hyperparameters of the artificial neural network. Within this context, each firefly does not merely represent an abstract mathematical vector; rather, it embodies a specific structural configuration of the ANN, where each firefly corresponds to a candidate ANN configuration defined as:

$$x_i = [N_1, N_2, \eta]$$

where: N_1, N_2 - are the numbers of units in the first and second hidden layers, respectively, η - is the learning rate.

The search space is bounded as:

$$x_i \in [lb, ub]$$

with: $lb = [50{,}30, 0.0001]$ and $ub = [280, 180, 0.01]$

Mean Square Error (MSE) is defined as the fitness value of the solution:

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(Y_i - \bar{Y}_i)^2 \quad (9)$$

where: Y_i - is Measured global solar radiation (W/m²), $\bar{Y}_i$ - is Predicted value (W/m²).

Fireflies with inferior quality (higher MSE) are attracted toward those with superior quality (lower MSE). As outlined in Eq. (8), this process iterates over a fixed number of generations, progressively enhancing fitness in line with the characteristics detailed in Table 2.

Table 2. Firefly algorithm parameters

Parameter	Generations	Population Size	α	γ
Value	20	15	0.5	1

After the generations are over, the best firefly (best set of parameters) is chosen as the best solution (number of units in the first and second layers and learning rate), as these values represent the appropriate neural network structure for the inputs used, and these values are used in training the final model.

2.2.3 Particle Swarm Optimization (PSO)

Among the most famous metaheuristic algorithms, the PSO algorithm, inspired by the collective behaviour of flocks of birds and schools of fish, defines each particle as a candidate solution within a range of possible solutions. Each particle updates its position by adjusting its velocity based on its own best experience and the swarm's best position [29,30].

A particle in PSO can be defined as:

$$P_i \in [a, b]$$

where: (i = 1,2,3…. d) and (a,b) ∈ R, d - is for dimensions, R - is for real numbers.

Each particle has its own speed and position, which are assigned randomly at the beginning, p_{best} represents the best position that the particle should maintain, and g_{best} known as the global best position among particles [30].

To determine the position and velocity of the particle, the following equations are used:

$$V_i(t+1) = V_i(t) + C_1 r_1\left(p_{best} - n_i(b)\right) + C_2 r_2\left(g_{best} - x_i(t)\right) \quad (10)$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \quad (11)$$

where: V_i - represents the velocity of the particle, X_i- represents the position of the particle, r_1, r_2- are Random numbers between [1,2], C_1, C_2 - are the leaning factors.

2.2.3.1. ANN–PSO Model

According to the PSO model introduced above, the algorithm is used to search for optimal ANN hyperparameters, with each particle representing a candidate network configuration within a bounded domain. An initial population of particles $x_i \in [lb, ub]$ is randomly generated such that, including the batch size B:

$$x_i = [\mathrm{N}_1, \mathrm{N}_2, \eta, \mathrm{B}]$$

where:
$lb = [64, 64, 0.0001, 32]$ and $ub[256, 256, 0.01, 128]$.

Each solution (or particle) represents a candidate configuration of the model parameters: the number of units in the first hidden layer, the number of units in the second hidden layer, the learning rate, and the batch size, respectively.

The performance of each particle is evaluated using the MSE, whereby particles move towards the best discovered solutions in each iteration. Particle positions and velocities are now updated according to Eqs. (10) and (11), based on the best values and algorithm parameters listed in Table 3, until convergence to the optimal solution representing the minimum MSE value.

Table 3. PSO algorithm parameters

Parameter	Batch size	Swarm Size	Maximum Iterations (maxiter)
Value	[28,132]	20	30

Once the MSE is identified corresponding to the best-performing particle, which encodes the optimal values for the number of units in the first and second hidden layers, the learning rate, and the batch size the final ANN model is constructed and trained using these optimized hyperparameters obtained by the PSO algorithm.

2.2.4 The White Whale Optimization Algorithm (WWOA)

The White Whale Optimization Algorithm is an evolutionary optimization algorithm inspired by natural whale hunting behaviour. It is easy to implement and has few parameters, relying on spiralling around a target and gradually approaching it. The algorithm uses two Eqs. (12) and (13) to simulate whale movement in the solution space [31].

$$X_{t+1} = X_t + A(|C \times X^* - X_t|) \quad (12)$$

The Eq. (12) represents the motion in the solution space. Where are: X_t - represents the current solution position, X^* - represents the best solution position, *A, C* - is the transactions directed towards exploration and exploitation.

The whales' movement in a narrow circle around the optimal solution is expressed by Eq. (13), where this equation represents a mathematical model for bubble formation (bubble-net feeding) [31]:

$$X_{t+1} = X^* - A(|C \times X^* - X_t|) \quad (13)$$

The type of movement (exploratory or exploitative) in the white whale algorithm is determined by the adaptive control parameter A, which regulates the update step size and the direction of movement during the search process. This parameter is defined as follows:

$$A = 2a \times r - a;\ C = 2 \times r \quad (14)$$

and

$$a = 2 - \frac{2i}{M} \quad (15)$$

where: i - the current iteration number, r - is a random number of intervals [0, 1], M - indicates the maximum number of iterations. In both the exploitation and exploration phases, a is a linearly decreasing control parameter decreases linearly with the number of iterations from 2 to 0.

When |A|<1: the movement of search agents is directed toward the best solution obtained so far, thereby enhancing exploitative behavior. Conversely, when |A|≥1: this permits wider movements for the search agents, thereby promoting exploratory behavior [31,32].

2.2.4.1. ANN–WWOA Model

For the ANN hyperparameter optimization, each whale x_i represents a candidate configuration in the predefined search space $x_i \in [lb, ub]$.

An initial population of whales (i.e., random candidate solutions) is generated such that:
$lb = [64{,}64, 0.0001{,}32]$ and $ub[256, 256, 0.01{,}128]$

Each whale X now moves within the search space toward the best-known solution, guided by Eqs. (12) and (13). In this manner, the whale progressively advances toward the optimal neural network configuration that minimizes the MSE on the solar radiation dataset, while preserving solution diversity to avoid local minima. Solutions are updated according to the properties specified in Table 4.

Table 4. WWOA parameters

Parameter	Batch size	Population size	Maximum Iterations
Value	[32,128]	20	30

Once the optimal whale (best set of hyperparameters) is identified, the final artificial neural network model is trained using these parameters, which fall within the bounds lb and ub.

Fig. 4 illustrates the general framework of the optimization process, which delineates the

sequential steps commencing with data preparation and pre-processing, followed by the initialization of candidate solutions; the fitness of these solutions is then evaluated using the MSE objective function, after which they are updated according to each algorithm's specific mechanism; FFA relies on an attraction-based update mechanism, PSO on a velocity update, and WWOA on a spiral search. This process continues iteratively until the best solution, representing the optimal model structure, is selected.

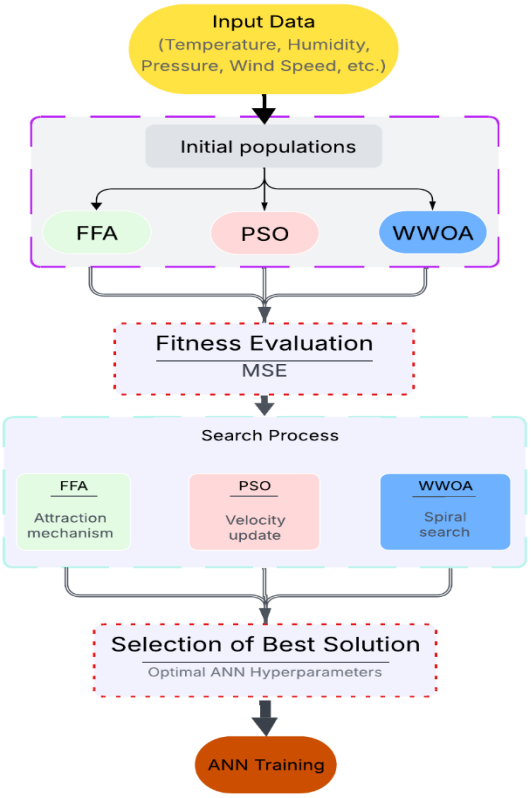

Fig. 4. Flowchart of the ANN optimization using metaheuristic algorithms

2.3 Assessment Indicators

Selecting the optimal predictive model is challenging due to the differing characteristics of each algorithm. Model selection relies on fundamental criteria such as ease of implementation, training efficiency, and overall predictive performance. In this study, the selection process was validated using graphical analyses alongside five comprehensive evaluation metrics: Mean Absolute Error (MAE), rMSE, MAPE, and coefficient of determination (R^2) [33]. The Scikit-learn library, a widely used Python package that supports the application of various machine learning models, allows for the calculation of the aforementioned statistical metrics. Lower values for MSE, MAE, RMSE, and MAPE indicate better predictive performance, while an R^2 value closer to 1 reflects higher model accuracy. Table 5 presents the mathematical formulas for these metrics [34].

Table 5. Statistical evaluation indicators used

Indices	Optimal value	Equation	
MAPE	0	$\frac{1}{m}\sum_{i=1}^{m}\left\|\frac{Y_i-X_i}{Y_i}\right\|$	(16)
RMSE	0	$\frac{1}{m}\sum_{i=1}^{m}(X_i-Y_i)^2$	(17)
MAE	0	$\frac{1}{m}\sum_{i=1}^{m}\|X_i-Y_i\|$	(18)
rRMSE	0	$\left(\frac{\mathrm{RMSE}}{\bar{Y}}\right)100$	(19)
R^2	1	$1-\frac{\sum_{i=1}^{m}(X_i-Y_i)^2}{\sum_{i=1}^{m}(\bar{Y}-Y_i)^2}$	(20)

3. RESULTS AND DISCUSSION

Enhanced ANN models were developed using multiple metaheuristic algorithms to achieve optimal structure, with the aim of predicting solar radiation using meteorological and astronomical data as constant inputs in all cases. The progressive incorporation of solar radiation components (DNI, DHI and GTI) was also adopted to evaluate the impact of each component on model performance and prediction accuracy. All models were evaluated using the statistical measures listed in Table 5.

In the first case, when using meteorological and astronomical data for the study site, it is clear that the models provided similar performance in all cases. There was a slight superiority of the ANN-WWOA model as it gave a balance between all indicators, where the results were for the ANN-WWOA model MAP =26.84%; rRMSE=20.79%; R^2=0.8831.

On the same input data, the ANN- PSO model gave the following results MAPE=28.29%; rRMSE=20.62%; R^2=0.8849. While the ANN-FFA model provided relatively lower performance compared to other models, as the results were MAPE=28.90%; rRMSE=21.68%; R^2=0.8728.

On the other hand, the ANN model yielded very poor results on the same dataset, with the estimated values being MAPE=44.44%; rRMSE=29.27%; R^2=0.7684.

Table 6 below summarizes the results of the studied models on the meteorological and astronomical dataset for the study site.

Table 6. Models result using only meteorological and astronomical data as inputs

Models	MAE (W/m^2)	rRMSE (%)	MAPE (%)	R^2
ANN	82.66	29.27	44.44	0.7684
ANN-FFA	59.28	21.68	28.90	0.8728
ANN-PSO	54.46	20.62	28.29	0.8848
ANN-WWOA	**53.71**	**20.79**	**26.84**	**0.8831**

It is evidently clear that optimization using the FFA, PSO, and WWOA enhances the accuracy of the horizontal global solar radiation estimation ANN model, where the coefficient of determination initially stands at R^2=0.7684, improving to a range between R^2=0.8728 and R^2=0.8831 post-optimization. The balance between the exploratory and exploitative characteristics of the WWOA gives it a slight advantage over other algorithms, with an estimated R^2 value of 0.8831. This balance also enables the algorithm to efficiently search the solution space and avoid falling into local minima, followed by a gradual focus on promising regions, where whale positions are intelligently updated towards the optimal solution. On the other hand, the PSO and FF Algorithms are simpler. The PSO algorithm tends to converge towards a solution without further optimization. In solar radiation prediction scenarios, strong correlations among input variables are not observed, which reduces the diversity of the PSO algorithm. In contrast, the FFA focuses on the brightness attraction mechanism, leading to an excessive narrowing of the exploration space, especially in solar radiation prediction problems characterized by complex, nonlinear, and multidimensional search spaces. This makes it fall into the trap of local optimum solutions. This behaviour does not reflect an inherent deficiency in the algorithm itself but rather highlights its relative limitations when applied to dynamic and nonlinear solar radiation forecasting tasks.

Table 7 below presents the optimal configuration of the ANN architecture obtained using the metaheuristic algorithms in the study on the dataset.

Table 7. Optimal ANN Parameters

Models	Layer 1	Layer 2	Learning rate	Batch size
ANN	128	64	0.001	64
ANN-FFA	178	69	0.0069	15
ANN-PSO	85	138	0.0087	36
ANN-WWOA	219	237	0.01	86

Table 7 shows a remarkable diversity in the ideal structure of an ANN. This reflects the varying capabilities of each model in representing the nonlinear relationships inherent in solar radiation prediction data. The results indicate that the WWOA algorithm tends towards a more complex structure in the hidden layers with a higher learning rate and larger batch size. reflecting a high representational capacity suitable for capturing complex relationships. This distinguishes the WWOA algorithm by its ease in exploring and exploiting the solution space. In contrast, the PSO algorithm exhibited an asymmetric distribution of nodes between the two layers, along with an average learning rate and an average batch size. This reflects the dynamic nature of the PSO algorithm in searching for solutions that achieve rapid convergence and an implicit orientation towards minimizing variance and reducing overlearning while maintaining acceptable accuracy. However, this often limits the comprehensive exploration of the solution space. On the other hand, the optimal structure obtained by the FF algorithm is characterized by a moderate number of nodes in the first layer and a relatively smaller number in the second layer, along with a low learning rate and a small batch size. This indicates a conservative learning behaviour resulting from the local brightness attraction mechanism, leading to an early convergence towards local solutions without sufficient exploration of deeper structures.

In general, these results indicate that structural variations in the model are not necessarily associated with improved predictive performance, but rather with the harmony among the network architecture, training parameters, and the nature of the data. This underscores the importance of hyperparameter optimization to achieve optimal predictive performance.

In the second part of the study, which concerns the relationship between the following three components of solar radiation (DNI, DHI and GTI) and GHI, the same data used in the first study were employed. One solar radiation component (DNI, DHI, or GTI) was added at a time, and the impact of this addition on improving the models' performance in predicting GHI was investigated. Table 8 presents the statistical results for each case.

Table 8. Models result for using meteorological and astronomical data with the addition of one component of solar radiation as inputs

Models	Components of solar radiation	MAE (W/m^2)	rRMSE (%)	MAPE (%)	R^2
ANN-FFA	DHI	43.38	17.24	25.55	0.9196
	DNI	24.07	9.69	13.07	0.9746
	GTI	**10.21**	**3.73**	**6.07**	**0.9962**
ANN-PSO	DHI	50.32	19.01	23.48	0.9022
	DNI	26.92	10.23	16.09	0.9717
	GTI	12.66	4.38	7.68	0.9948
ANN-WWOA	DHI	45.72	18.54	24.24	0.9071
	DNI	25.60	9.76	15.40	0.9742
	GTI	11.77	4.21	6.58	0.9952

The results in Table 8 indicate that all models performed well in predicting GHI. A comparison of their performance shows that the ANN-FFA model slightly outperforms the ANN-PSO and ANN-WWOA models across most statistical metrics.

The ANN-FFA model achieved the best results ever when adding the GTI component, as the results were MAPE=6.07%; rRMSE=3.73%; R^2=0.9962.

On the other hand, when adding the same component to the ANN-WWOA model, the results were as follows MAPE=6.58%; R^2=0.9952.

In this case, the ANN-PSO model performed slightly less than the other models, as the results were MAPE=7.68%; rRMSE=4.38%; R^2=0. 9948.

DNI is one of the main components of solar radiation and gave good results when added to the models, as the results for the ANN-FFA model were MAPE=13.07%; rRMSE=9.69%; R^2=0.9746.

The ANN-WWOA model also achieved similar performance MAPE=15.40%; rRMSE=9.76%; R^2=0. 9742.

As for the ANN-PSO model MAPE=16.09%; rRMSE=10.23%; R^2=0.9717.

When DHI was added, the results were not as accurate as desired compared to the other components in all models, as the ANN-FFA model gave relatively better results. The results were as follows MAPE=25.55%; rRMSE=17.24%; R^2=0.9196. The ANN-WWOA model also did not have very satisfactory results and achieved average performance, as MAPE=24.24%; R^2=0. 9071.

The results in Table 8 confirm that including an additional radiation component in the model inputs clearly contributed to improving predictive performance, indicating that this variable adds an important informational dimension that was not adequately represented in the original model formulation. This enabled the ANN-FFA model to quickly exploit information via the "brightness attraction" mechanism, making it particularly effective in complex, highly interconnected search spaces.

In contrast, the WWO algorithm showed less effectiveness when the optimization goal was precisely defined and the underlying relationships were clear. Its relative strength becomes apparent in ambiguous or nonlinear scenarios, as in the first case of this study, which relied entirely on climate data and had uncertain, complex input-output relationships. In such conditions, WWOA's broad and comprehensive search capability enables it to explore the difficult solution space with high efficiency, making it well-suited for complex and ambiguous optimization problems.

The PSO algorithm demonstrated consistently acceptable performance across both cases. It performed particularly well in the weather-only scenario, owing to its efficient information-sharing mechanism among particles, which supports effective exploration and rapid adaptation in environments with weakly defined or noisy relationships.

Finally, Figs. 5 and 6 display scatter plots of predicted versus actual GHI values. In these plots, blue markers represent the model predictions, while the red diagonal line indicates ideal one-to-one agreement between observed and estimated values.

Fig. 6 shows that the data points precisely match along a 1:1 ideal proportionality line, indicating a strong agreement between predicted and observed solar radiation values. This agreement is significantly enhanced by the addition of the GTI component, which improves model performance by reducing random scattering and systematic bias. The improved accuracy is due to the supplementary information provided by the GTI, which relates to the dynamic behaviour of solar radiation, particularly its interaction with an inclined surface under realistic atmospheric conditions. Consequently, the model becomes not only more accurate but also more physically consistent, underscoring the importance of incorporating field-based variables into solar radiation forecasting frameworks.

In contrast, Fig. 5 shows the ANN-WWOA model, which relies on a limited dataset and exhibits significantly weaker agreement, with numerous outliers that deviate markedly from the isotropy line. This demonstrates the inability of conventional weather data to fully describe the dynamics of solar radiation. The increased dispersion indicates that the input set lacks sufficient information to capture

the full variability of solar radiation, leading to a deterioration in predictive performance, particularly in critical situations and during sudden data fluctuations. This comparison clearly illustrates that incorporating direct solar radiation components, especially GTI, is an effective strategy for enhancing the reliability of solar radiation forecasting models.

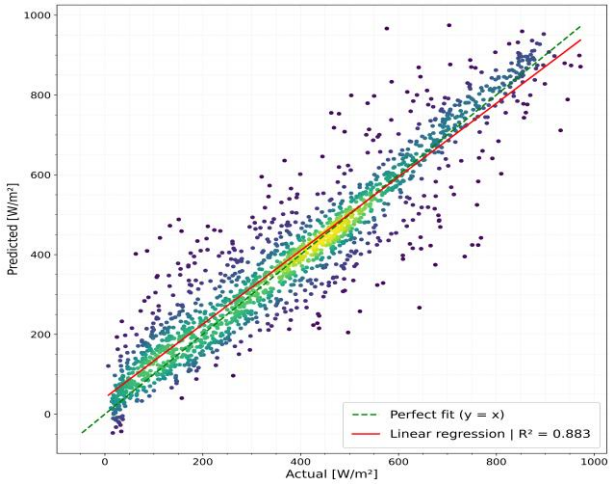

Fig. 5. Scatter plot for the ANN-WWOA model for meteorological and astronomical data

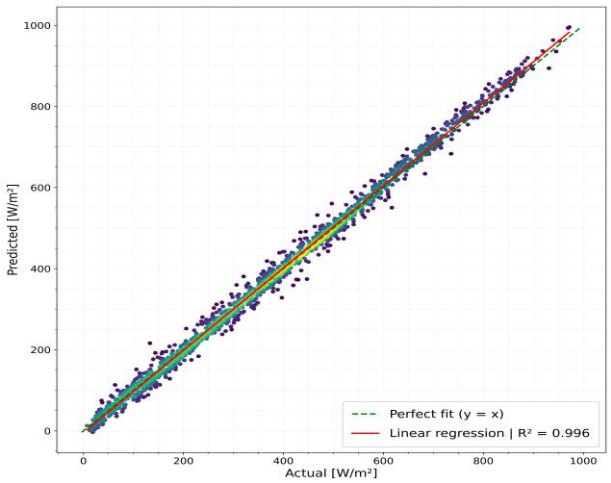

Fig. 6. Scatterplot for ANN-FFA model with GTI

To analyze the time behaviour of both actual and expected values, Figs. 7 and 8 illustrate the performance of the two models on a time series of data.

When the GTI component was incorporated as an additional input, the ANN-FFA model achieved excellent performance, with the actual values (blue line) largely matching the expected values (orange line) in most cases, as shown in Fig. 7, demonstrating its high predictive power across various time periods. In contrast, the ANN-WWOA model utilizing astronomical and meteorological data, as depicted in Fig. 8, exhibits significant discrepancies between actual and predicted values.

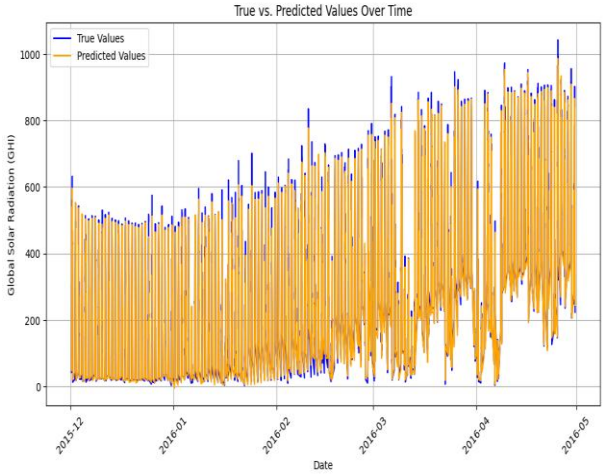

Fig. 7. Plot of true values and predictions for the ANN-FFA model with GTI

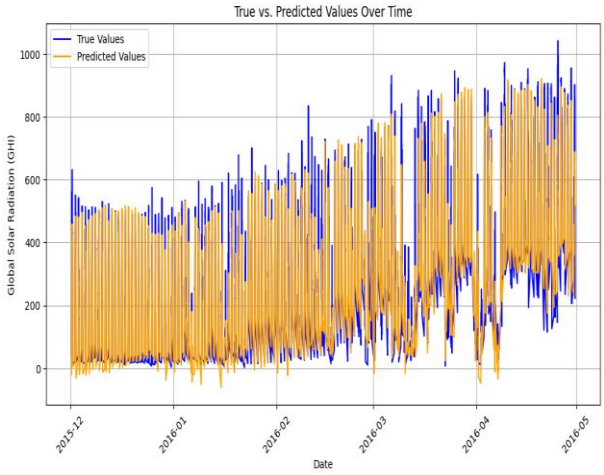

Fig. 8. Plot of true values and predictions for the ANN-WWOA model for meteorological and astronomical data

To understand the ability of the model to represent the statistical properties of the data, Figs. 9 and 10 represent the Density Plot of True and Predicted Values (Histogram).

Fig. 9 shows that the ANN-FFA model closely matches observed GHI values in the 100–500 W/m² range but struggles at low irradiance (<100 W/m²), where rapid transitions (e.g., sunrise or cloud edges) are hard to capture. The ANN-WWOA model performs weakest in the 0–200 W/m² range (Fig. 10), likely due to high atmospheric variability during morning/evening hours, yet achieves acceptable accuracy in the 200–600 W/m² daylight window. Performance degrades again at high irradiance (>600 W/m²), suggesting limited robustness under peak conditions. Overall, ANN-FFA demonstrates superior fidelity, especially in transitional regimes critical for real-time solar forecasting.

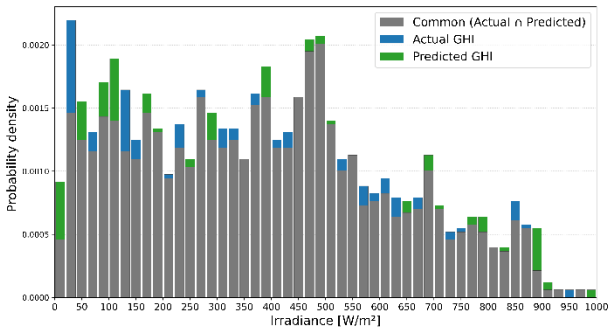

Fig. 9. The density plot of true and predicted values for ANN-FFA model with GTI

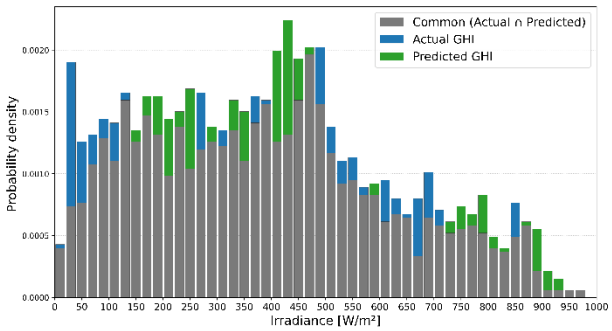

Fig. 10. The density plot of true and predicted values for the ANN-WWOA model for meteorological and astronomical data

4. CONCLUSION

This study investigated 10-minute ahead forecasting of GHI using ANNs optimized with three metaheuristic algorithms: the FFA, PSO, and WWOA. Using meteorological and astronomical data, the developed models were trained to identify the best ANN architectures for short-term solar irradiance forecasting. Furthermore, the relationship between GHI and its three main components, DNI, DHI, and GTI, was systematically examined using the developed models.

In the first scenario, the model was built using only meteorological and astronomical variables, without including measurements of direct solar radiation. The performance of the models was evaluated using a set of approved statistical measures. Results indicated limited predictive capability across all configurations. Remarkably, the ANN–WWOA model, among the best-performing in this scenario, achieved R^2=0.8831, reflecting substantial error and insufficient reliability for short-term forecasting. These findings underscore that meteorological and astronomical data alone are inadequate for accurate GHI prediction at high temporal resolution, particularly under rapidly fluctuating atmospheric conditions.

To examine the relationship between solar radiation components and GHI, the same models described previously were employed, with one solar radiation component introduced as an additional input at a time. This approach was adopted to assess the influence of each component on the accuracy of the results. Using the previous statistical measures, it was shown that the GTI component made a significant improvement and high accuracy in the results for the models, in which ANN-FFA slightly outperformed the ANN-WWOA and ANN-PSO models. The ANN-FFA model gave the following results: R^2=0.9962. These results proved that GTI Incline is the most solar radiation component that has a relationship with GHI. The other components of solar radiation also provided improvements, as DNI gave good results in all models. The results for the ANN-FFA model were estimated at R^2 = 0.9746, while DHI showed the least improvement for the studied models.

The results of the investigation indicate that predicting Global Horizontal Irradiance over short periods using machine learning models requires access to solar radiation components, especially GTI. The study also showed that improving the artificial neural network model using the FFA, PSO and WWOA showed effectiveness in improving the accuracy of the model. Based on this, future research should expand the scope by developing additional models and more advanced metaheuristic inference algorithms, and broaden the study to include additional climatic variables, with the aim of achieving more accurate and effective predictions. This path opens up broad prospects for more effective practical applications in analysing and optimizing solar energy resources.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] M. Ahmedbelbachir, Renewable Energies, Transition and Prospects: The Case of Algeria. *The Eurasia Proceedings of Educational and Social Sciences*, 32, 2023: 62–70. https://doi.org/10.55549/epess.1412817

[2] J. Skerlic, D. Nikolic, D. Cvetkovic, A. Miskovic, Optimal Position of Solar Collectors: A Review. *Applied Engineering Letters*, 3(4), 2018: 129–134. https://doi.org/10.18485/aeletters.2018.3.4.3

[3] J. Piri, S. Shamshirband, D. Petković, C.W. Tong, M.H. ur Rehman, Prediction of the solar radiation on the Earth using support vector

regression technique. *Infrared Physics & Technology*, 68, 2015: 179–185. https://doi.org/10.1016/j.infrared.2014.12.006

[4] L. Huang, J. Kang, M. Wan, L. Fang, C. Zhang, Z. Zeng, Solar radiation prediction using different machine learning algorithms and implications for extreme climate events. *Frontiers in Earth Science*, 9, 2021: 596860. https://doi.org/10.3389/feart.2021.596860

[5] M. Li, Z. Liao, C.F.M. Coimbra, Spectral solar irradiance on inclined surfaces: A fast Monte Carlo approach. *Journal of Renewable and Sustainable Energy*, 12(5), 2020: 053705. https://doi.org/10.1063/5.0011635

[6] K. Bouchouicha, N. Bailek, A. Razagui, M. El-Shimy, M. Bellaoui, N.E.I. Bachari, Comparison of artificial intelligence and empirical models for energy production estimation of 20 MWp solar photovoltaic plant at the Saharan Medium of Algeria. *International Journal of Energy Sector Management*, 15(1), 2021: 119–138. https://doi.org/10.1108/IJESM-12-2019-0017

[7] K. Barhmi, C. Heynen, S. Golroodbari, W. van Sark, A Review of Solar Forecasting Techniques and the Role of Artificial Intelligence. *Solar*, 4(1), 2024: 99–135. https://doi.org/10.3390/solar4010005

[8] C. Voyant, G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, A. Fouilloy, Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, 105, 2017: 569–582. https://doi.org/10.1016/j.renene.2016.12.095

[9] S. Zhao, L. Wu, Y. Xiang, J. Dong, Z. Li, X. Liu, Z. Tang, H. Wang, X. Wang, J. An, F. Zhang, Z. Li, Coupling meteorological stations data and satellite data for prediction of global solar radiation with machine learning models. *Renewable Energy*, 198, 2022: 1049–1064. https://doi.org/10.1016/j.renene.2022.08.111

[10] Y. Zhou, Y. Liu, D. Wang, X. Liu, Y. Wang, A review on global solar radiation prediction with machine learning models in a comprehensive perspective. *Energy Conversion and Management*, 235, 2021: 113960. https://doi.org/10.1016/j.enconman.2021.113960

[11] A. Al-rubaye, M.M. Al-Khuzaie, Artificial Neural Networks for Solar Radiation Prediction: Case Study, Al-Qadisiyah, Iraq. *The Journal of Engineering Research*, 20(2), 2023: 106–112. https://doi.org/10.53540/tjer.vol20iss2pp106-112

[12] N.D. Kaushika, R.K. Tomar, S.C. Kaushik, Artificial neural network model based on interrelationship of direct, diffuse and global solar radiations. *Solar Energy*, 103, 2014: 327–342. https://doi.org/10.1016/j.solener.2014.02.015

[13] Ü. Ağbulut, A.E. Gürel, Y. Biçen, Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison. *Renewable and Sustainable Energy Reviews*, 135, 2021: 110114. https://doi.org/10.1016/j.rser.2020.110114

[14] E.S. Solano, C.M. Affonso, Solar Irradiation Forecasting Using Ensemble Voting Based on Machine Learning Algorithms. *Sustainability*, 15(10), 2023: 7943. https://doi.org/10.3390/su15107943

[15] F.P. Marinho, P.A.C. Rocha, A.R.R. Neto, F.D.V. Bezerra, Short-Term Solar Irradiance Forecasting Using CNN-1D, LSTM, and CNN-LSTM Deep Neural Networks: A Case Study With the Folsom (USA) Dataset. *Journal of Solar Energy Engineering*, 145(4), 2023: 041002. https://doi.org/10.1115/1.4056122

[16] C. Vanlalchhuanawmi, S. Deb, M.M. Islam, T.S. Ustun, Solar radiation prediction: A multi-model machine learning and deep learning approach. *AIP Advances*, 15(5), 2025: 055201. https://doi.org/10.1063/5.0237246

[17] B. Zerouali, N. Bailek, S. Qaysi, S. Difi, N. Alarifi, A. Elbeltagi, C.A.G. Santos, K. He, Y.M. Youssef, Hybrid machine learning optimization for solar radiation forecasting. *Physics and Chemistry of the Earth, Parts A/B/C*, 140, 2025: 104052. https://doi.org/10.1016/j.pce.2025.104052

[18] Z. Bounoua, L.O. Chahidi, A. Mechaqrane, Estimation of daily global solar radiation using empirical and machine-learning methods: A case study of five Moroccan locations. *Sustainable Materials and Technologies*, 28, 2021: e00261. https://doi.org/10.1016/j.susmat.2021.e00261

[19] E.L.H.I.E.I. Soukeur, D. Chaabane, K. Amarouche, N.E.I. Bachari, Creating an artificial neural network time series model for

the prediction of daily solar radiation in Oran. *Desalination and Water Treatment*, 255, 2022: 163–171. https://doi.org/10.5004/dwt.2022.28337

[20] D. Benatiallah, A. Zahouani, A. Yahiaoui, H. Djeldjli, B. Nasri, A. Benatiallah, B. Benabdelkrim, Estimating of solar radiation based on machine learning approaches under Algerian desert climate. *Instrumentation Mesure Métrologie*, 23(5), 2024: 363–374. https://doi.org/10.18280/i2m.230504

[21] D. Halima, B. Djelloul, G. Mehdi, T. Camel, B. Ali, B. Bouchra, Solar Radiation Estimation Based on a New Combined Approach of Artificial Neural Networks (ANN) and Genetic Algorithms (GA) in South Algeria. *Computers, Materials and Continua*, 79(3), 2024: 4725–4740. https://doi.org/10.32604/cmc.2024.051002

[22] M.R. Yaiche, A. Bouhanik, S.M.A. Bekkouche, A. Malek, T. Benouaz, Revised solar maps of Algeria based on sunshine duration. *Energy Conversion and Management*, 82, 2014: 114–123. https://doi.org/10.1016/j.enconman.2014.02.063

[23] D.S. Kumar, G.M. Yagli, M. Kashyap, D. Srinivasan, Solar irradiance resource and forecasting: a comprehensive review. *IET Renewable Power Generation,* 14(10), 2020: 1641–1656. https://doi.org/10.1049/iet-rpg.2019.1227

[24] C.A. Gueymard, Direct and indirect uncertainties in the prediction of tilted irradiance for solar engineering applications. *Solar Energy*, 83(3), 2009: 432–444. https://doi.org/10.1016/j.solener.2008.11.004

[25] R. Lippmann, An introduction to computing with neural nets. *IEEE ASSP Magazine*, 4(2), 1987: 4–22. https://doi.org/10.1109/MASSP.1987.1165576

[26] A.K. Yadav, S.S. Chandel, Solar radiation prediction using Artificial Neural Network techniques: A review. *Renewable and Sustainable Energy Reviews*, 33, 2014: 772–781. https://doi.org/10.1016/j.rser.2013.08.055

[27] E.P.P.A. Derks, M.S. Sánchez Pastor, L.M.C. Buydens, Robustness analysis of radial basis function and multilayer feed-forward neural network models. *Chemometrics and Intelligent Laboratory Systems*, 28(1), 1995: 49–60. https://doi.org/10.1016/0169-7439(95)80039-C

[28] I. Fister, I. Fister, X.-S. Yang, J. Brest, A comprehensive review of firefly algorithms. *Swarm and Evolutionary Computation*, 13, 2013: 34–46. https://doi.org/10.1016/j.swevo.2013.06.001

[29] M. Imran, R. Hashim, N.E.A. Khalid, An Overview of Particle Swarm Optimization Variants. *Procedia Engineering*, 53, 2013: 491–496. https://doi.org/10.1016/j.proeng.2013.02.063

[30] A.G. Gad, Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review. *Archives of Computational Methods in Engineering*, 29, 2022: 2531–2561. https://doi.org/10.1007/s11831-021-09694-4

[31] S. Mirjalili, A. Lewis, The Whale Optimization Algorithm. *Advances in Engineering Software*, 95, 2016: 51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008

[32] J. Wei, Y. Gu, Y. Yan, Z. Li, B. Lu, S. Pan, N. Cheong, LSEWOA: An Enhanced Whale Optimization Algorithm with Multi-Strategy for Numerical and Engineering Design Optimization Problems. *Sensors*, 25(7), 2025: 2504. https://doi.org/10.3390/s25072054

[33] V. Gayathry, D. Kaliyaperumal, S.R. Salkuti, Seasonal solar irradiance forecasting using artificial intelligence techniques with uncertainty analysis. *Scientific Reports*, 14, 2024: 17945. https://doi.org/10.1038/s41598-024-68531-3

[34] M. Salama, Optimization of Regression Models Using Machine Learning: A Comprehensive Study with Scikit-learn. *IUSRJ International Uni-Scientific Research Journal*, 5(16), 2024: 119–129. https://doi.org/10.59271/s45500.024.0624.16

© 2026 by the authors. This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License (CC BY-NC 4.0).