International Journal of Robotics and Control Systems
Vol. 5, No. 1, 2025, pp. 179-196
ISSN 2775-2658
http://pubs2.ascee.org/index.php/ijrcs

Parametric Analysis of Climate Factors for Monthly Weather Prediction in Ghardaïa District Using Machine Learning-Based Approach: ANN-MLPs

Abdennasser Dahmani [a,b,1,*], Yamina Ammi [b,2], Kouidri Ikram [c,3], Sofaine Kherrour [d,4], Salah Hanini [b,5], Raheem Al-Sabur [e,6], Maamar Laidi [b,7], Alfian Ma'arif [f,g,8], Abdel-Nasser Sharkawy [h,i,9,*]

[a] Department of Mechanical Engineering, Faculty of Sciences and Applied Sciences, University of Bouira, Bouira 10000, Algeria
[b] Laboratory of Biomaterials and Transport Phenomena (LBMPT), University of Medea, Urban Pole, 26000, Medea, Algeria
[c] Department of Mechanical Engineering, Faculty of Science and Technology, GIDD Industrial Engineering and Sustainable Development Laboratory, University of Relizane, Bourmadia, 48 000, Relizane, Algeria
[d] Applied Research Unit in Renewable Energies, URAER, Centre for the Development of Renewable Energies, CDER, 47133, Ghardaïa, Algeria
[e] Department of Mechanical Engineering, College of Engineering, University of Basrah, Basrah 61001, Iraq
[f] Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta 55191, Indonesia
[g] Department of Engineering Profession, Universitas Muhammadiyah Yogyakarta, Yogyakarta 55183, Indonesia
[h] Mechanical Engineering Department, Faculty of Engineering, South Valley University, Qena 83523, Egypt
[i] Mechanical Engineering Department, College of Engineering, Fahad Bin Sultan University, Tabuk 47721, Saudi Arabia
[1] dahmaniabdennasser@gmail.com; [2] ammi.yamina@yahoo.fr; [3] ikramkouidri295@gmail.com;
[4] sofianekherrour@gmail.com; [5] s_hanini2002@yahoo.fr; [6] raheem.musawel@uobasrah.edu.iq; [7] maamarw@yahoo.fr;
[8] alfianmaarif@ee.uad.ac.id; [9] abdelnassersharkawy@eng.svu.edu.eg
* Corresponding Authors

ARTICLE INFO

Article history
Received September 29, 2024
Revised November 06, 2024
Accepted December 10, 2024

Keywords
Ghardaïa District;
Earth-Sun Distance;
Weather Prediction Model;
Solar Radiation;
Machine Learning

ABSTRACT

In the rapidly developing field of smart cities, accurately predicting weather conditions plays a vital role in various sectors, including industry, tourism, agriculture, social planning, architecture, and economic development. Unfortunately, the instruments used (such as pyranometers, barometers, and thermometers) often suffer from low accuracy, high computational costs, and a lack of robustness. This limitation affects the reliability of weather predictions and their application across these critical areas. This study proposes artificial neural network-multilayer perceptrons (ANN-MLPs). A dataset of 480 data points was used, with 80% allocated for the training phase, 10% for the validation phase, and 10% for the testing phase. The best results were obtained with the structure 6-17-1 (6 inputs, 17 hidden neurons, and 1 output neuron) to predict weather condition data in the Ghardaïa district. Weather conditions parameters include air temperature, relative humidity, wind speed, and cumulative precipitation. Results showed that the most relevant input factors are, in order of importance: earth-sun distance ($D_{T\text{-}S}$) with a relative importance (RI) of 31.10%, factor conversion (δ) with an RI of 26.05%, and solar radiation (SR) with an RI of 16.26%. The contribution of the elevation of the sun (H_l) has an RI of 13.29%. The optimal configuration includes seventeen neurons in the hidden layer with a logistic sigmoid activation function and a Levenberg–Marquardt learning algorithm, resulting in a root mean square error (RMSE) of 3.3043% and a correlation coefficient (R) of 0.9683. The proposed model can predict both short- and long-term climate factors such as solar radiation, air temperature, and wind energy in areas with similar conditions.

1. Introduction

Weather predicting involves the identification and prediction of climatic conditions with a certain degree of accuracy using various technologies. Many real-time systems rely on weather conditions to make necessary adjustments in their operations. Accurate predicting enables preventive measures to safeguard life and property from potential damage. Quantitative forecasts, such as temperature, humidity, and rainfall, are crucial in the agricultural sector and for traders in commodity markets. Utility companies use temperature forecasts to predict energy demand in the coming days. Since outdoor activities can be significantly impacted by heavy rain, snow, and cold weather, forecasts are essential for planning activities around these events and preparing in advance to manage and endure them [1], [2].

Many regions in Algeria face significant challenges in obtaining climate factors measurements due to the high costs of measurement equipment (such as pyranometers and solarimeters) and the complexities of maintenance and calibration. Although there are some meteorological stations throughout the country, measurements are often unavailable continuously due to power outages, especially during the summer, or due to limited recorded variables. Therefore, it is essential to develop effective methods for predicting climate factors using more readily available meteorological data. To achieve this, we utilize proven techniques for assessing climate factors components, such as empirical modeling and intelligent techniques, including artificial neural networks [3].

Nowadays, a variety of computing techniques are available to enhance predicting accuracy. These methods include the Linear Regression, Exponential Smoothing, Decision Trees, Moving Averages, Statistical Clustering, Time series methods, and Artificial intelligence-based techniques [4].

Weather predicting necessitates advanced computing techniques that can analyze nonlinear data and develop rules and patterns to derive insights from historical data for predicting future conditions [5]. Utilizing Artificial Neural Networks - Multilayer Perceptrons (ANN-MLPs) can yield more accurate results. Although the error may not be eliminated, accuracy is expected to improve compared to previous predictions [6].

Weather predicting is a dynamic process where model outputs may be required for daily, weekly, or monthly weather planning [7]. Therefore, accuracy is a critical component of this prediction. Various factors that can enhance accuracy are discussed.

The main objective of this research is to develop a method for optimizing the hyperparameters of traditional machine learning models using multilayer perceptrons (ANN-MLPs), thereby increasing the reliability of monthly weather condition predictions. We used ANN-MLPs models to produce reliable one-month interval weather predictions for the Ghardaïa station in Algeria. To the best of our knowledge, no studies in the literature have simultaneously predicted the four weather conditions examined in this study: cumulative precipitation, wind speed, relative humidity, and average temperature. The following overview outlines the framework for this research: Section 2 presents some literature review, Section 3 discusses the materials and methods and covers the construction of the model, and Section 4 presents the results and discussion. The paper concludes with a summary of the findings as presented in Section 5.

2. Literature Survey

This section provides an overview of various weather predicting techniques that use neural networks. Pang et al. [8] developed ANN and RNN models to assess the effectiveness of Deep Learning algorithms in predicting solar radiation using meteorological data from an Alabama station. The study found that the RNN model performed better than the ANN model. The accuracy of predictions can be further improved by increasing data granularity or applying a moving-window algorithm. Additionally, cloud cover was identified as a significant factor influencing prediction accuracy. Bailek et al. [9] developed a new Angström-Prescott model for predicting global solar radiation on a horizontal surface in the southwest region of Algeria. This model is compared with

other existing models using sunshine duration data, with the aim of improving prediction accuracy for solar energy applications in the region. Amiri et al. [10] The study suggests using the "Weights" method to evaluate the significance of input features for predicting solar irradiation through Artificial Neural Networks (ANNs). Satellite data was used for training, while ground data from Bouzaréah station was used for testing. The method helped rank input variables by their importance, leading to better model performance by excluding less relevant parameters. Dahmani et al. [11] examine the development of AI models for predicting hourly global solar irradiation. The models were optimized using the BFGS quasi-Newton algorithm and statistica software, with data sourced from two Algerian stations in different climatic zones. Accuracy was evaluated using metrics like the correlation coefficient, mean absolute error, and RMSE. This research highlights the importance of accurate prediction models in improving renewable energy management and planning. Sharkawy et al. [12] used a multilayer feedforward neural network (MLFFNN) to predict the output power of a solar PV power station in Egypt, using module temperature and solar radiation as inputs. Data from five days was used for training, and a sixth day for validating the trained model. The results demonstrated the accuracy and efficiency of the model in predicting power, with the LM algorithm being slightly more effective than the EBP algorithm.

Afzali et al. [13] developed Artificial Neural Networks (ANNs) to predict daily and monthly ambient air temperatures in Kerman city, Iran. Historical data from 1961 to 2004, including mean, minimum, and maximum temperatures, were used as inputs in Feed Forward and Elman Networks. The analysis showed that the ANN approach is effective for temperature prediction, with the Elman network providing more accurate results for one-day-ahead mean temperature and one-month-ahead maximum temperature predictions. Trang et al. [14] detailed review of artificial neural network (ANN) methods, such as recurrent neural networks (RNN) and long short-term memory (LSTM), for forecasting air temperature, covering research from 2005 to 2020. The review emphasizes that while ANN models are effective and popular for their speed and capability in managing complex problems, there is no agreed-upon best method. These approaches are particularly suited for short-term temperature predicting. Samer AlSadi and Tamer Khatib [15] The study explores predicting relative humidity using a feedforward artificial neural network (FFNN). It utilizes weather data from Malaysia for training the FFNN, focusing on predicting relative humidity based on sunshine ratio and cloud cover. The evaluation of the neural network involves statistical. The proposed model demonstrates strong accuracy in hourly relative humidity prediction.

Kuzugudenli [16] focuses on predicting relative humidity, an important climate parameter, using annual total precipitation, average ambient temperature, and altitude. The study developed both regression and artificial intelligence models based on data from 177 meteorological stations across Turkey. It found that the artificial neural network model offered superior predictive accuracy compared to the multiple linear regression model. These models are suggested to be valuable for similar climate conditions. Muhammad et al. [17] focused on using artificial neural networks (ANNs) to predict wind speed by optimizing network topology, specifically the number of hidden layers and neurons. The topology was determined using principal component analysis (PCA) for hidden layers and K-means clustering for neuron selection. The chosen model outperformed other methods in wind speed prediction. Sharkawy et al. [18] applied Multi-input single-output neural network models were designed and evaluated for estimating output power in wind turbine farms. Three types of artificial neural networks (MLFFNN, CFNN, and RNN) were used to estimate the total power of wind turbines in Egypt, using wind speed, surface temperature, and pressure as inputs. The results showed that the RNN was the most accurate with the lowest mean square error (MSE) compared to the other methods. Zhenhao et al. [19] proposed a two-phase deep learning model for effective short-term wind direction forecasting. In the first phase, a hybrid data processing strategy involving data reconstruction, outlier deletion, dimension reduction, and sequence decomposition is employed to extract key information. In the second phase, an optimized echo state network is developed for accurate wind direction prediction, with hyper-parameters fine-tuned using an improved flower pollination algorithm (IFPA). The model's effectiveness is validated through experiments using real wind farm data. Ahcene Bouach [20] presents a forecasting tool using artificial neural networks (ANNs) for predicting monthly

precipitation over a 12-month period. The study evaluates two normalization methods (ANN-SS and ANN-MM) and four approaches for selecting input variables (no selection, ANN-WO, ANN-CO, and ANN-VE) to enhance model performance, with a particular focus on medium-term forecasting, an area often overlooked in prior studies.

Sharadqah et al. [21] used artificial neural networks for predicting rainfall in Jordan county, employing the training algorithm Levenberg-Marquardt. The network structural design included three layers (input, hidden, layer). The study reported that the model achieved highly satisfactory results with minimal errors. Ghamariadyan and Imteaz [22] developed a predicting model using a wavelet-aided artificial neural network (WANN) for predicting rainfall in Australia. A comparative analysis with classical models like ANN, ARIMA, MLR, and the Australian Community Climate Earth-System Simulator–Seasonal (ACCESS-S) showed that the WANN model offered superior accuracy. Abhishek et al. [23] discuss weather forecasting using Artificial Neural Networks (ANNs), highlighting the growing importance of accurate predictions across various sectors. The authors present a model that employs ANN techniques to enhance forecasting accuracy. Their study demonstrates that neural networks, with their ability to recognize complex patterns in large datasets, can effectively predict weather conditions by learning from historical data. The findings suggest that ANN-based models have significant potential for improving the precision of weather forecasts.

This paper aims to examine and analyze the weather patterns in Ghardaïa, Algeria, and propose predicting models that can accurately predict future weather data. The study seeks to develop models that assist decision-makers in taking appropriate measures to address environmental concerns and manage related demands effectively.

3. Materials and Methods

3.1. Study Region and Climatic Conditions

This study focuses on the city of Ghardaïa, located in the south of Algeria, approximately 600 km from the capital, Algiers. Ghardaïa covers an area of 19,729 km² and is situated at a latitude of +32.37°, a longitude of +3.77°, and an altitude of 450 meters above sea level. The region features a climate that is a mix of arid and semi-arid conditions, with mild winters and hot summers. Summer temperatures range from 14°C to 47°C, while winter temperatures vary from 2°C to 37°C. Daily solar energy in Ghardaïa fluctuates between a 607 Wh/m²/day and a 7,574 Wh/m²/day, with an annual average of approximately 5,656 Wh/m²/day [24]. The geographical location of the study area is illustrated on the map of Algeria (Fig. 1).

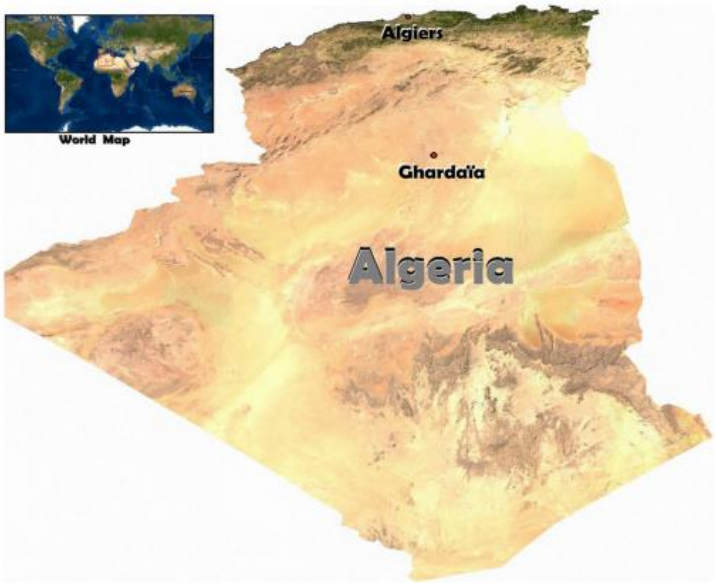

Fig. 1. The map of Ghardaïa, Algeria, [25]

3.2. Description Database Collection and Pre-Processing

The database used in this study includes average monthly data on air temperature, relative humidity, wind speed, global solar radiation, and cumulative precipitation, collected by the Applied Research Unit for Renewable Energies (URAER) in Ghardaïa. The data spans a 10-year period from January 2010 to December 2019. Instead of using the numbers 1 to 12 to represent the months, it is more effective to work with specific phenomenological parameters. These include the monthly earth-sun distance, denoted as $(D_{T\text{-}S})_1$, $(D_{T\text{-}S})_2$, ..., $(D_{T\text{-}S})_{12}$, (Fig. 2), and the elevation of the Sun at noon (solar) during the monthly solstices and equinoxes, denoted as H_1, H_2, H_3, ..., H_{12}, for each season (inputs 1 and 2).

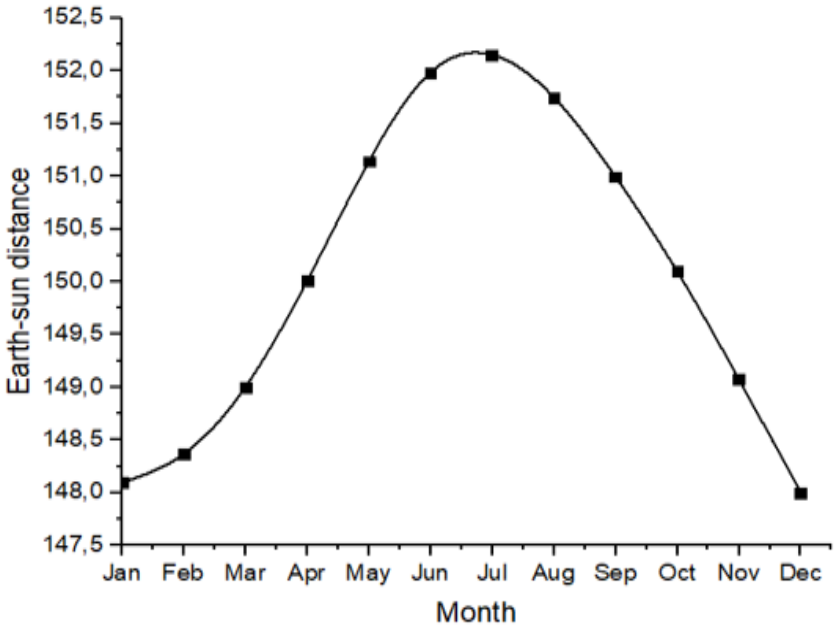

Fig. 2. Distance earth-sun for each month, [26]

At the spring and autumn equinoxes (March 21 and September 21), solar radiation at noon is perpendicular to the equator (latitude 0°) [27], [28]. During these times, days and nights are of equal duration everywhere on the globe (Fig. 3 a). This makes it the easiest time to calculate the Sun's height at noon, as it is equal to the complementary angle of the latitude (L°). This can be calculated as follows:

$$H = 90° - L° \tag{1}$$

At the summer solstice (June 21), the Earth is tilted towards the Sun, and at noon, the Sun's rays are perpendicular to the Tropic of Cancer (latitude 23°27 N). In regions within the Arctic Circle (latitude 66°33'N), the Sun never sets. A person living at this latitude (23°27 below the North Pole). A person living at latitude 66°33 N (90°-23°27) would see the Sun move around the north, descend to touch the horizon at midnight, and then rise again in the eastern sky. The height of the Sun at noon during the summer solstice is 23°27' higher than at the equinoxes [29], [30] (Fig. 3 b). This can be calculated as follows:

$$H = 90° - L° + 23°27 = 90° - L° + 23.45 \tag{2}$$

At the winter solstice (December 22), the Earth's tilt is reversed, and the Tropic of Capricorn (latitude 23°27'S) receives perpendicular solar radiation [31], [32]. The height of the Sun at noon during the winter solstice is 23°27' lower than at the equinoxes (Fig. 3 c). This can be calculated as follows:

$$H = 90° - L° - 23°27' = 90° - L° - 23.45 \tag{3}$$

For years: It is also advantageous to work with phenomenological parameters relative to the mechanisms of seasonal climatic variations (input 3 and 4): each year will be characterized by 4 pairs (J* and H*)i with {J* = Normalized days =Day/max (day) via month and H* = Normalized hours = Hour/24 hours} $(J^* \text{ and } H^*)_{\text{March equinox}}$, $(J^* \text{ and } H^*)_{\text{June solstice}}$, $(J^* \text{ and } H^*)_{\text{September equinox}}$, $(J^* \text{ and } H^*)_{\text{December solstice}}$.

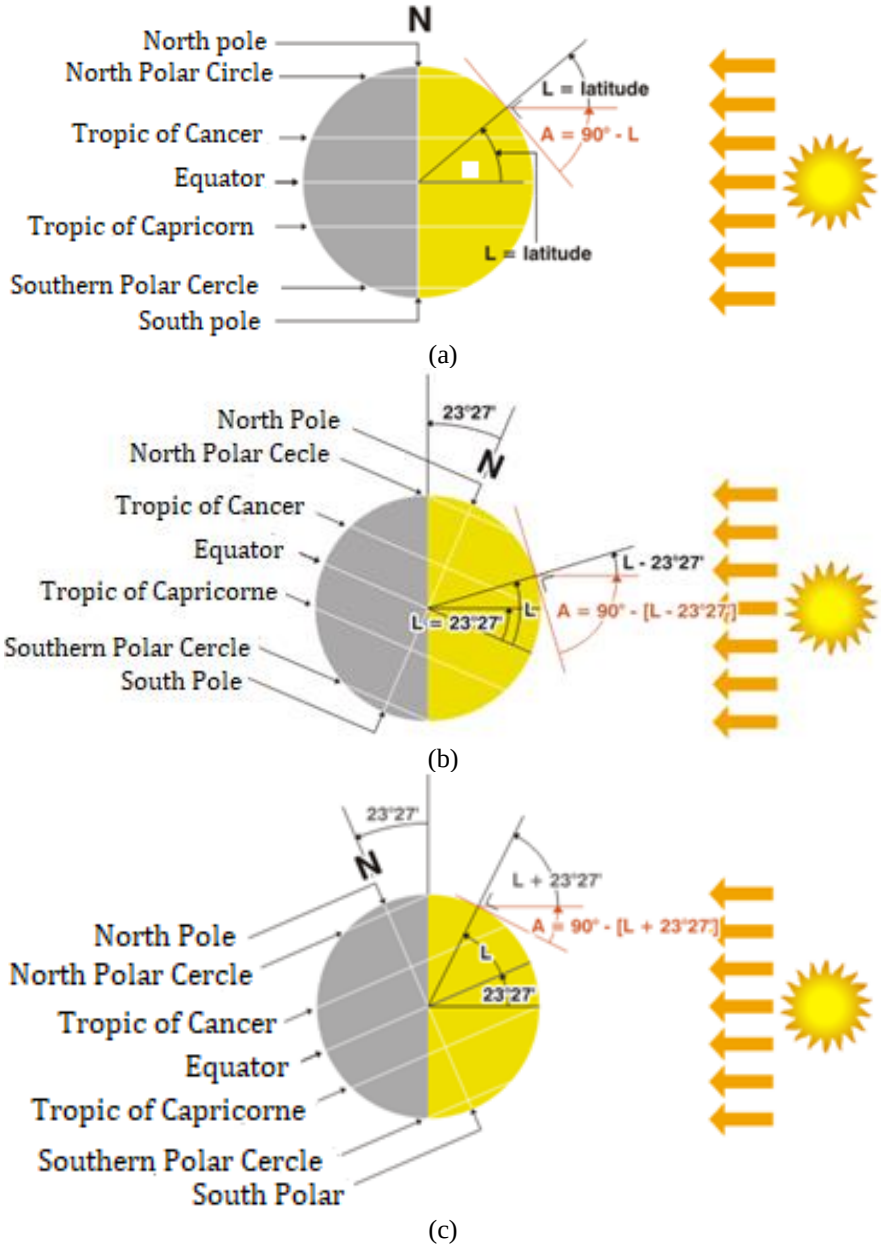

Fig. 3. Elevation of sun earth: a) spring and autumn equinoxes, b), summer solstice c), winter solstice

3.3. Database Normalization

The database consists of 480 experimental data points used to predict variations in weather conditions. It includes six inputs: the distance between the Earth and the Sun for each month, the annual elevation, days, hours, global solar radiation, and the conversion factor. These inputs are normalized as follows:

- The normalized distance between the Earth and the Sun = (distance between the Earth and the Sun / 152,000,000) = (distance between the Earth and the Sun / maximum distance between the Earth and the Sun).

- The normalized annual elevation of the sun = (annual evolution / 77.1858) = (annual elevation / maximum annual elevation).
- The normalized month is characterized by two inputs: J* (normalized days) = number of days / maximum number of days in a month, and H* (normalized hours) = hour / 24 hours.
- The normalized conversion factor (δ) = (conversion factor / 4) = (conversion factor / maximum conversion factor), with δ = 0.25 for average air temperature, δ = 0.5 for relative humidity, δ = 0.75 for cumulative precipitation, and δ = 1 for wind speed.
- The normalized weather variation = (weather variation for each parameter / maximum weather variation for each parameter).

The statistical analysis of the total data includes the Domain, average ("mean"), and standard deviation ("SD"), as shown in Table 1.

Table 1. Statistical analysis of normalized inputs and output data

Inputs and Outputs of NN		Symbol	Domain	Mean	SD
Inputs	Earth-Sun-Distance	(DT-S) I	0.9666 - 1.0000	0.9851	0.0101
	elevation of the sun	HI	0.3924 - 1.0000	0.6858	0.2066
	Day	D*	0.6452 - 0.7666	0.6951	0.0365
	Hour	H*	0.0792 - 0.9850	0.5283	0.2392
	Solar radiation	SR	0.3820 - 1.0000	0.7284	0.1977
	Conversion factor	δ	0.2500 - 1.0000	0.6250	0.2798
Output	Weather conditions	WC	0.0000 - 1.0000	0.5820	0.3134

3.4. Artificial Neural Networks (ANNs) Based Approach

Artificial neural networks (ANNs) represent a sophisticated paradigm inspired by the human brain's information processing capabilities. These networks excel in modeling complex, nonlinear systems that defy straightforward analytical representation. ANNs leverage their ability to adapt and learn from environmental inputs, making them invaluable for tasks where traditional mathematical models fall short [33]–[35]. Central to ANNs is their multilayer structure, typically comprising input, hidden, and output layers. Each layer houses neurons that process incoming signals through weighted connections, adjusted iteratively during training to minimize prediction errors. The transfer functions, such as tangent sigmoid (tansig), logarithmic sigmoid (logsig), sinusoidal (sin), and exponential, applied in hidden layers, enhance their capability to capture intricate relationships within data [36]–[38].

Training ANNs involves optimizing parameters like learning algorithms and network topology, crucial for achieving accurate predictions. The process often includes fine-tuning the number of hidden neurons and selecting optimal division subsets for training and testing phases. This iterative refinement, guided by trial and error, culminates in an optimized ANN model capable of predicting outcomes with high precision. The Multilayer Perceptron (MLP) is the predominant architecture, featuring a straightforward structure comprising three layers: Firstly, the input layer receives and processes input data. Secondly, the hidden layer(s) undertakes information processing from the input layer. Lastly, the output layer generates the final model output [39], [40]. The modeling process entailed designing and fine-tuning the neural network architecture according to the steps illustrated in Fig. 4.

This study used various error metrics to evaluate prediction model accuracy, including the Correlation Coefficient (R), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Standard Error of Prediction (SEP). The Correlation Coefficient (R) assesses the linear relationship between predicted and actual values, while MAE measures the average absolute error. RMSE evaluates the average magnitude of errors, emphasizing larger discrepancies, and SEP indicates prediction precision around the regression line. These metrics together offer a comprehensive view of model performance [41]–[43].

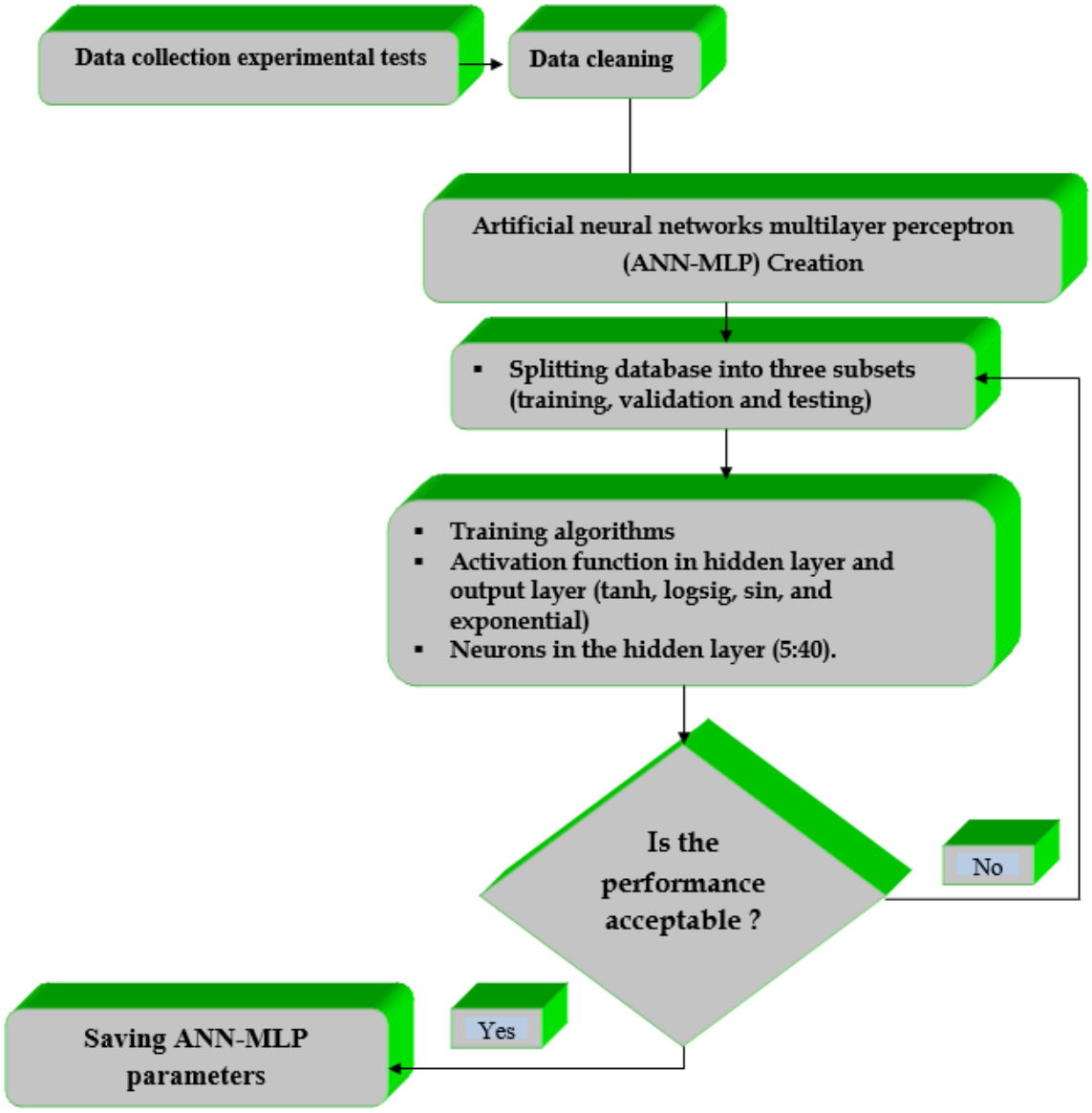

Fig. 4. Flow diagram for ANN-MLPs development

$$R = \frac{\sum_{i=1}^{n} \left(Y_{i,exp} - \overline{Y_{i,exp}}\right)\left(Y_{i,cal} - \overline{Y_{i,cal}}\right)}{\sqrt{\sum_{i=1}^{n} \left(Y_{i,exp} - \overline{Y_{i,exp}}\right)^2 \sum_{i=1}^{n} \left(Y_{i,cal} - \overline{Y_{i,cal}}\right)^2}} \quad (4)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \left|Y_{i,cal} - Y_{i,exp}\right| \quad (5)$$

$$RMSE(\%) = \sqrt{\frac{\sum_{i=1}^{n} \left(Y_{i,cal} - Y_{i,exp}\right)^2}{n}} \quad (6)$$

$$SEP(\%) = \frac{RMSE}{Y_e} \times 100 \quad (7)$$

Where, n is the number of data points, $Y_{i,exp}$ the experimental data points of weather conditions, $Y_{i,cal}$ the calculated data points of weather conditions, and $\overline{Y_{i,exp}}$ the mean experimental data.

4. Results and Discussion

4.1. Experimental Results

A learning algorithm is a method used to adjust network coefficients (weights and biases) to minimize the error function between the neural network's output and the correct output for a given set of inputs, thus solving the problem. When nonlinearities are used, the gradient of the error function can be calculated by conventional procedures [11], [44]. To determine the best learning algorithm, a variety of algorithms were evaluated, including Levenberg–Marquardt, Fletcher–Reeves conjugate gradient, Powell–Beale conjugate gradient, Polak-Ribiére conjugate gradient, BFGS Quasi-Newton, Scaled conjugate gradient, Random order incremental training, One-step secant, Variable learning rate, and Gradient descent with adaptive learning rate. Different activation functions such as logistic sigmoid (logsig), tangent hyperbolic (tanh), triangular (tribas), and soft max (softmax) were also studied.

Table 2. Comparison various learning algorithms with different activation functions in the hidden layer

Training Algorithm		Transfer Function	Number of epochs	Testing phase "R"
Levenberg–Marquardt	Trainlm	Tangsig	1500	0.8362
		Logsig		0.9683
		Tribas		0.7897
		Softmax		0.8550
Bayesian regularization	Trainbr	Tangsig	1500	0.7525
		Logsig		0.8145
		Tribas		0.8309
		Softmax		0.9130
Fletcher–Reeves conjugate gradient	Traincgf	Tangsig	1500	0.8478
		Logsig		0.9173
		Tribas		0.8398
		Softmax		0.6518
Powell–Beale conjugate gradient	Traincgb	Tangsig	1500	0.8970
		Logsig		0.8143
		Tribas		0.8755
		Softmax		0.8255
Polak-Ribiére conjugate gradient	Traincgp	Tangsig	1500	0.7764
		Logsig		0.8349
		Tribas		0.8498
BFGS Quasi-Newton	trainbfg	Tangsig	1500	0.8831
		Logsig		0.9173
		Tribas		0.7923
		Softmax		0.8735
Scaled conjugate gradient	trainscg	Tangsig	1500	0.8055
		Logsig		0.7681
		Tribas		0.8227
		Softmax		0.8502
Random order incremental training	trainr	Tangsig	1500	0.8446
		Logsig		0.6633
		Tribas		0.7887
		Softmax		0.5550
One step secant	trainoss	Tangsig	1500	0.7733
		Logsig		0.8556
		Tribas		0.7924
		Softmax		0.8434
Variable learning rate	traingdx	Tangsig	1500	0.8565
		Logsig		0.4498
		Tribas		0.8080
		Softmax		0.5962
Gradient descent with adaptive learning rate	traingda	Tangsig	1500	0.7482
		Logsig		0.5314
		Tribas		0.7925
		Softmax		0.5926

Additionally, the number of neurons in the hidden layer varied between 5 and 40 for each learning algorithm, with 1500 epochs. Table 2 provides a detailed comparison of the performance of different training algorithms and activation functions. In this study, the BFGS Quasi-Newton algorithm was selected as the learning algorithm because it achieved a higher correlation coefficient compared to the other learning algorithms.

Hence, the configuration of the ANN-MLPs for the modeling of weather conditions is mentioned in Fig. 5. Its more detailed architecture is illustrated in Table 3. The weight matrices and bias vectors of the enhanced ANN-MLPs models are applied as follows:

- w_I : weight matrix for connections from the input layer to the hidden layer (17 rows x 6 columns).
- w_h : weight matrix for connections from the hidden layer to the output layer (17 rows x 1 column.
- b_h : bias vector for the hidden layer neurons (17 rows).
- b_o : bias vector for the output layer neuron (1 row).

Fig. 5 illustrates the optimized ANN-MLP models, and the assimilation of weather conditions (WC) can be represented by a mathematical model incorporating all inputs Xi, as described by the following equations:

The instance outputs Z_j of the hidden layer are:

$$Z_j = f_H\left[\sum_{i=1}^{06} w_{ji}^{I} x_i + b_j^{H}\right] = \frac{\exp\left(\sum_{i=1}^{06} w_{ji}^{I} x_i + b_j^{H}\right) - \exp\left(-\sum_{i=1}^{06} w_{ji}^{I} x_i + b_j^{H}\right)}{\exp\left(\sum_{i=1}^{06} w_{ji}^{I} x_i + b_j^{H}\right) + \exp\left(-\sum_{i=1}^{06} w_{ji}^{I} x_i + b_j^{H}\right)} \quad (8)$$

Where, j=1, 2, …, 17

The output "weather conditions"

$$\text{Weather condition} = f_0\left[\sum_{j=1}^{17} w_{1j}^{H} Z_j + b_1^{o}\right] = \sum_{j=1}^{17} w_{1j}^{H} Z_j + b_1^{o} \quad (9)$$

The combination of Equations (1) and (2) yields the following mathematical formula, which terms the weather conditions by incorporating all input variables.

$$\text{weather conditions} = \sum_{j=1}^{17} w_{1j}^{H} \frac{\exp\left(\sum_{i=1}^{06} w_{ji}^{I} x_i + b_j^{H}\right) - \exp\left(-\sum_{i=1}^{06} w_{ji}^{I} x_i + b_j^{H}\right)}{\exp\left(\sum_{i=1}^{06} w_{ji}^{I} x_i + b_j^{H}\right) + \exp\left(-\sum_{i=1}^{06} w_{ji}^{I} x_i + b_j^{H}\right)} + b_1^{o} \quad (10)$$

Table 3. Structure of the optimized optimal (ANN-MLPs) model

Training Algorithm	Input layer	Hidden layer		Output layer	
	Neurons number	Neurons number	Activation function	Neurons number	Activation function
Levenberg–Marquardt	06	17	logistic Sigmoid	1	Linear (purelin)

The results from the optimal-neural networks model (ANN-MLPs) followed the Dames guideline [45], which suggests that different input profiles can be processed by each neuron. As the number of neurons increases, the network becomes better at fitting the presented data but at the cost of reduced generalizability. Dames [45] also outlined empirical rules for determining the number of hidden neurons, proposing:

- Equal to that of the input layer.

- Equal to 75% of it.
- Equal to the square root of the product of the number of neurons in the input and output layer.

However, this rule contradicts our findings with this database. To prevent the neural model's complexity from exceeding the available data size, we adhere to the rule proposed by Roubehie Fissa et al. [46], which states: (Number of input neurons * Number of hidden neurons) + (Number of hidden neurons * Number of output neurons) ≤ Database size. This rule aligns with our study, as shown in the calculation: $(6 * 17) + (17 * 1) = 119 \leq 480$ (database size).

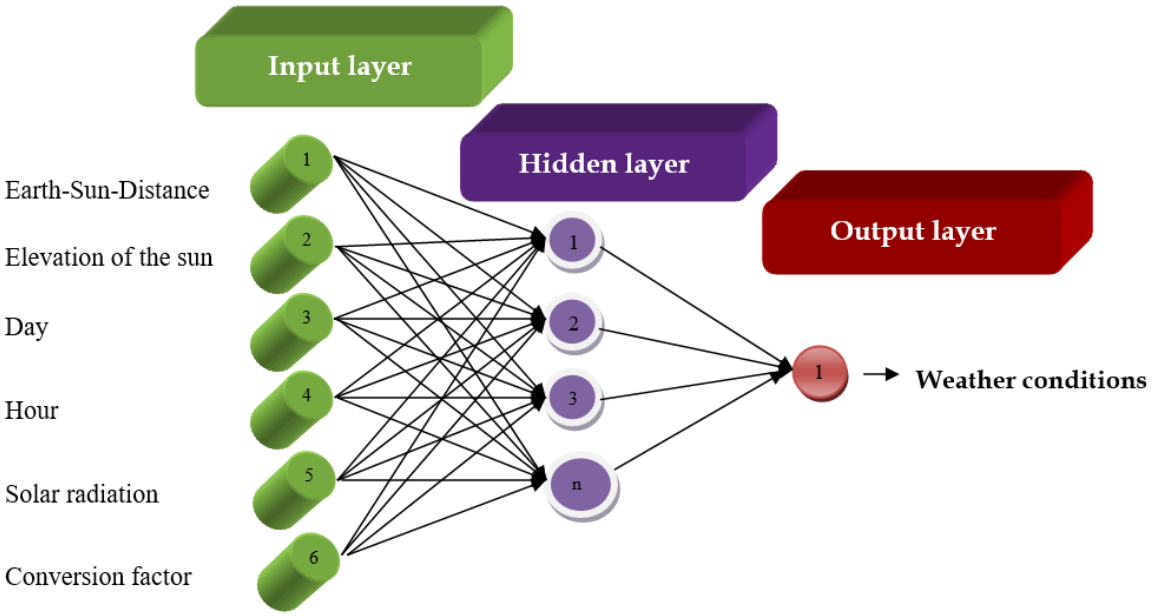

Fig. 5. Artificial neural network for predicting weather conditions

According to this analysis, the parameters and plot of the linear regression were conveniently generated using the MATLAB 2020b function "postreg" (Fig. 6 (a), (b), (c), and (d)). An artificial neural network-multilayer perceptrons (ANN-MLPs) model was developed to predict weather conditions in the Ghardaïa region of Algeria. Fig. 6 (a), (b), (c), and (d) present a comparison between the experimental and calculated values of weather conditions, showing vectors of agreement near the optimization of the FFNN-MLP profiles for the optimal model obtained (train Levenberg-Marquardt - Logsig). The correlation coefficients are as follows: R = 0.9888 for the training phase, R = 0.9793 for the validation phase, R = 0.9683 for the test phase, and R = 0.9867 for the total phase. Generally, correlation coefficients are considered excellent when $0.9 \leq R \leq 1$, demonstrating the robustness of the neural network models and their ability to predict weather conditions accurately.

Table 4 displays the performance statistics and statistical parameters of the optimal ANN-MLPs model across the training, validation, and testing phases. The correlation coefficients (R) for the training and validation phases are remarkably high, at 0.9888 and 0.9773, respectively, demonstrating a strong alignment between the experimental and predicted results. The testing phase also shows a high correlation coefficient (R) of 0.9683, reflecting a strong match between the experimental and predicted weather conditions and demonstrating the model's robust interpolation capability.

Table 4. Statistical performance of the optimal ANN-MLPs model

Phase	R	RMSE	MAE	SEP
Training phase	0.9888	2.7077	1.8183	16.9764
Validation phase	0.9773	3.2869	2.2895	18.6613
Testing phase	0.9683	3.3043	2.0676	27.9413

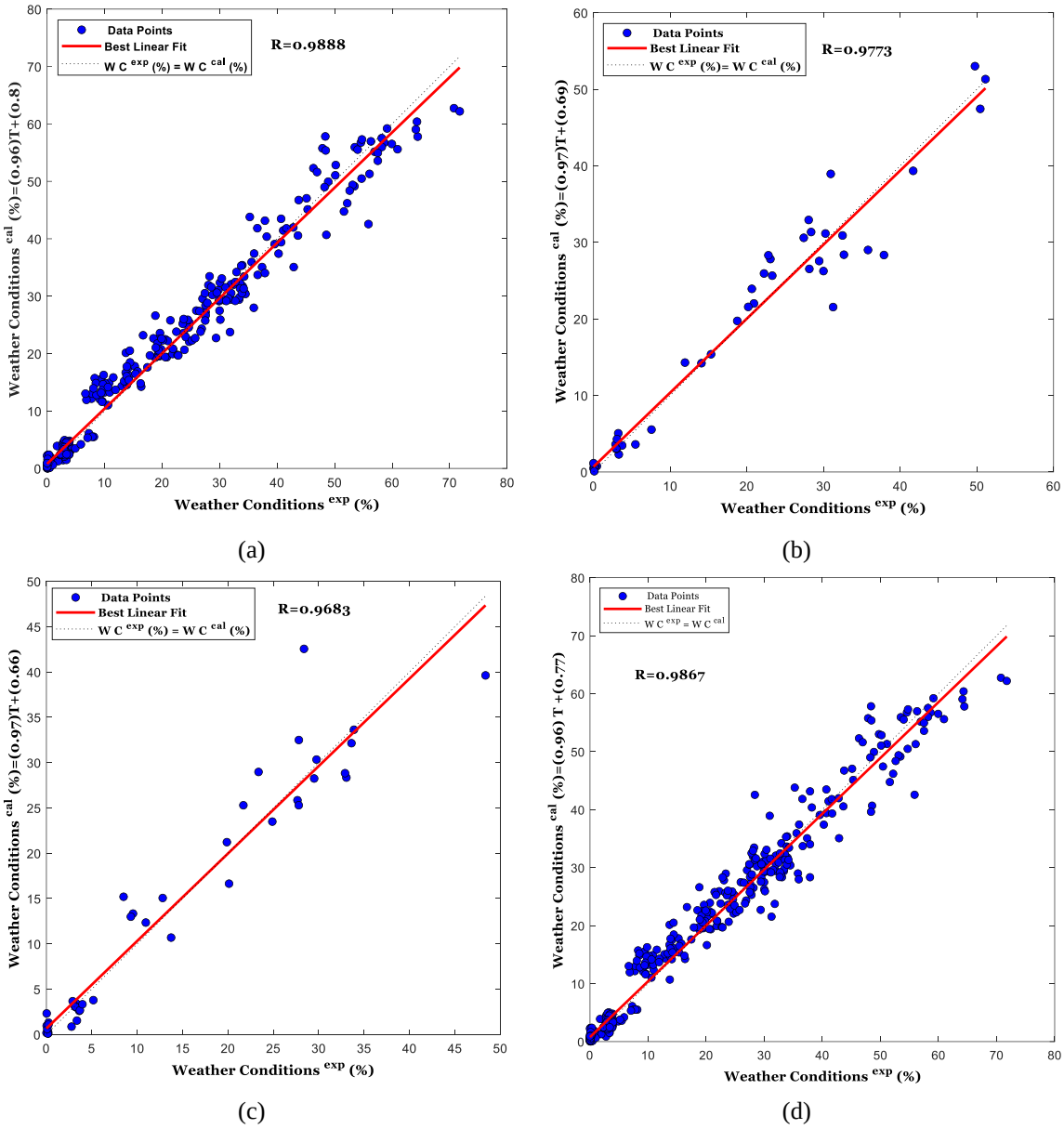

Fig. 6. Comparison between experimental and predicted values: (a) training phase, (b) validation phase, (c) testing phase, and (d) total phase

4.2. Sensitivity Analysis

The contribution of various input variables—earth-sun distance ($D_{T\text{-}S}$), elevation of the Sun (H_I), normalized days (D*), normalized hours (H*), solar radiation (SR), and conversion factor (δ) to the output (weather conditions) was evaluated using a sensitivity analysis via the "Weight" method. This method, initially proposed by Garson [47] and Goh [47], is based on the division of connection weights (between input and hidden layers, and between hidden and output layers). The contribution results are shown in Fig. 7. The most significant variables influencing the prediction of weather conditions are earth-sun distance ($D_{T\text{-}S}$) with RI = 31.10%, factor conversion with RI = 26.05%, and solar radiation (SR) with RI = 16.26%. The contributions of the elevation of the Sun (HI) with RI = 13.29%, normalized hours (H*) with RI = 6.95%, and normalized days (D*) with RI = 6.35% also have a notable effect. The sensitivity analysis results indicate that all input parameters have a relative importance higher than 5%, underscoring the impact of the selected parameters on the output.

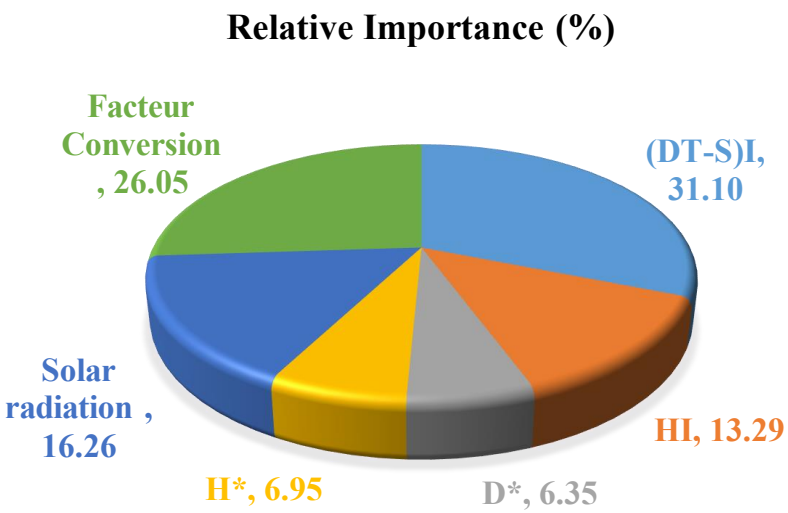

Fig. 7. Relative importance of input variables on the value of the calculated weather conditions

4.3. Applicability Domain

In this study, the Leverage mathematical technique was applied to detect outliers, utilizing residual values and a Hat matrix. The process for calculating the Hat values and the key steps involved are as follows [48]–[50]:

$$H = X(X^tX)^{-1}X^t \tag{11}$$

In this context, X represents the $m \times n$ matrix, where m denotes the quantity of samples and n corresponds to the model's factors (input variables). The Hat values are derived from the main diagonal of the H matrix.

$$Hat = diagonal(H) \tag{12}$$

The Williams diagram, which plots normalized residuals against leverage (Hat values), is a graphical tool used to identify outliers. In this diagram, the critical leverage value (H*), often considered the threshold, is typically determined using the following equation [51]:

$$H^* = \frac{3(n+1)}{m} \tag{13}$$

The normalized residuals are determined by comparing the experimental weather condition data with the values predicted by the model.

$$(R\_Norm)_i = \frac{\left(\text{Weather conditions}_i^{exp} - \text{Weather conditions}_i^{cal}\right)}{\sqrt{\text{Var}(\text{Weather conditions}^{exp} - \text{Weather conditions}^{cal})}} \quad i = 1, \dots m \tag{14}$$

Fig. 8 presents the Williams plot for the applicability domain in the testing phase. In this chart, 47 out of 48 data points (98%) lie within the horizontal lines (±3 range), with only 1 point (2%) falling outside the suspected limit.

$H^* = \frac{3(n+1)}{m} = \frac{3(7+1)}{48} = 0.5$, This suggests that the development of the optimal ANN-MLPs model and its predictions remain within acceptable limits, resulting in a statistically valid neural network. Consequently, it can be confirmed that there are "Good Haut Levier" points for the testing phase.

Fig. 9 represents Williams range plot of ANN-MLPs optimal neural model for the total phase. This plot contains 472 out of 480 data points (98.33%) are within the ±3 range, and 8 points (1.66%) are outside the suspected limit.

$H^* = \frac{3(n+1)}{m} = \frac{3(7+1)}{480} = 0.05$, Due to its high predictive capability, the proposed model can be used to screen existing databases to identify weather conditions. The applicability domain can therefore serve as a valuable tool for filtering out dissimilar weather condition data points.

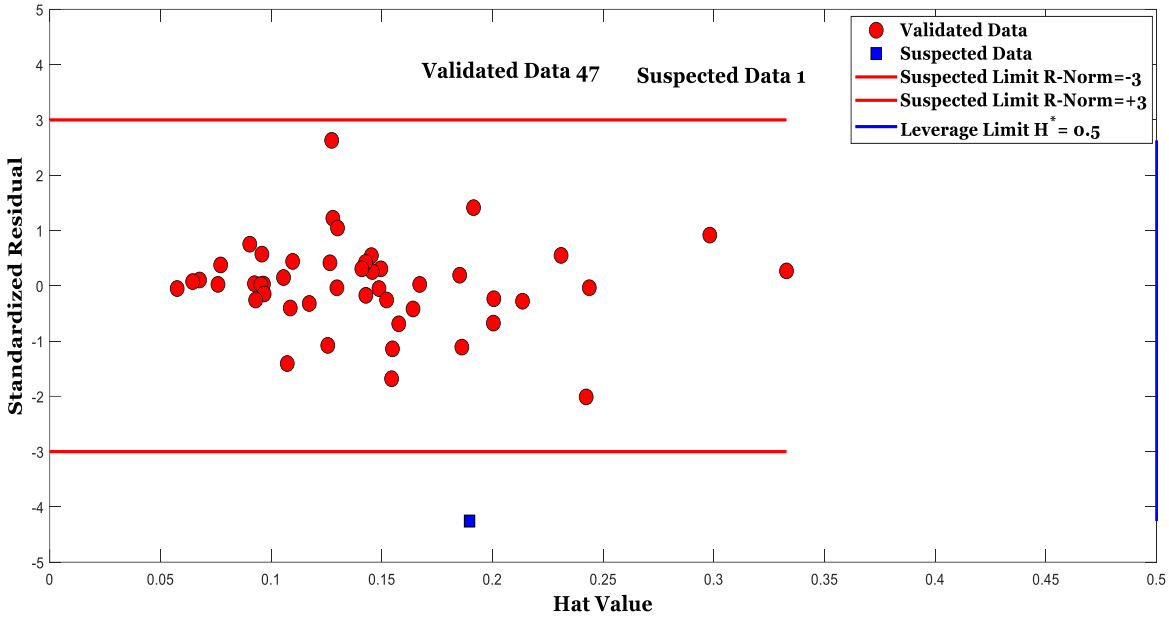

Fig. 8. Williams range plot for optimal ANN-MLPs in testing phase

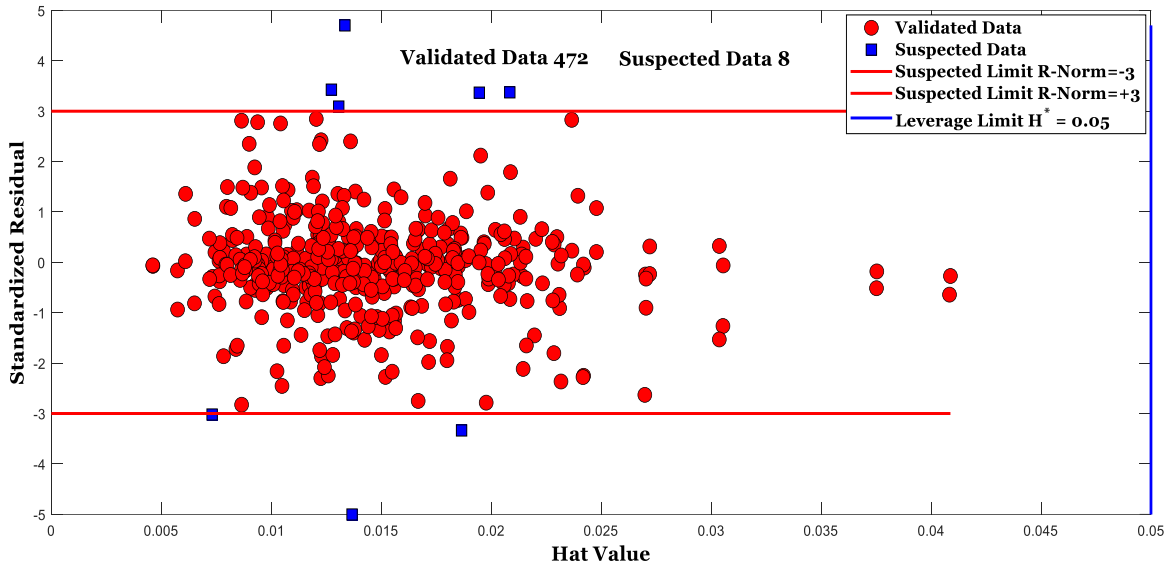

Fig. 9. Williams range plot for optimal ANN-MLPs in total phase

5. Conclusion and Future Work

In this paper, we develop an artificial neural network-multilayer perceptrons (ANN-MLPs) model to predict weather conditions using various meteorological variables (wind speed, precipitation, humidity, average temperature, and precipitation). The dataset includes the Earth-Sun distance, elevation of the Sun, normalized days, normalized hours, solar radiation, and the conversion factor. Accurate weather condition predictions can be achieved using efficient multilayer perceptrons (MLPs) techniques. An optimal ANN-MLPs configuration features a structure with 6 neurons in the input layer, 17 neurons in the hidden layer, and 1 neuron in the output layer. The network was trained using

the Levenberg-Marquardt algorithm ("train-LM") with a "logsig" activation function in the hidden layer and a "Purelin" activation function in the output layer. The optimal ANN-MLPs model exhibited a high level of consistency between the predicted and experimental data points during the testing phase, with a correlation coefficient of R=0.9683 and a root mean square error (RMSE) of 3.3043%.

The sensitivity analysis using the weight method effectively identified the true significance of each variable in predicting the impact of weather conditions, with the Earth-Sun distance ($D_{T\text{-}S}$) emerging as particularly influential (RI = 31.10%). This confirms the relevance of the variables chosen for this study. The applicability domain and outlier diagnostics for the optimized ANN-MLPs neural model demonstrated that its development and predictions lie within the valid application domain. This supports the model's statistical validity and highlights the presence of "Good High Leverage" points during the validation and testing phases. As a result, the developed model is both reliable and capable of delivering accurate weather predictions. Future works will consider other different types of neural networks and machine learning approaches.

Supplementary Materials: Not Applicable.

Author Contribution: All authors contributed equally to the main contributor to this paper. The paper was supervised by Abdel-Nasser Sharkawy. All authors read and approved the final paper.

Funding: This research received no external funding.

Acknowledgment: The authors thank their institutions and Universities for their support and encouragement.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] L. Chen *et al*., "FuXi: a cascade machine learning forecasting system for 15-day global weather forecast," *npj climate and atmospheric science*, vol. 6, no. 1, pp. 1–11, 2023, https://doi.org/10.1038/s41612-023-00512-1.

[2] M. T. Hasan, K. M. Fattahul Islam, M. S. Rahman, and S. Li, "Weather Forecasting Using Artificial Neural Network," *Artificial Intelligence and Security*, vol. 11633, no. 22, pp. 171–180, 2019, https://doi.org/10.1007/978-3-030-24265-7_15.

[3] M. Laidi, S. Hanini, A. Rezrazi, M. R. Yaiche, A. A. El Hadj, and F. Chellali, "Supervised artificial neural network-based method for conversion of solar radiation data (case study: Algeria)," *Theoretical and Applied Climatology*, vol. 128, no. 1-2, pp. 439-451, 2016, https://doi.org/10.1007/s00704-015-1720-7.

[4] A. Dahmani, Y. Ammi, and S. Hanini, "A Novel Non-Linear Model Based on Bootstrapped Aggregated Support Vector Machine for the Prediction of Hourly Global Solar Radiation," *Smart Grids and Sustainable Energy*, vol. 9, no. 1, p. 3, 2023, https://doi.org/10.1007/s40866-023-00179-w.

[5] S. Scher and G. Messori, "Predicting weather forecast uncertainty with machine learning," *Quarterly Journal of the Royal Meteorological Society*, vol. 144, no. 717, pp. 2830–2841, 2018, https://doi.org/10.1002/qj.3410.

[6] A. -N. Sharkawy *et al*., "Solar PV Power Estimation and Upscaling Forecast Using Different Artificial Neural Networks Types: Assessment, Validation, and Comparison," *IEEE Access*, vol. 11, pp. 19279-19300, 2023, https://doi.org/10.1109/ACCESS.2023.3249108.

[7] J. Sillmann *et al*., "Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities," *Weather and Climate Extremes*, vol. 18, pp. 65–74, 2017, https://doi.org/10.1016/j.wace.2017.10.003.

[8] Z. Pang, F. Niu, and Z. O'Neill, "Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons," *Renewable Energy*, vol. 156, pp. 279–289, 2020, https://doi.org/10.1016/j.renene.2020.04.042.

[9] N. Bailek *et al*., "Developing a new model for predicting global solar radiation on a horizontal surface located in Southwest Region of Algeria," *NRIAG Journal of Astronomy and Geophysics*, vol. 9, no. 1, pp. 341–349, 2020, https://doi.org/10.1080/20909977.2020.1746892.

[10] B. Amiri, R. Dizène, and K. Dahmani, "Most relevant input parameters selection for 10-min global solar irradiation estimation on arbitrary inclined plane using neural networks," *International Journal of Sustainable Energy*, vol. 39, no. 8, pp. 779–803, 2020, https://doi.org/10.1080/14786451.2020.1758104.

[11] A. Dahmani *et al*., "Assessing the Efficacy of Improved Learning in Hourly Global Irradiance Prediction," *Computers, Materials & Continua*, vol. 77, no. 2, pp. 2579-2594, 2023, https://doi.org/10.32604/cmc.2023.040625.

[12] A. Sharkawy, M. M. Ali, H. H. H. Mousa, A. S. Ali, G. T. Abdel-Jaber, "Short-Term Solar PV Power Generation Day-Ahead Forecasting," *International Journal of Robotics and Control Systems*, vol. 2, no. 3, pp. 562–580, 2022, https://doi.org/10.31763/ijrcs.v2i3.780.

[13] M. Afzali, A. Afzali, and G. Zahedi, "The Potential of Artificial Neural Network Technique in Daily and Monthly Ambient Air Temperature Prediction," *International Journal of Environmental Science and Development*, vol. 3, no. 1, pp. 33–38, 2012, https://doi.org/10.7763/IJESD.2012.V3.183.

[14] T. T. K. Tran, S. M. Bateni, S. J. Ki, and H. Vosoughifar, "A review of neural networks for air temperature forecasting," *Water*, vol. 13, no. 9, p. 1294, 2021, https://doi.org/10.3390/w13091294.

[15] S. AlSadi and T. Khatib, "Modeling of relative humidity using artificial neural network," *Journal of Asian Scientific Research*, vol. 2, no. 2, p. 81, 2012, https://archive.aessweb.com/index.php/5003/article/view/3329.

[16] E. Kuzugudenli, "Relative humidity modeling with artificial neural networks," *Applied Ecology and Environmental Research*, vol. 16, no. 4, pp. 5227–5235, 2018, https://www.aloki.hu/pdf/1604_52275235.pdf.

[17] M. I. C. Rachmatullah, J. Santoso, and K. Surendro, "Determining the number of hidden layer and hidden neuron of neural network for wind speed prediction," *PeerJ Computer Science*, vol. 7, p. e724, 2021, https://doi.org/10.7717/peerj-cs.724.

[18] A. N. Sharkawy, A. G. Ameen, S. Mohamed, G. T. Abdel-Jaber, and I. Hamdan, "Design, Assessment, and Modeling of Multi-Input Single-Output Neural Network Types for the Output Power Estimation in Wind Turbine Farms," *Automation*, vol. 5, no. 2, pp. 190–212, 2024, https://doi.org/10.3390/automation5020012.

[19] Z. Tang, G. Zhao, and T. Ouyang, "Two-phase deep learning model for short-term wind direction forecasting," *Renewable Energy*, vol. 173, pp. 1005–1016, 2021, https://doi.org/10.1016/j.renene.2021.04.041.

[20] A. Bouach, "Artificial neural networks for monthly precipitation prediction in north-west Algeria: a case study in the Oranie-Chott-Chergui basin," *Journal of Water & Climate Change*, vol. 15, no. 2, pp. 582–592, 2024, https://doi.org/10.2166/wcc.2024.494.

[21] S. Sharadqah, A. M. Mansour, M. A. Obeidat, R. Marbello, and S. M. Perez, "Nonlinear Rainfall Yearly Prediction based on Autoregressive Artificial Neural Networks Model in Central Jordan using Data Records: 1938-2018," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 12, no. 2, pp. 240-247, 2021, https://dx.doi.org/10.14569/IJACSA.2021.0120231.

[22] M. Ghamariadyan and M. A. Imteaz, "Prediction of seasonal rainfall with one-year lead time using climate indices: A wavelet neural network scheme," *Water Resources Management*, vol. 35, no. 15, pp. 5347–5365, 2021, https://doi.org/10.1007/s11269-021-03007-x.

[23] K. Abhishek, M. P. Singh, S. Ghosh, and A. Anand, "Weather forecasting model using artificial neural network," *Procedia Technology*, vol. 4, pp. 311–318, 2012, https://doi.org/10.1016/j.protcy.2012.05.047.

[24] R. Yacef, A. Mellit, S. Belaid, and Z. Şen, "New combined models for estimating daily global solar radiation from measured air temperature in semi-arid climates: application in Ghardaïa, Algeria," *Energy Conversion and Management*, vol. 79, pp. 606–615, 2014, https://doi.org/10.1016/j.enconman.2013.12.057.

[25] M. Guermoui, S. Benkaciali, K. Gairaa, K. Bouchouicha, T. Boulmaiz, and J. W. Boland, "A novel ensemble learning approach for hourly global solar radiation forecasting," *Neural Computing and Applications*, vol. 34, no. 4, pp. 2983–3005, 2022, https://doi.org/10.1007/s00521-021-06421-9.

[26] H. Bouzemlal *et al*., "New method of prediction of climatic conditions by artificial neural networks (case of the province of Médéa)," *Algerian Journal of Environmental Science and Technology*, vol. 9, no. 4, pp. 1419-1428, 2023, https://www.aljest.net/index.php/aljest/article/view/912.

[27] R. W. Mueller, "Solar irradiance, global distribution," *Solar Energy*, pp. 553–583, 2013, https://doi.org/10.1007/978-1-4614-5806-7_447.

[28] J. A. Duffie and W. A. Beckman, "Solar engineering of thermal processes," *John Wiley & Sons*, 2013, https://doi.org/10.1002/9781118671603.

[29] B. Amiri, "Estimation temporelle du rayonnement solaire au nord et au Sahara Algérien à partir de données de mesures sélectionnées," *PhD Thesis, Université des Sciences et de la Technologie d'Alger*, 2021, https://dspace.usthb.dz/handle/123456789/9228.

[30] D. Markovics, M. J. Mayer, "Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction," *Renewable and Sustainable Energy Reviews*, vol. 161, p. 112364, 2022, https://doi.org/10.1016/j.rser.2022.112364.

[31] Z. Şen, "Solar Radiation Deterministic Models," *Solar Energy Fundamentals and Modeling Techniques*, pp. 47-99, 2008, https://doi.org/10.1007/978-1-84800-134-3_3.

[32] R. Foster, M. Ghassemi, and A. Cota, "Solar energy: renewable energy and the environment," *CRC press*, p. 382, 2009, https://doi.org/10.1201/9781420075670.

[33] C. M. Siham, H. Salah, L. Maamar, and K. Latifa, "Artificial neural networks based prediction of hourly horizontal solar radiation data: case study," *International Journal of Applied Decision Sciences*, vol. 10, no. 2, p. 156, 2017, https://doi.org/10.1504/IJADS.2017.084312.

[34] A. Dahmani, Y. Ammi, and S. Hanini, "Neural network for prediction solar radiation in Relizane region (Algeria)-Analysis study," *International Journal of Energetica*, vol. 7, no. 2, pp. 8–18, 2022, https://www.ajol.info/index.php/ijeca/article/view/264400.

[35] M. A. Hassan, M. Abubakr, A. Khalil, "A profile-free non-parametric approach towards generation of synthetic hourly global solar irradiation data from daily totals," *Renewable Energy*, vol. 167, pp. 613-628, 2021, https://doi.org/10.1016/j.renene.2020.11.125.

[36] Y. Ammi, L. Khaouane, and S. Hanini, "A Model Based on Bootstrapped Neural Networks for Modeling the Removal of Organic Compounds by Nanofiltration and Reverse Osmosis Membranes," *Arabian Journal for Science and Engineering*, vol. 43, no. 11, pp. 6271–6284, 2018, https://doi.org/10.1007/s13369-018-3484-8.

[37] A. Dahmani, Y. Ammi, S. Hanini, and Z. Driss, "Developed nonlinear model based on bootstrap aggregated neural networks for predicting global hourly scale horizontal irradiance," *International Journal of Energetica*, vol. 8, no. 1, pp. 31–37, 2023, https://www.ajol.info/index.php/ijeca/article/view/263636.

[38] A. Dahmani, Y. Ammi, S. Hanini, M. Redha Yaiche, and H. Zentou, "Prediction of Hourly Global Solar Radiation: Comparison of Neural Networks/Bootstrap Aggregating," *Kemija u industriji: Časopis kemičara i kemijskih inženjera Hrvatske*, vol. 72, no. 3–4, pp. 201–213, 2023, https://doi.org/10.15255/KUI.2022.065.

[39] A.-N. Sharkawy, M. Ali, H. Mousa, A. Ali, and G. Abdel-Jaber, "Machine Learning Method for Solar PV Output Power Prediction," *SVU-International Journal of Engineering Sciences and Applications*, vol. 3, no. 2, pp. 123–130, 2022, https://doi.org/10.21608/svusrc.2022.157039.1066.

[40] A.-N. Sharkawy, "Principle of Neural Network and Its Main Types: Review," *Journal of Advances in Applied & Computational Mathematics*, vol. 7, pp. 8–19, 2020, https://doi.org/10.15377/2409-5761.2020.07.2.

[41] Y. Ammi, L. Khaouane, and S. Hanini, "Stacked neural networks for predicting the membranes performance by treating the pharmaceutical active compounds," *Neural Computing and Applications*, vol. 33, no. 19, pp. 12429–12444, 2021, https://doi.org/10.1007/s00521-021-05876-0.

[42] K. Ikram, K. Djilali, D. Abdennasser, R. Al-Sabur, B. Ahmed, and A. N. Sharkawy, "Comparative analysis of fouling resistance prediction in shell and tube heat exchangers using advanced machine learning techniques," *Research on Engineering Structures and Materials*, vol. 10, no. 1, pp. 253–270, 2024, http://dx.doi.org/10.17515/resm2023.858en0816.

[43] N. Bailek *et al*., "A new empirical model for forecasting the diffuse solar radiation over Sahara in the Algerian Big South," *Renewable Energy*, vol. 117, pp. 530–537, 2018, https://doi.org/10.1016/j.renene.2017.10.081.

[44] A. Benyekhlef, B. Mohammedi, S. Hanini, M. Boumahdi, A. Rezrazi, and M. Laidi, "A Contribution to the Modelling of Fouling Resistance in Heat Exchanger-Condenser by Direct and Inverse Artificial Neural Network," *Kemija u industriji: Časopis kemičara i kemijskih inženjera Hrvatske*, vol. 70, no. 11–12, pp. 639–650, 2021, https://doi.org/10.15255/KUI.2020.076.

[45] J. Kočí, V. Kočí, J. Maděra, R. Černý, "Effect of applied weather data sets in simulation of building energy demands: Comparison of design years with recent weather data," *Renewable and Sustainable Energy Reviews*, vol. 100, pp. 22-32, 2019, https://doi.org/10.1016/j.rser.2018.10.022.

[46] M. R. Fissa, Y. Lahiouel, L. Khaouane, and S. Hanini, "QSPR estimation models of normal boiling point and relative liquid density of pure hydrocarbons using MLR and MLP-ANN methods," *Journal of Molecular Graphics and Modelling*, vol. 87, pp. 109–120, 2019, https://doi.org/10.1016/j.jmgm.2018.11.013.

[47] A. T. C. Goh, "Back-propagation neural networks for modeling complex systems," *Artificial Intelligence in Engineering*, vol. 9, no. 3, pp. 143–151, 1995, https://doi.org/10.1016/0954-1810(94)00011-S.

[48] M. Mansourian, L. Saghaie, A. Fassihi, A. Madadkar-Sobhani, and K. Mahnam, "Linear and nonlinear QSAR modeling of 1, 3, 8-substituted-9-deazaxanthines as potential selective A 2B AR antagonists," *Medicinal Chemistry Research*, vol. 22, pp. 4549–4567, 2013, doi: https://doi.org/10.1007/s00044-012-0453-8.

[49] M. Hamadache, O. Benkortbi, S. Hanini, and A. Amrane, "QSAR modeling in ecotoxicological risk assessment: application to the prediction of acute contact toxicity of pesticides on bees (Apis mellifera L.)," *Environmental Science and Pollution Research*, vol. 25, pp. 896–907, 2018, https://doi.org/10.1007/s11356-017-0498-9.

[50] I. Euldji, A. Belghait, C. Si-Moussa, O. Benkortbi, and A. Amrane, "A new hybrid quantitative structure property relationships-support vector regression (QSPR-SVR) approach for predicting the solubility of drug compounds in supercritical carbon dioxide," *AIChE Journal*, vol. 69, no. 8, p. e18115, 2023, https://doi.org/10.1002/aic.18115.

[51] Y. Ammi, C. Si-Moussa, and S. Hanini, "Machine Learning and Neural Networks for Modelling the Retention of PPhACs by NF/RO," *Kemija u industriji: Časopis kemičara i kemijskih inženjera Hrvatske*, vol. 72, no. 11–12, pp. 617–626, 2023, https://doi.org/10.15255/KUI.2022.085.