scientific reports

OPEN

An innovative machine learning based on feed-forward artificial neural network and equilibrium optimization for predicting solar irradiance

Ting Xu[1], Mohammad Hosein Sabzalian[2], Ahmad Hammoud[3,4], Hamed Tahami[5], Ali Gholami[6] & Sangkeum Lee[7✉]

As is known, having a reliable analysis of energy sources is an important task toward sustainable development. Solar energy is one of the most advantageous types of renewable energy. Compared to fossil fuels, it is cleaner, freely available, and can be directly exploited for electricity. Therefore, this study is concerned with suggesting novel hybrid models for improving the forecast of Solar Irradiance (I_S). First, a predictive model, namely Feed-Forward Artificial Neural Network (FFANN) forms the non-linear contribution between the I_S and dominant meteorological and temporal parameters (including humidity, temperature, pressure, cloud coverage, speed and direction of wind, month, day, and hour). Then, this framework is optimized using several metaheuristic algorithms to create hybrid models for predicting the I_S. According to the accuracy assessments, metaheuristic algorithms attained satisfying training for the FFANN by using 80% of the data. Moreover, applying the trained models to the remaining 20% proved their high proficiency in forecasting the I_S in unseen environmental circumstances. A comparison among the optimizers revealed that Equilibrium Optimization (EO) could achieve a higher accuracy than Wind-Driven Optimization (WDO), Optics Inspired Optimization (OIO), and Social Spider Algorithm (SOSA). In another phase of this study, Principal Component Analysis (PCA) is applied to identify the most contributive meteorological and temporal factors. The PCA results can be used to optimize the problem dimension, as well as to suggest effective real-world measures for improving solar energy production. Lastly, the EO-based solution is yielded in the form of an explicit formula for a more convenient estimation of the I_S.

Background

Recent advances in computational and engineering domains have provided reliable responses to various problems in human modern life[1–5]. Intelligent tools, sophisticated simulation packages, and soft computing approaches are evident examples of these advances[6–9]. In the field of energy, engineers have successfully employed these tools and methodologies to improve the sustainable development of renewable energy systems[10–13]. Recently, solar energy has been introduced as an outstanding renewable source due to its numerous benefits like environmental friendliness, universality, high capacity, and inexhaustible supply[14,15]. Scholars attempted to evaluate the solar energy production pattern by forecasting related parameters such as I_S. However, the evaluation of these

[1]School of Economics and Management, Hubei Engineering University, Hubei 430000, China. [2]Department of Mechanical Engineering, Faculty of Engineering, University of Santiago of Chile (USACH), Avenida Libertador Bernardo O'Higgins 3363, 9170022 Santiago, Chile. [3]Department of Medical and Technical Information Technology, Bauman Moscow State Technical University, Moscow, Russia. [4]Department of Mathematics and Natural Sciences, Gulf University for Science and Technology, Mishref Campus, Kuwait. [5]School of Industrial and Information Engineering, Politecnico Di Milano, 20133 Milan, Italy. [6]Department of Electrical Engineering, Faculty of Technology and Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran. [7]Department of Computer Engineering, Hanbat National University, Daejeon 34158, South Korea. ✉email: sangkeum@hanbat.ac.kr

factors required reliable approaches to handle non-linear calculations because of the existing many involved parameters[16,17]. Today, Machine Learning (ML) models proved to have an impressive approach to handling non-linear calculations[18–20]. In terms of forecasting tasks, ML methods can be used to conduct complicated mathematical relationships and provide exact solutions. Artificial Neural Network (ANN)[21,22], Support Vector Machine (SVM)[23,24], decision trees[25,26], and neuro-fuzzy tools[27,28] are among the most popular ML algorithms utilized for prediction aims related to solar energy calculations.

Literature review

ML algorithms provided fast, inexpensive, and reliable solutions, which motivated experts to take advantage of them in the forecasting tasks[29–33]. Kim, Seong and Choi[34] utilized ANN Models to forecast the energy consumption of an actual air handling unit and the appropriate result was obtained. Bhatt and Gandhi[35] used two different statistical and ANN models to forecast the energy consumption in the wind power plants and the error of the feed-forward neural network was determined to be around 9.85%. Yin, Jia, Wu, Dai and Tang[36] used a feedforward ANN model for forecasting tasks in the case of energy demand and the mean relative error value of the forecast was determined to be 1.58%. Malvoni, De Giorgi and Congedo[37] utilized an SVM method to forecast the data from the Photovoltaic (PV) power. This method was also utilized in Ref. for forecasting wind energy production in Estonia and compared to other traditional methods like Behavior-Driven Development (BDD), appropriate results were obtained.

Optimization-oriented efforts form lots of studies in the engineering literature[38–40]. In particular, many scholars have suggested metaheuristic algorithms for optimization purposes[41,42]. They can also serve to optimize traditional ML methods like ANN and ANFIS[43]. These algorithms were concerned in the case of renewable energy analysis in many studies previously[44,45] such as solar power energy[46] and wind energy[47,48]. Computational problems such as local minima can be removed by using the metaheuristic-based hybrids[29]. In the following, previous works related to the use of metaheuristic algorithms to optimize ML methods in the case of energy forecasting tasks are briefly summarized. Moayedi and Mosavi[49] utilized an innovative metaheuristic approach (Electromagnetic Field Optimization (EFO)) to optimize a neural network and proved that the EFO-supervised neural network algorithm can appropriately mine a dataset of nonlinearly tuning the network elements. Abedinia, Amjady and Ghadimi[50] utilized a productive engine consisting of a metaheuristic optimizer, namely shark smell optimization to optimize ANN. They claimed that this hybrid method had better performance compared to other conventional predictors such as ANN with lower normalized Root Mean Square Errors (RMSEs) by about 27% compared to ANN and other traditional methods. Galván, Valls, Cervantes and Aler[51] used a multi-objective Particle Swarm Optimization (PSO) method to enhance the ANN method and observed that the PSO optimizer had outstanding results compared to traditional ANN. Tran, Luong and Chou[52] introduced a new model namely Evolutionary Neural Machine Inference Model (ENMIM) consisting of different models of Least Squares Support Vector Regression (LSSVR), and the Radial Basis Function Neural Network (RBFNN) together with Symbiotic Organism Search (SOS) for obtaining optimized tuning parameters. This approach was proved to be a promising alternative for the energy management tasks. Halabi, Mekhilef and Hossain[53] demonstrated that the algorithm introduced in[54] can be coupled with an ANFIS system in the case of I_S predictions. Louzazni, Khouya, Amechnoue, Gandelli, Mussetta and Crăciunescu[55] have proven the competency of the algorithm of firefly to evaluate the solar energy harvesting parameters and observed that the firefly algorithm was very reliable in the case of solar energy forecasting. Bechouat, Younsi, Sedraoui, Soufi, Yousfi, Tabet and Touafek[56] also concerned the PSO and Genetic Algorithm (GA) in this case. Zhou, Moayedi and Foong[57] have studied the limitation of neural computing approaches, for example, local minima, and suggested a novel metaheuristic method namely Teaching–Learning-Based Optimization (TLBO) for enhancing a Multi-Layer Perceptron Neural Network (MLPNN). They observed that, by using TLBO method, the prediction error is reduced by 19.89% compared to the ANN approach. Vaisakh and Jayabarathi[58] utilized a hybrid approach called the deer hunting optimization algorithm as well as grey wolf optimization to adjust the structure of ANNs, which was used for solar energy calculations. Their achievements reflected a notable improvement attained by the tested optimizer. Abedinia, Amjady and Ghadimi[59] have used a neural network algorithm enhanced by a metaheuristic algorithm as the hybrid method for the forecasting tasks in the case of solar energy harvesting, and appropriate results were obtained. Abdalla, Rezk and Ahmed[60] have successfully utilized Wind-Driven Optimization (WDO) to track the elements of photovoltaic systems and justified that the mentioned algorithm had better results compared to many traditional optimization techniques such as cuckoo search.

Motivation, novelty, and objective

The above literature shows the necessity of utilizing modern tools and techniques for coping with intricate engineering problems[61–65]. In this sense, different ML models have great contributions to the concept of renewable energy, particularly for SE-related predictions. On the other hand, metaheuristic algorithms have been recommended for optimal development of ML models. Based on the previous literature, incorporating metaheuristic optimizers with ML models such as ANN helps to avoid computational drawbacks, and therefore, is becoming a research hotspot in this way[66]. However, a gap of knowledge emerges when these studies mostly focus on earlier metaheuristic methods such as PSO and GA[67,68], because the metaheuristic family is being extended by new potential members. This gap calls for evaluating the capability of newer hybrid models to improve SE-related predictions. Hence, in this research, a novel potential metaheuristic technique named EO is employed through an FFANN framework to analyze the meteorological and temporal conditions and predict the I_S. The EO algorithm here is responsible for best-tuning the FFANN's weights (and biases) which connect the I_S to the environmental conditions. Moreover, to comparatively validate the performance of the EO, this algorithm is evaluated versus three benchmark optimizers including WDO, Optics Inspired Optimization (OIO), and Social Spider Algorithm

(SOSA), as well as three algorithms of the EFO, Shuffled Complex Evolution (SCE), and Shuffled Frog Leaping Algorithm (SFLA) used in an earlier study by Moayedi and Mosavi[49]. Accuracy assessment is carried out using different criteria to rank them and distinguish the most competent model. Since the used models have not been applied to this problem before, the findings of this study can assist solar energy experts in the appropriate selection of predictive models. For more convenience, a mathematical formula is also extracted from the EO-FFANN model to eliminate the need for computer-aided implementations in predicting the I_S. Moreover, a well-known statistical technique called Principal Component Analysis (PCA) is applied to identify the most contributive meteorological and temporal parameters, and therefore, to optimize the dimension of the problem.

To sum up, the main strengths and novelties of this study can be highlighted as follows:

– Evaluating the applicability of ensemble learning theory for predicting the I_S as a crucial parameter of renewable (solar) energy,
– Employing the EO metaheuristic algorithm to create a novel FFANN-based model whose absence is considered a gap of knowledge in the literature on I_S prediction,
– Introducing the optimal configurations (i.e., population size and No. of iterations) for the used models,
– Exposing various environmental and temporal conditions as key parameters in the I_S prediction and determining the principal dataset components using the PCA method which has not been performed in the previous literature. In addition to optimizing the problem dimension, the results of the PCA can be regarded for suggesting real-world measures (attributing to the key parameters) to maximize solar energy production.
– Conducting a comparative assessment by evaluating six other metaheuristic algorithms in this study (i.e., WDO, OIO, and SOSA) and previous literature (i.e., EFO, SCE, and SFLA). It makes this study a suitable benchmark for future applications of hybrid models and appropriate model selection by energy experts,
– Developing a monolithic explicit formula from the proposed EO-FFANN model to be used as a convenient method for predicting the I_S.

Overall, the achievements of this research can greatly contribute to the body of knowledge (from both data and methodology perspectives) that deals with solar energy modeling. Performing several optimization ideas carried out in this study can be helpful to reduce the complexities (i.e., computational costs) in the way of proper I_S prediction.

In the following, the study continues by introducing the used materials and methods in Sect. 2, presenting the results and discussion in Sect. 3, followed by providing conclusions in Sect. 4.

Materials and methods
Dataset and splitting
From previous studies, it is evident that the amount of received I_S is a function of various meteorological conditions[69,70]. In this work, this amount is represented by a so-called parameter Global Horizontal Irradiance (GI_H) which is measured for Yemen. Along with the GI_H, the records of five meteorological factors, namely: Air Temperature (T), Relative Humidity (H), Surface Pressure (P), Wind Direction (WD), and Wind Speed (WS) are downloaded from the Solcast community (https://solcast.com/). All measurements are hourly within one year (2021-05-31 to 2022-06-01). Figure 1 shows the time series of the T, H, P, WD, WS, and GI_H.

In addition to these five parameters, three temporal inputs, namely Month (m), Day (d), and Hour (h) are also considered influential parameters. When this dataset is exposed to the considered ML models, the influential parameters (i.e., m, d, h, T, H, P, WD, and WS) play the role of inputs, while the GI_H is the target of the system. Therefore, the used models explore the relationship between the temporal and meteorological parameters to understand and predict the hourly GI_H. Table 1 gives the results of the statistical analysis performed on the used dataset.

As per Table 1, a total of 8803 records exist in the dataset. These records are split into two sub-sets for creating the training and testing sets. The training set is required to provide the training material for the models, and the testing set examines the generalizability of the models. Based on previous works, 80:20 ratio is applied to split the dataset, meaning that 7042 records exist in the training set, and 1761 records exist in the testing set.

Applied algorithms
Overview of EO
As the name implies, the EO is an optimization technique that mimics specific laws of physics to obtain an optimum solution[71]. It is a capable metaheuristic algorithm for dealing with problems with different levels of complexity. The search units of the EO are called particles each of which receives an initial concentration value as in Eq. (1):

$$C_j = LB + r \times (UB - LB), \tag{1}$$

where r is a random value in [0, 1]. Moreover, LB and UB are the lower and upper bounds of the space.

Similar to other population-based optimizers, the quality of the particles is reflected by a fitness value. They are then sorted, and the algorithm hires four of them which are distinguished by the highest fitness value. A fifth particle is also considered that represents the mean of these four particles.

The exponential term (F) of the algorithm is defined by Eqs. (2), (3), (4).
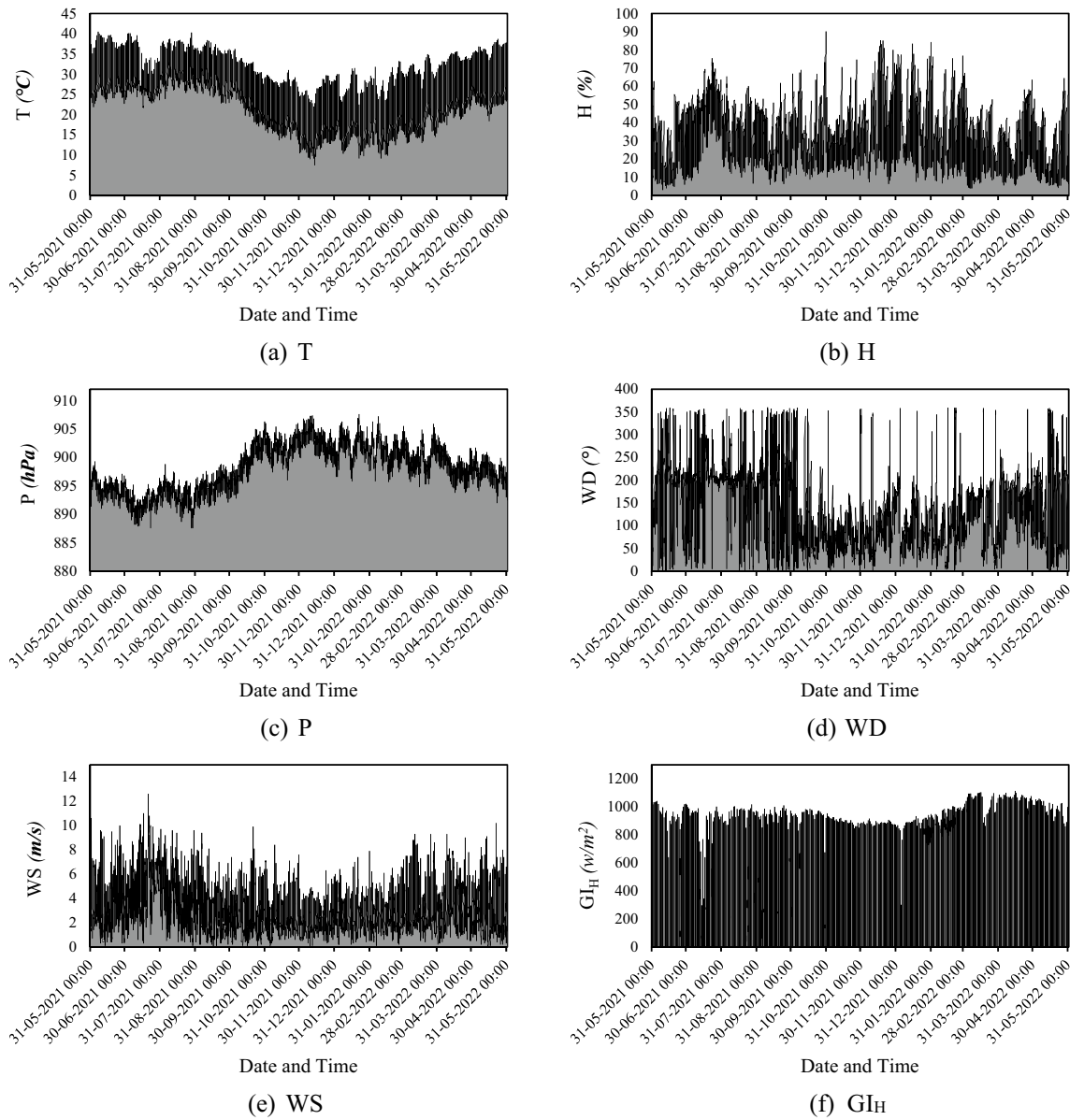
$$F = CP_1 sign(r - 0.5)(e^{-\beta t} - 1), \tag{2}$$

Figure 1. Time-series of the GI_H and meteorological parameters.

	T (°C)	H (%)	P (hPa)	WD (°)	WS (m/s)	GI_H (w/m²)
Mean	25.64	30.64	898.12	132.60	3.38	273.85
Standard error	0.07	0.18	0.04	0.86	0.02	3.80
Median	26.10	28.10	898.30	121.00	3.00	10.00
Mode	28.30	16.30	901.20	56.00	1.90	0.00
Standard deviation	6.78	16.49	3.99	80.83	1.98	356.42
Sample variance	45.94	271.87	15.93	6532.69	3.92	127,036.86
Kurtosis	−0.69	−0.20	−0.88	−0.24	0.50	−0.84
Skewness	−0.15	0.63	−0.10	0.60	0.85	0.88
Range	33.00	87.00	20.00	360.00	12.60	1113.00
Minimum	7.50	3.10	887.60	0.00	0.00	0.00
Maximum	40.50	90.10	907.60	360.00	12.60	1113.00
Sum	225,706.00	269,723.30	7,906,112.80	1,167,256.00	29,760.90	2,410,715.00
Count	8803.00	8803.00	8803.00	8803.00	8803.00	8803.00

Table 1. Correlation factors showing the proportionality of the parameters.

$$t = (1 - R_{iter})^{(CP_2 \times R_{iter})}, \tag{3}$$

$$R_{iter} = \frac{iter}{T}, \tag{4}$$

in which β stands for the turnover rate, CP_1 and CP_2 are controlling parameters for the exploration and exploitation phases, respectively.

Assuming G_{CP} and GP as a controlling parameter and the generation probability, respectively, generation rate is calculated by Eqs. (5) and (6).

$$R_G = G_0 G_{CP} (C_{eq} - \beta C) F, \tag{5}$$

$$G_{CP} = \begin{cases} 0.5r_1 & r_2 \geq GP \\ 0 & otherwise \end{cases}, \tag{6}$$

where C_{eq} is the equilibrium pool, and r_1 and r_2 are random numbers in [0, 1].

Based on the above calculations, the solution is updated as in Eq. (7)[72]:

$$C_j = C_{eq} + (C_j - C_{eq}) \times F + (1 - F)\frac{R_G}{\beta V}, \tag{7}$$

where V is the considered unit.

Comparative algorithms
Wind-driven optimization was first introduced by Bayraktar, Komurcu and Werner[73] in 2010 for electromagnetics applications. The WDO relies on the air parcel's movement in hyper-dimensional space. These movements are supposed to be affected by four natural forces of Coriolis force, gravitational force, frictional force, and pressure gradient force. Also, by taking into consideration the ideal gas equation, the position (as well as the velocity) of the air parcels is updated to find the best responses. Scholars like Moayedi, Bui and Ngo[74] and Bayraktar[75] have successfully used the WDO for optimizing the neural parameters.

As a physic-based scheme, the OIO was suggested by Kashan[76] in 2014. It is inspired by optics (a law in physics) which works by a group of artificial light-related stuff. After randomly generating the fixed number of individuals, the initial position of the light points is determined. Each point is then put in front of an artificial mirror and its image is created in the search space with a certain distance from the main axis. The position of the image is then updated to be a new solution This process continues until a stopping criterion is satisfied[77].

The SOSA, as the name implies, takes the idea from the food-seeking action of social spider, introduced by James and Li[78] in 2015. In this method, the solution space is considered a hyper-dimensional spider web that the agents (i.e., spiders) can move on it. As assumptions, the agents have regular interaction with each other and every position in this area corresponds to a possible solution[79]. Each spider distinguishes itself by the position and fitness value. The agents possess a memory to hold three basic attributes: all possible vibration intensities are positive, (ii) the larger the fitness values mean more intense vibrations, and (iii) once the best solution is getting close, the vibration does not experience excessive increase.

Mathematical details pertaining to the above algorithms can be found in the literature (like the WDO[60,80], OIO[81,82], and SOSA[83,84]).

Evaluation method
Statistical indices are normally used for evaluating the accuracy of ML models. In this work, RMSE along with Mean Absolute Error (MAE) is used to indicate the prediction error. Given $GI_{Hi_{real}}$ and $GI_{Hi_{predict}}$ as the real and predicted GI_Hs, respectively, Eqs. (8) and (9) formulate the RMSE and MAE as follows:

$$RMSE = \sqrt{\frac{1}{S} \sum_{i=1}^{S} (GI_{Hi_{real}} - GI_{Hi_{predict}})^2}, \tag{8}$$

$$MAE = \frac{1}{S} \sum_{i=1}^{S} \left| GI_{Hi_{real}} - GI_{Hi_{predict}} \right|, \tag{9}$$

where S stands for the size of the set.

Moreover, a so-called correlation indicator "Pearson Correlation Coefficient (R)" is designated as per Eq. (10) to reflect the agreement between reality and prediction.

$$R = \frac{\sum_{i=1}^{S}(GI_{Hi_{predict}} - \overline{GI_{H_{predict}}})(GI_{Hi_{real}} - \overline{GI_{H_{real}}})}{\sqrt{\sum_{i=1}^{S}(GI_{Hi_{predict}} - \overline{GI_{H_{predict}}})^2}\sqrt{\sum_{i=1}^{S}(GI_{Hi_{real}} - \overline{GI_{H_{real}}})^2}} \times 100, \tag{10}$$

Results and discussion
Hybridization of algorithms
When the FFANN is hybridized with a metaheuristic algorithm, the basic idea is to optimize its weights and biases to establish the best relationship between the target and input parameters. In this research, the FFANN is optimized by the EO algorithm, as well as OIO, WDO, and SOSA. The metaheuristic algorithms are able to find the solution in an iterative process.

The used FFANN is represented by an MLPNN (8,6,1) model indicating a three-layered neural network with 8 input neurons in the first layer, 6 hidden neurons in the middle layer, and 1 output neuron in the last layer. The activation functions in the middle and last layers are Tansig and Purelin, respectively. This configuration is obtained after an extensive trial-and-error effort. A topology of the used FFANN is embedded in Fig. 2. According to this figure, this network has a total of 61 weights and biases which are optimized by the metaheuristic algorithm. In this process, the training dataset is used solely. First, the mathematical equation of the FFANN is extracted and is considered as the problem function. Next, a metaheuristic algorithm is run to tune the FFANN equation (i.e., weights and biases) so that the training RMSE is minimized by 1000 iterations. In each iteration, new 61 variables construct the FFANN, and the training RMSE is calculated. Note that, each of the EO, OIO, WDO, and SOSA algorithms were implemented with different population sizes (varying from 50 to 700) and it was observed that the best population size for them is 400, 200, 100, and 200, respectively.

Optimization results
Figure 3 shows the optimization process for the used algorithms (with the mentioned best population sizes) that are iterated 1000 times. From the comparison of the curves, it is immediate that the EO has reached a higher quality of solution due to the minimum RMSE error. While the solutions of the WDO and OIO are very close, the SOSA has found the solution with considerably higher error. Note that this process was carried out using the training set only because the testing set should be kept away from the models in this stage. In the next two sections, the training and testing results are assessed using the accuracy methods.

Training accuracy
Figure 4 forms part of the training results as the final RMSEs are the training RMSEs. Having the order of algorithms as OIO-FFANN, WDO-FFANN, SOSA-FFANN, and EO-FFANN, training RMSEs were 161.22, 152.16, 230.61, and 142.38 w/m^2.

Figure 4 illustrates the error values for the 7042 records in the training set. Each single value in this figure indicates the pure difference between $GI_{Hi_{real}}$ and $GI_{Hi_{predict}}$. Hence, the lower the value, the higher the accuracy. In a glance, it can be seen that the results of the EO are better positioned around the Y = 0 line. Quantitatively speaking, the training MAEs were 127.52, 119.33, 189.56, and 110.09 w/m^2.

The calculated values of the RMSE and MAE indicated an acceptable level of error for all used models. As for the R index, the values were 0.89, 0.90, 0.76, and 0.91 which demonstrate a significant level of agreement between the reality and prediction results of all four models. However, again, the superiority of the EO algorithm is obvious in terms of the R, too. It was the only model that achieved a correlation larger than 90%.
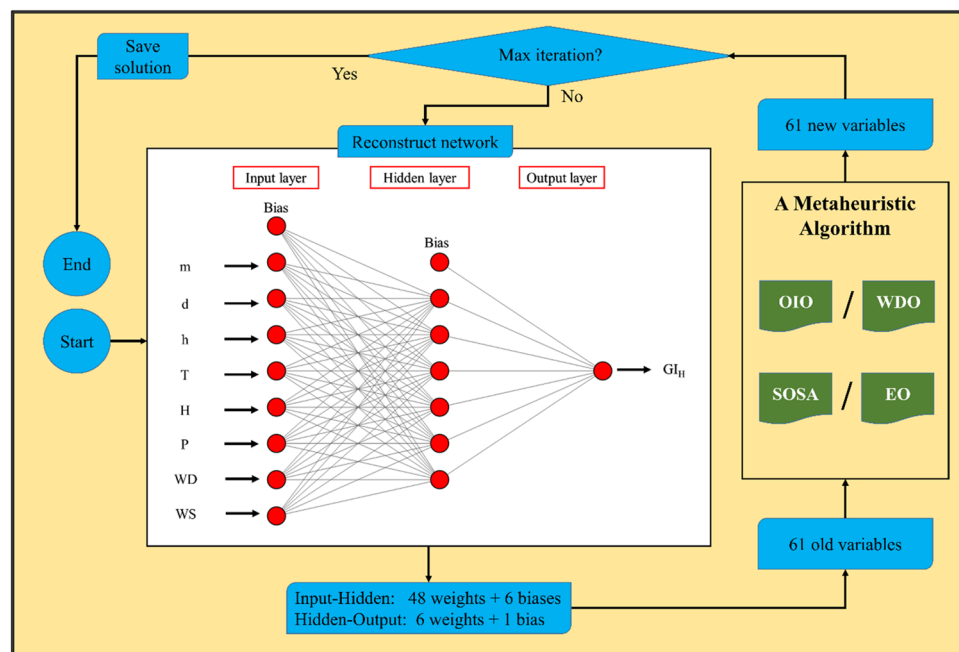
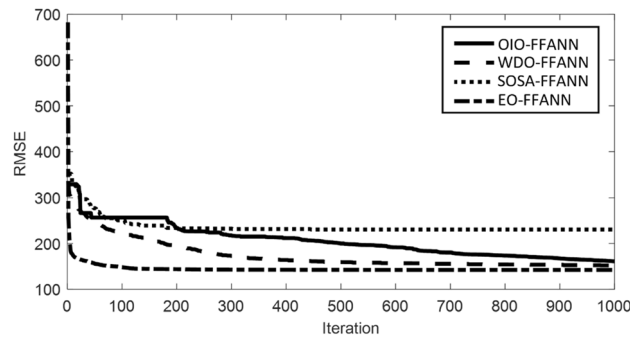

Figure 2. Optimization process of the hybrid models.

Figure 3. Optimization curves of the OIO-FFANN, WDO-FFANN, SOSA-FFANN, and EO-FFANN algorithms.
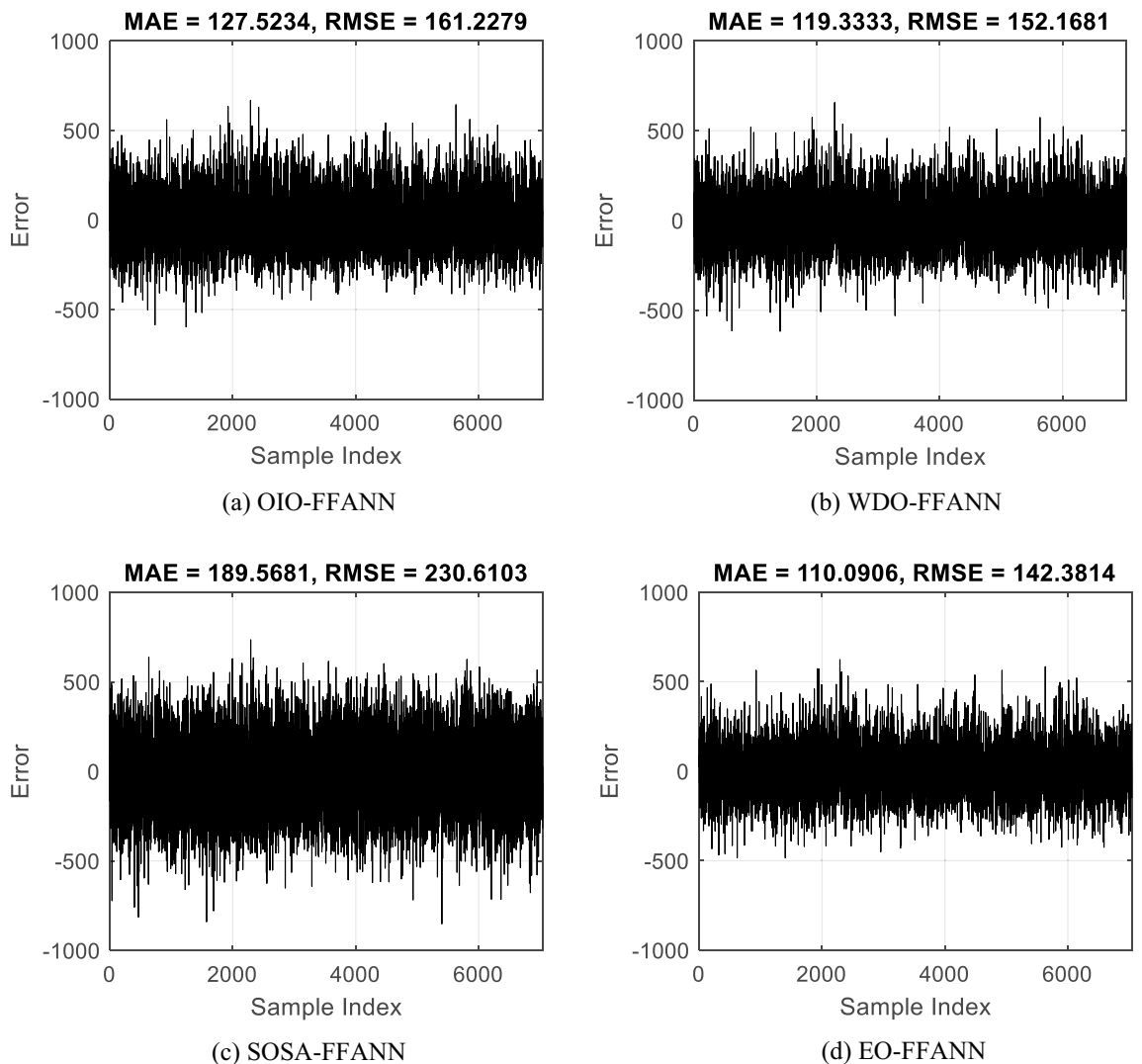

(a) OIO-FFANN

(b) WDO-FFANN

(c) SOSA-FFANN

(d) EO-FFANN

Figure 4. Training errors of (**a**) OIO-FFANN, (**b**) WDO-FFANN, (**c**) SOSA-FFANN, and (**d**) EO-FFANN.

Testing accuracy

This section shows the performance of the OIO-FFANN, WDO-FFANN, SOSA-FFANN, and EO-FFANN when they are subjected to the 1761 records in the testing set. This process demonstrates the power of the trained models in dealing with unseen environmental conditions for estimating hourly GI_H.

From the obtained RMSEs of 161.63, 151.57, 230.16, and 141.61 w/m^2, it is quantitatively inferred that the testing results enjoy a satisfying level of accuracy. Figure 5 illustrates the statistics of the testing errors. In these
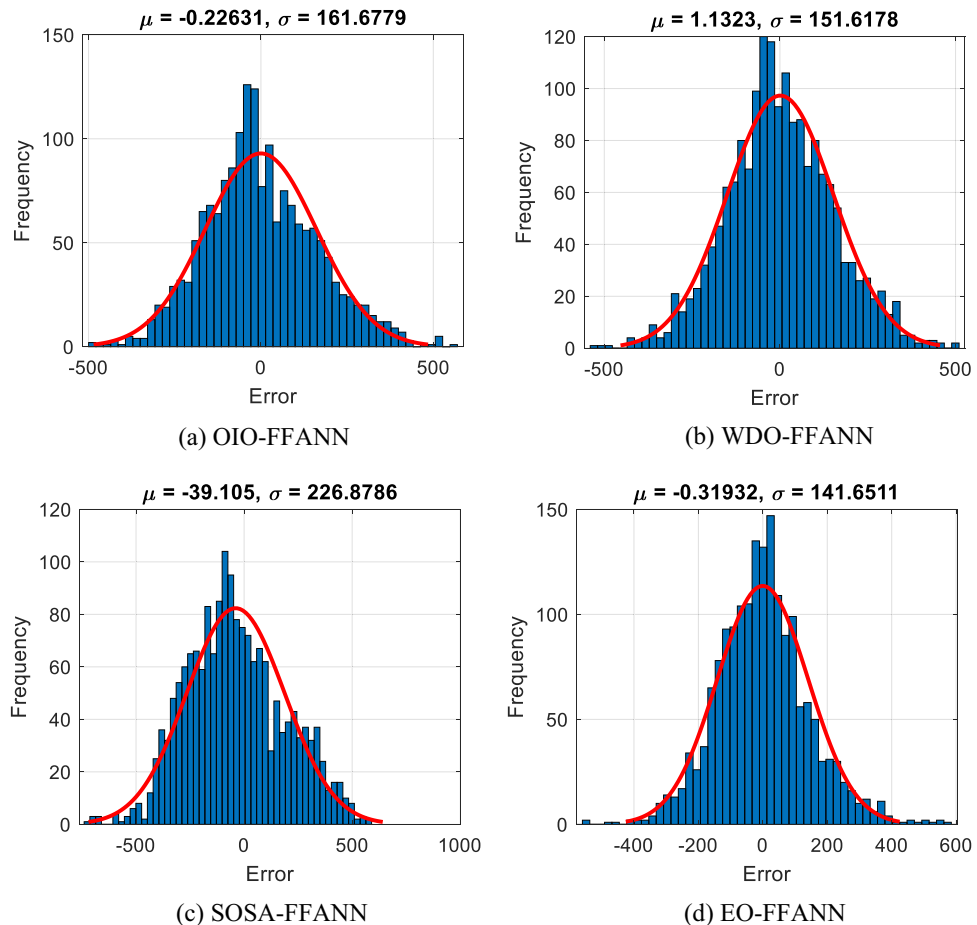
Figure 5. Testing histogram of errors of (**a**) OIO-FFANN, (**b**) WDO-FFANN, (**c**) SOSA-FFANN, and (**d**) EO-FFANN.

histogram charts, the higher the frequency of 0 error, the better the accuracy. As is seen, the distribution is almost normal for all models. It professes the high quality of testing results. Besides, the MAEs of 127.52, 118.72, 188.20, and 108.07 w/m^2 indicate a low level of average errors.

Figure 6 shows the correlation diagrams of the testing set. The values on the horizontal and vertical axis represent the $GI_{Hi_{real}}$ and $GI_{Hi_{predict}}$, respectively. Hereupon, the ideal prediction happens when all points are positioned on the line x = y, and the R-value is 1. As per Fig. 6, all four models have performed a nice prediction and their calculated Rs were 0.89, 0.90, 0.77, and 0.91. Similar to the training stage, EO-FFANN is the only algorithm with a correlation above 90%.

Accuracy comparison

It was in general shown that the EO-based model outperformed the benchmarks in both training and testing stages. In this section, the models are subjected to a more detailed comparison to rank them. For this purpose, Taylor diagrams are generated and presented in Fig. 7. These figures can simultaneously show the correlation (Correlation Coefficient) and error (RMSD = RMSE). As is seen, in both training and testing sets, the same pattern is obtained, and it means there is no discrepancy between the training and testing qualities. The EO-FFANN is distinguished by the lowest error and highest correlation, followed by WDO-FFANN and OIO-FFANN. As for SOSA-FFANN, this model has a considerable weakness in its performance in comparison with three other models. As per Fig. 7, the point of the SOSA-FFANN is separated from the others.

PCA importance analysis

In this section, an importance assessment is applied to the used dataset. The results of such efforts can be of great importance for the proper selection of input factors from the statistical point of view. The PCA technique[85] is used to determine the most contributive factors for the GI_H prediction. Figure 8 shows the obtained scree plot, according to which, four components have an eigenvalue larger than 1. These four components are considered as principal components, and based on Table 2, cumulatively account for about 75% of the variance in the dataset.

Table 3 shows the results of the Varimax rotation method. In each of the four components, the factors with loading above + 0.70 and below −0.70 are selected. As is seen, Component 1 reflects T and H, Component 2
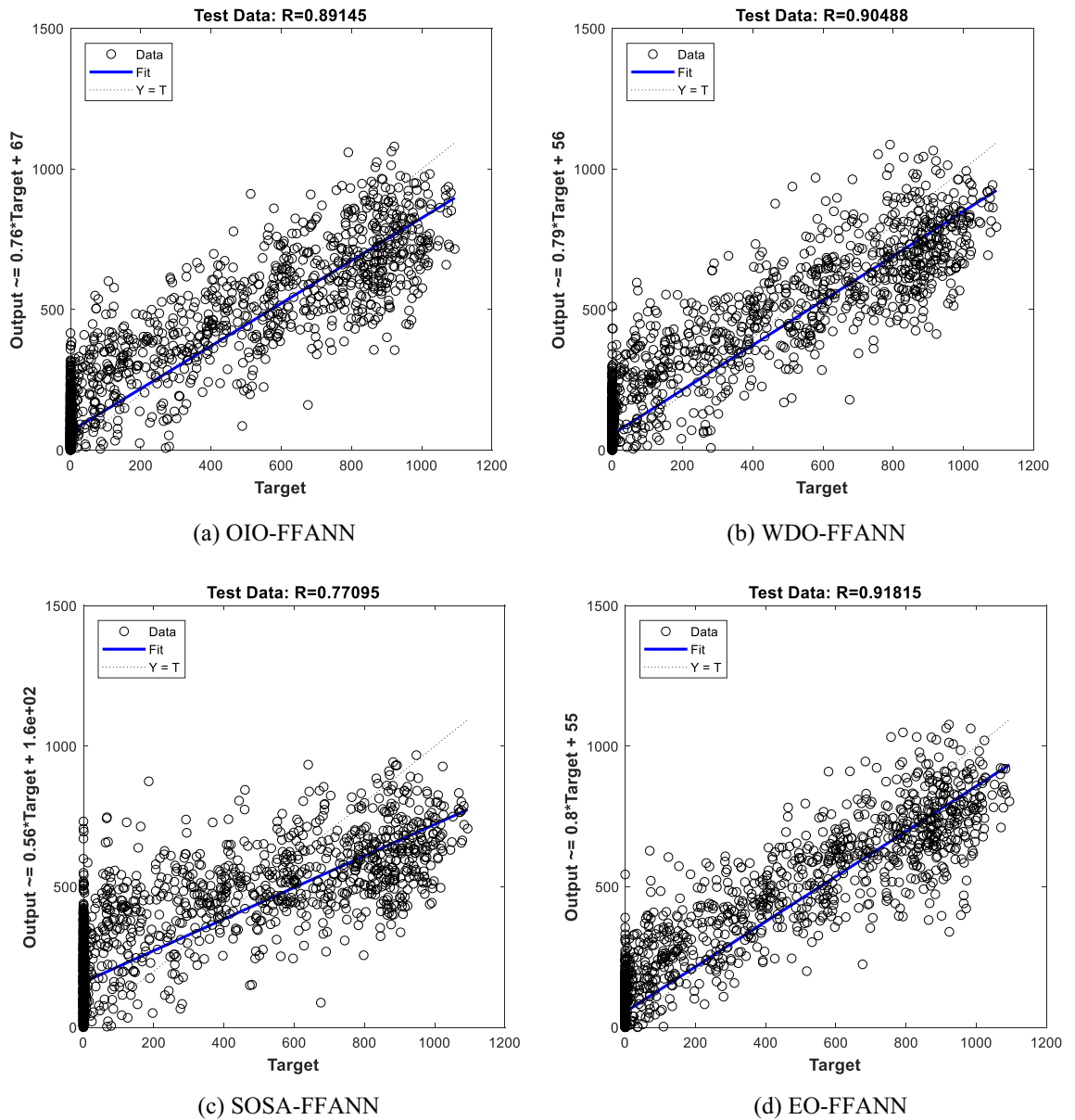
Figure 6. Correlation diagrams for testing set of (**a**) OIO-FFANN, (**b**) WDO-FFANN, (**c**) SOSA-FFANN, and (**d**) EO-FFANN.

reflects P and WD, Component 3 reflects d, and Component 4 reflects m. Hence, it can be deduced that h and WS can be discarded for optimizing the dataset.

A monolithic formula

In order to eliminate the need for implementing computer-based programs, this section provides a mathematical expression that is derived from the proposed model i.e., EO-FFANN, for predicting the GI_H. The reason for considering EO-FFANN is that this model achieved the highest accuracy in the previous assessments. The formula is a monolithic relationship; however, it has two steps and the GI_H needs to be calculated in the second step.

Referring to the FFANN topology in Fig. 2, this equation is constructed from 61 weights and biases. The general inputs of this equation are m, d, h, T, H, P, WD, and WS that feed Eq. (11). With the help of Table 4, the outcomes of this equation are N_i (i = 1, 2, …, 6) that feed Eq. (12) for calculating the GI_H. In other words, Eq. (11) and Table 4 together express the process between the input and hidden layers of the FFANN, while Eq. (12) expresses the process between the hidden and output layers (see Fig. 2).

$$N_i = Tansig(W_{i1} \times m + W_{i2} \times d + W_{i3} \times h + W_{i4} \times T + W_{i5} \times H + W_{i6} \times P + W_{i7} \times WD + W_{i8} \times WS + b_i),$$
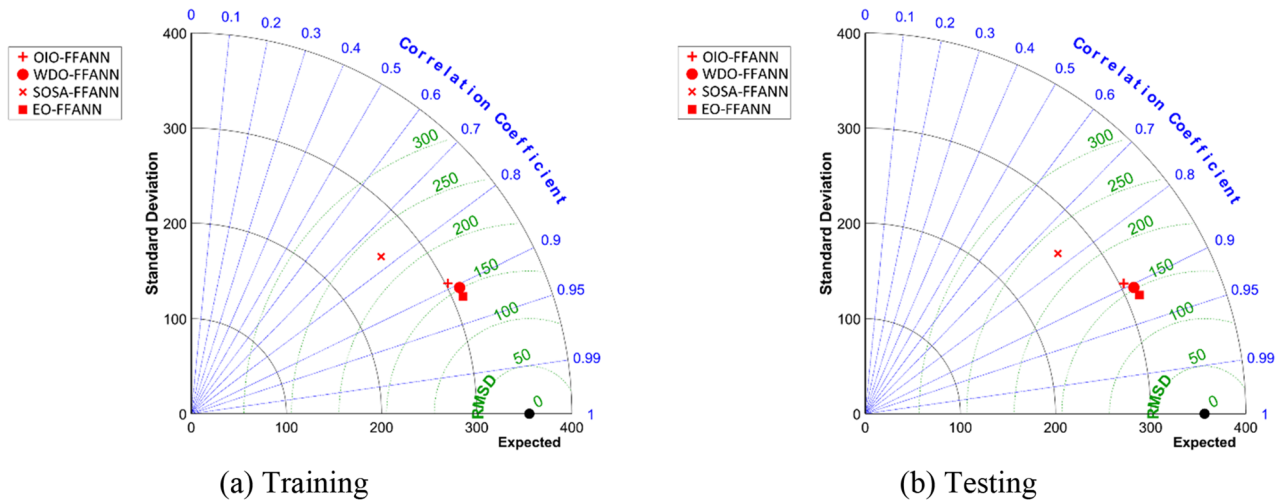
$$(11)$$

(a) Training

(b) Testing

Figure 7. Taylor diagrams for (**a**) training and (**b**) testing sets.
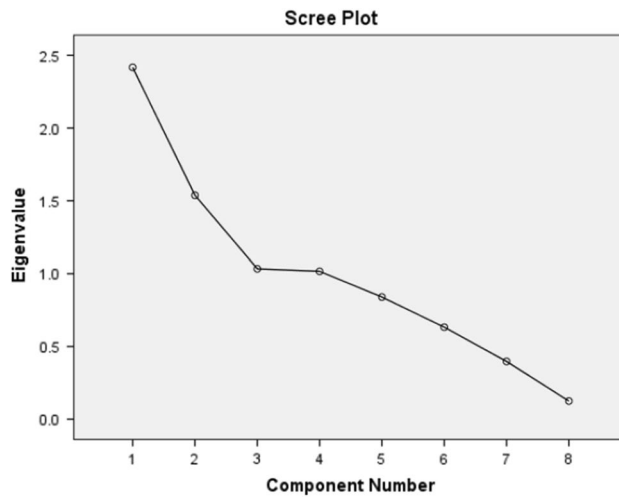

Figure 8. Scree plot of the PCA analysis.

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.417	30.213	30.213	2.417	30.213	30.213	2.141	26.764	26.764
2	1.537	19.218	49.431	1.537	19.218	49.431	1.796	22.451	49.215
3	1.034	12.919	62.350	1.034	12.919	62.350	1.040	12.999	62.214
4	1.016	12.706	75.056	1.016	12.706	75.056	1.027	12.842	75.056
5	0.840	10.501	85.557						
6	0.633	7.910	93.466						
7	0.397	4.967	98.434						
8	0.125	1.566	100.000						

Table 2. Total variance explained.

$$GI_H = 0.4187 \times N_1 + 0.5094 \times N_2 - 0.4479 \times N_3 + 0.3594 \times N_4$$
$$+ 0.3102 \times N_5 - 0.6748 \times N_6 - 0.7620, \tag{12}$$

	Component			
	1	2	3	4
m	−0.039	0.019	0.030	**0.983**
d	0.031	−0.012	**0.936**	0.032
h	0.655	−0.138	0.126	−0.039
T	**0.720**	0.565	−0.110	0.162
H	*−0.736*	0.094	0.316	−0.042
P	−0.301	*−0.868*	−0.010	−0.093
WD	−0.346	**0.794**	0.012	−0.099
WS	0.663	0.255	0.189	−0.113

Table 3. Rotated component matrix from Varimax method. Significant values bold and italics.

i	W_{i1}	W_{i2}	W_{i3}	W_{i4}	W_{i5}	W_{i6}	W_{i7}	W_{i8}	b_i
1	0.6970	−0.4906	1.0125	0.6472	0.3959	0.4563	0.4315	0.5881	−1.7514
2	0.8025	0.0927	−0.0289	0.9087	0.5097	−0.9258	−0.3617	0.5838	−1.0509
3	−0.6994	0.8579	0.5630	0.2920	0.4559	−0.4183	0.8442	−0.5871	0.3503
4	0.6707	0.7543	−0.5810	−0.7533	−0.1749	−0.7363	−0.7554	−0.0166	0.3503
5	0.3433	−0.8881	−0.2029	0.9056	0.4033	−1.0450	−0.1589	−0.1411	1.0509
6	−0.7074	0.8272	0.7308	0.7629	−0.5780	0.5686	−0.2082	0.2572	−1.7514

Table 4. Optimized internal parameters of the FFANN.

Strength, limitations, and future guidelines

This study presented novel applications of metaheuristic-empowered ML models for predicting I_S. A valid dataset with various meteorological and temporal factors was applied for this purpose. The models were optimized in terms of their hyper-parameters such as the FFANN topology and population size of the metaheuristic algorithms. Therefore, it can be claimed that the used models are among the most optimum ones. The desirable level of accuracy obtained in this study proved the applicability of the applied models, however, a comparison showed that the EO-FFANN shows greater promise. This model achieved improvement when it is compared to previous studies. For instance, Moayedi and Mosavi[49] used the EFO algorithm, along with the SCE and SFLA, to optimize a similar FFANN. These models reached an R-value (non-percentage) of 0.82132, 0.78046, and 0.75212, respectively, which are lower than the R values of the EO-FFANN in this work.

Presenting a simplified formula is another outcome of this study which enables the users to predict the GI_H without the need for computer-aided facilities. Furthermore, regarding the performed trial and error efforts in different stages, it can be said that this solution is captured carefully among numerous candidates.

In Sect. 3.6, the PCA model was applied to the dataset and its results highlighted the T, H, P, WD, d, and m as the most contributive input factors. As is known, reducing the dataset inputs from 8 to 6 results in lightening the computational burden due to the reduction in the problem dimension[86]. Considering this idea is highly recommended for future efforts towards improving the solution for the GI_H prediction.

However, this study encountered some limitations, too. About the used dataset, it includes the records from 2021-05-31 to 2022-06-01. Hence, updating this dataset with the most recent data (e.g., late 2022 and early 2023) could be of great interest to future efforts. It may help in enhancing the generalizability of the suggested models for new climate conditions. As far as the models are concerned, although the applied metaheuristic algorithms are among the recent members of this family, more algorithms have been developed lately. Comparing the results of the EO with the most recent metaheuristic algorithms would greatly help in updating the solutions, and probably, increasing the accuracy of GI_H prediction.

Conclusions

The importance of analyzing environmental conditions is evident in the forecast of renewable energy potentials. This work was dedicated to optimizing solar energy simulation using state-of-the-art ML and feature selection strategies. An FFANN was optimally trained using different metaheuristic algorithms for predicting solar irradiance from meteorological and temporal parameters (including humidity, temperature, pressure, cloud coverage, speed and direction of wind, month, day, and hour). Assessing the prediction results revealed that the EO performs more accurately than the three optimization algorithms evaluated in this study (OIO, WDO, and SOSA), as well as three optimization algorithms (EFO, SCE, and SFLA) from the earlier literature. Therefore, the mathematical representation of the EO-FFANN was presented in the form of a predictive formula to be reliably used for practical GI_H predictions. Moreover, the PCA method could successfully analyze the datasets and address the T, H, P, WD, d, and m as the input factors that are most essential in forecasting solar irradiance. This part of the results can be regarded in the real world for enhancing the generation of solar energy. In conclusion,

the findings of this study professed the efficiency of the PCA and metaheuristic techniques for optimizing the performance of ML models. However, some ideas were presented for future work toward coping with the limitations of the study, most notably updating the used dataset and predictive models.

Data availability

All data analysed during this study can be downloaded from the Solcast community (https://solcast.com/). Also, the used codes will be made available upon reasonable request from the authors.

References

1. Wu, Z. et al. Effect of dielectric relaxation of epoxy resin on dielectric loss of medium-frequency transformer. *IEEE Trans. Dielectr. Electr. Insul.* **29**, 1651–1658 (2022).
2. Lu, S. et al. Adaptive control of time delay teleoperation system with uncertain dynamics. *Front. Neurorobot.* **16**, 928863. https://doi.org/10.3389/fnbot.2022.928863 (2022).
3. Lu, S. et al. Multiscale feature extraction and fusion of image and text in VQA. *Int. J. Comput. Intell. Syst.* **16**, 54. https://doi.org/10.1007/s44196-023-00233-6 (2023).
4. Lu, S. et al. The multi-modal fusion in visual question answering: a review of attention mechanisms. *PeerJ Comput. Sci.* **9**, e1400. https://doi.org/10.7717/peerj-cs.1400 (2023).
5. Zheng, H. et al. A multi-scale point-supervised network for counting maize tassels in the wild. *Plant Phenomics* (2023).
6. Cheng, B. et al. Situation-aware dynamic service coordination in an IoT environment. *IEEE/ACM Trans. Netw.* **25**, 2082–2095. https://doi.org/10.1109/TNET.2017.2705239 (2017).
7. Yang, X., Wang, X., Wang, S. & Puig, V. Switching-based adaptive fault-tolerant control for uncertain nonlinear systems against actuator and sensor faults. *J. Franklin Inst.* **360**, 11462–11488 (2023).
8. Jiang, J., Zhang, L., Wen, X., Valipour, E. & Nojavan, S. Risk-based performance of power-to-gas storage technology integrated with energy hub system regarding downside risk constrained approach. *Int. J. Hydrogen Energy* **47**, 39429–39442 (2022).
9. Li, S. & Liu, Z. Scheduling uniform machines with restricted assignment. *Math. Biosci. Eng* **19**, 9697–9708 (2022).
10. Zhang, J. et al. Enhanced efficiency with CDCA co-adsorption for dye-sensitized solar cells based on metallosalophen complexes. *Solar Energy* **209**, 316–324 (2020).
11. Huang, S., Huang, M. & Lyu, Y. Seismic performance analysis of a wind turbine with a monopile foundation affected by sea ice based on a simple numerical method. *Eng. Appl. Comput. Fluid Mech.* **15**, 1113–1133 (2021).
12. Chen, H., Wu, H., Kan, T., Zhang, J. & Li, H. Low-carbon economic dispatch of integrated energy system containing electric hydrogen production based on VMD-GRU short-term wind power prediction. *Int. J. Electrical Power Energy Syst.* **154**, 109420 (2023).
13. Zhu, D. et al. Feedforward Frequency Deviation Control in PLL for Fast Inertial Response of DFIG-Based Wind Turbines. *IEEE Trans. Power Electronics* (2023).
14. Gong, J., Li, C. & Wasielewski, M. R. Advances in solar energy conversion. *Chem. Soc. Rev.* **48**, 1862–1864. https://doi.org/10.1039/C9CS90020A (2019).
15. Zhang, W., Zheng, Z. & Liu, H. Droop control method to achieve maximum power output of photovoltaic for parallel inverter system. *CSEE J. Power Energy Syst.* **8**, 1636–1645 (2021).
16. Blal, M. et al. A prediction models for estimating global solar radiation and evaluation meteorological effect on solar radiation potential under several weather conditions at the surface of Adrar environment. *Measurement* **152**, 107348. https://doi.org/10.1016/j.measurement.2019.107348 (2020).
17. Chakchak, J. & Cetin, N. S. Investigating the impact of weather parameters selection on the prediction of solar radiation under different genera of cloud cover: A case-study in a subtropical location. *Measurement* **176**, 109159 (2021).
18. Mishra, M., Dash, P. B., Nayak, J., Naik, B. & Swain, S. K. Deep learning and wavelet transform integrated approach for short-term solar PV power prediction. *Measurement* **166**, 108250 (2020).
19. Patel, S. K. et al. Ultra-Wideband, Polarization-Independent, Wide-Angle Multilayer Swastika-Shaped Metamaterial Solar Energy Absorber with Absorption Prediction using Machine Learning. *Adv. Theory Simul.*, 2100604 (2022).
20. Zazoum, B. Solar photovoltaic power prediction using different machine learning methods. *Energy Rep.* **8**, 19–25 (2022).
21. Patel, D., Patel, S., Patel, P. & Shah, M. Solar radiation and solar energy estimation using ANN and Fuzzy logic concept: A comprehensive and systematic study. *Environ. Sci. Pollut. Res.*, 1–15 (2022).
22. Heng, S. Y. et al. Artificial neural network model with different backpropagation algorithms and meteorological data for solar radiation prediction. *Sci. Rep.* **12**, 1–18 (2022).
23. Ghimire, S., Nguyen-Huy, T., Deo, R. C., Casillas-Perez, D. & Salcedo-Sanz, S. Efficient daily solar radiation prediction with deep learning 4-phase convolutional neural network, dual stage stacked regression and support vector machine CNN-REGST hybrid model. *Sustain. Mater. Technol.* **32**, e00429 (2022).
24. Guermoui, M., Abdelaziz, R., Gairaa, K., Djemoui, L. & Benkaciali, S. New temperature-based predicting model for global solar radiation using support vector regression. *Int. J. Ambient Energy* **43**, 1397–1407 (2022).
25. Singh, N., Jena, S. & Panigrahi, C. K. A novel application of decision tree classifier in solar irradiance prediction. *Mater. Today: Proc.* **58**, 316–323 (2022).
26. Shorabeh, S. N. et al. A decision model based on decision tree and particle swarm optimization algorithms to identify optimal locations for solar power plants construction in Iran. *Renew. Energy* **187**, 56–67 (2022).
27. Guerra, M. I., de Araújo, F. M., de Carvalho Neto, J. T. & Vieira, R. G. Survey on adaptative neural fuzzy inference system (ANFIS) architecture applied to photovoltaic systems. *Energy Syst.*, 1–37 (2022).
28. Fraihat, H. et al. Solar radiation forecasting by pearson correlation using LSTM neural network and ANFIS method: application in the west-central Jordan. *Future Internet* **14**, 79 (2022).
29. Moayedi, H. et al. Novel hybrids of adaptive neuro-fuzzy inference system (ANFIS) with several metaheuristic algorithms for spatial susceptibility assessment of seismic-induced landslide. *Geomatics, Natl. Hazards Risk* **10**, 1879–1911. https://doi.org/10.1080/19475705.2019.1650126 (2019).
30. Jahanafroozi, N. et al. New heuristic methods for sustainable energy performance analysis of HVAC systems. *Sustainability* **14**, 14446 (2022).
31. Liu, L., Zhang, S., Zhang, L., Pan, G. & Yu, J. Multi-UUV maneuvering counter-game for dynamic target scenario based on fractional-order recurrent neural network. *IEEE Trans. Cybern.* (2022).
32. Luo, J., Wang, G., Li, G. & Pesce, G. Transport infrastructure connectivity and conflict resolution: a machine learning analysis. *Neural Comput. Appl.* **34**, 6585–6601 (2022).

33. Cheng, Y. *et al.* A dual-branch weakly supervised learning based network for accurate mapping of woody vegetation from remote sensing images. *Int. J. Appl. Earth Obs. Geoinf.* **124**, 103499 (2023).
34. Kim, J.-H., Seong, N.-C. & Choi, W. Forecasting the Energy Consumption of an Actual Air Handling Unit and Absorption Chiller Using ANN Models. *Energies* **13**, 4361 (2020).
35. Bhatt, G. A. & Gandhi, P. R. in *3rd International Conference on Trends in Electronics and Informatics (ICOEI)*. 622–627.
36. Yin, Z., Jia, B., Wu, S., Dai, J. & Tang, D. Comprehensive forecast of urban water-energy demand based on a neural network model. *Water* **10**, 385 (2018).
37. Malvoni, M., De Giorgi, M. G. & Congedo, P. M. Data on support vector machines (SVM) model to forecast photovoltaic power. *Data Brief* **9**, 13–16. https://doi.org/10.1016/j.dib.2016.08.024 (2016).
38. Tian, H., Li, R., Salah, B. & Thinh, P.-H. Bi-objective optimization and environmental assessment of SOFC-based cogeneration system: Performance evaluation with various organic fluids. *Process Saf. Environ. Prot.* **178**, 311–330 (2023).
39. Li, R., Xu, D., Tian, H. & Zhu, Y. Multi-objective study and optimization of a solar-boosted geothermal flash cycle integrated into an innovative combined power and desalinated water production process: Application of a case study. *Energy* **282**, 128706 (2023).
40. Zhang, Z., Altalbawy, F. M., Al-Bahrani, M. & Riadi, Y. Regret-based multi-objective optimization of carbon capture facility in CHP-based microgrid with carbon dioxide cycling. *J. Clean. Prod.* **384**, 135632 (2023).
41. Afzal, A. *et al.* Optimizing the thermal performance of solar energy devices using meta-heuristic algorithms: A critical review. *Renew. Sustain. Energy Rev.* **173**, 112903 (2023).
42. Stoean, C. *et al.* Metaheuristic-based hyperparameter tuning for recurrent deep learning: Application to the prediction of solar energy generation. *Axioms* **12**, 266. https://doi.org/10.3390/axioms12030266 (2023).
43. Alkhazaleh, H. A. *et al.* Prediction of thermal energy demand using fuzzy-based models synthesized with metaheuristic algorithms. *Sustainability* **14**, 14385 (2022).
44. Houssein, E. H. in *Advanced Control and Optimization Paradigms for Wind Energy Systems* 165–187 (Springer, 2019).
45. Corizzo, R., Ceci, M., Fanaee-T, H. & Gama, J. Multi-aspect renewable energy forecasting. *Inf. Sci.* **546**, 701–722. https://doi.org/10.1016/j.ins.2020.08.003 (2021).
46. Bessa, R. J., Trindade, A., Silva, C. S. P. & Miranda, V. Probabilistic solar power forecasting in smart grids using distributed information. *Int. J. Electrical Power Energy Syst.* **72**, 16–23. https://doi.org/10.1016/j.ijepes.2015.02.006 (2015).
47. Liu, H., Chen, C., Lv, X., Wu, X. & Liu, M. Deterministic wind energy forecasting: A review of intelligent predictors and auxiliary methods. *Energy Convers. Manag.* **195**, 328–345 (2019).
48. Cavalcante, L., Bessa, R. J., Reis, M. & Browell, J. LASSO vector autoregression structures for very short-term wind power forecasting. *Wind Energy* **20**, 657–675. https://doi.org/10.1002/we.2029 (2017).
49. Moayedi, H. & Mosavi, A. An innovative metaheuristic strategy for solar energy management through a neural networks framework. *Energies* **14**, 1196 (2021).
50. Abedinia, O., Amjady, N. & Ghadimi, N. Solar energy forecasting based on hybrid neural network and improved metaheuristic algorithm. *Comput. Intell.* **34**, 241–260 (2018).
51. Galván, I. M., Valls, J. M., Cervantes, A. & Aler, R. Multi-objective evolutionary optimization of prediction intervals for solar energy forecasting with neural networks. *Inf. Sci.* **418**, 363–382 (2017).
52. Tran, D.-H., Luong, D.-L. & Chou, J.-S. Nature-inspired metaheuristic ensemble model for forecasting energy consumption in residential buildings. *Energy* **191**, 116552 (2020).
53. Halabi, L. M., Mekhilef, S. & Hossain, M. Performance evaluation of hybrid adaptive neuro-fuzzy inference system models for predicting monthly global solar radiation. *Appl. Energy* **213**, 247–261 (2018).
54. Zhao, Y., Moayedi, H., Bahiraei, M. & Foong, L. K. Employing TLBO and SCE for optimal prediction of the compressive strength of concrete. *Smart Struct. Syst.* **26**, 753. https://doi.org/10.12989/sss.2020.26.6.753 (2020).
55. Louzazni, M. *et al.* Metaheuristic algorithm for photovoltaic parameters: comparative study and prediction with a firefly algorithm. *Appl. Sci.* **8**, 339 (2018).
56. Bechouat, M. *et al.* Parameters identification of a photovoltaic module in a thermal system using meta-heuristic optimization methods. *Int. J. Energy Environ. Eng.* **8**, 331–341 (2017).
57. Zhou, G., Moayedi, H. & Foong, L. K. Teaching–learning-based metaheuristic scheme for modifying neural computing in appraising energy performance of building. *Eng. Comput.* **37**, 3037–3048. https://doi.org/10.1007/s00366-020-00981-5 (2021).
58. Vaisakh, T. & Jayabarathi, R. Analysis on intelligent machine learning enabled with meta-heuristic algorithms for solar irradiance prediction. *Evolutionary Intelligence*, 1–20 (2020).
59. Abedinia, O., Amjady, N. & Ghadimi, N. Solar energy forecasting based on hybrid neural network and improved metaheuristic algorithm. *Comput. Intell.* **34**(1), 241–260. https://doi.org/10.1111/coin.12145 (2018).
60. Abdalla, O., Rezk, H. & Ahmed, E. M. Wind driven optimization algorithm based global MPPT for PV system under non-uniform solar irradiance. *Solar Energy* **180**, 429–444 (2019).
61. Lu, C. *et al.* Split-core magnetoelectric current sensor and wireless current measurement application. *Measurement* **188**, 110527 (2022).
62. Li, P., Hu, J., Qiu, L., Zhao, Y. & Ghosh, B. K. A distributed economic dispatch strategy for power–water networks. *IEEE Trans. Control Netw. Syst.* **9**, 356–366 (2021).
63. Zheng, X. *et al.* Combustion characteristics and thermal decomposition mechanism of the flame-retardant cable in urban utility tunnel. *Case Stud. Thermal Eng.* **44**, 102887 (2023).
64. Zhang, L. *et al.* Development of geopolymer-based composites for geothermal energy applications. *J. Clean. Product.* **419**, 138202 (2023).
65. Cheng, B., Zhu, D., Zhao, S. & Chen, J. Situation-aware IoT service coordination using the event-driven SOA paradigm. *IEEE Trans. Netw. Serv. Manag.* **13**, 349–361 (2016).
66. Alkabbani, H., Ahmadian, A., Zhu, Q. & Elkamel, A. Machine learning and metaheuristic methods for renewable power forecasting: a recent review. *Front. Chem. Eng.* **3**, 665415. https://doi.org/10.3389/fceng.2021.665415 (2021).
67. Gumar, A. K. & Demir, F. Solar photovoltaic power estimation using meta-optimized neural networks. *Energies* **15**, 8669 (2022).
68. Aissaoui, A., Belhaouas, N., Hadjrioua, F., Bakria, K. & Aloui, I. in *Artif. Intell. Renew. Energy Trans.* 4, 592–603 (Springer).
69. Feng, Y., Cui, N., Chen, Y., Gong, D. & Hu, X. Development of data-driven models for prediction of daily global horizontal irradiance in northwest China. *J. Clean. Product.* **223**, 136–146 (2019).
70. Zang, H. *et al.* Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations. *Renew. Energy* **160**, 26–41 (2020).
71. Faramarzi, A., Heidarinejad, M., Stephens, B. & Mirjalili, S. Equilibrium optimizer: A novel optimization algorithm. *Knowl.-Based Syst.* **191**, 105190. https://doi.org/10.1016/j.knosys.2019.105190 (2020).
72. Elsheikh, A. H., Shehabeldeen, T. A., Zhou, J., Showaib, E. & Abd Elaziz, M. Prediction of laser cutting parameters for polymethylmethacrylate sheets using random vector functional link network integrated with equilibrium optimizer. *J. Intell. Manuf.* https://doi.org/10.1007/s10845-020-01617-7 (2020).
73. Bayraktar, Z., Komurcu, M. & Werner, D. H. in *2010 IEEE Antennas and Propagation Society International Symposium*. 1–4 (IEEE).
74. Moayedi, H., Bui, D. T. & Ngo, P. T. T. Shuffled Frog Leaping Algorithm and Wind-Driven Optimization Technique Modified with Multilayer Perceptron. (2020).
75. Bayraktar, Z. Adaptive Wind Driven Optimization Trained Artificial Neural Networks. *arXiv preprint* arXiv:1911.08942 *(2019)*.

76. Kashan, A. H. A new metaheuristic for optimization: optics inspired optimization (OIO). *Comput. Op. Res.* **55**, 99–125 (2015).
77. ÖZDEMİR, M. & Öztürk, D. Comparative performance analysis of optimal PID parameters tuning based on the optics inspired optimization methods for automatic generation control. *Energies* **10**, 2134 (2017).
78. James, J. & Li, V. O. A social spider algorithm for global optimization. *Appl. Soft Comput.* **30**, 614–627 (2015).
79. James, J. & Li, V. O. A social spider algorithm for solving the non-convex economic load dispatch problem. *Neurocomputing* **171**, 955–965 (2016).
80. Sankar, V. U., Basha, C. H., Mathew, D., Rani, C. & Busawon, K. in *Soft Computing for Problem Solving* 925–940 (Springer, 2020).
81. Jalili, S. & Husseinzadeh Kashan, A. Optimum discrete design of steel tower structures using optics inspired optimization method. *Struct. Des. Tall Spec. Build.* **27**, e1466 (2018).
82. Özdemir, M. T. & Öztürk, D. (ICNES, 2016).
83. El-Bages, M. & Elsayed, W. Social spider algorithm for solving the transmission expansion planning problem. *Electric Power Syst. Res.* **143**, 235–243 (2017).
84. Ewees, A. A., El Aziz, M. A. & Elhoseny, M. in *2017 8th international conference on computing, communication and networking technologies (ICCCNT).* 1–6 (IEEE).
85. Abdi, H. & Williams, L. J. Principal component analysis. *Wiley Interdisc. Rev.: Comput. Stat.* **2**, 433–459 (2010).
86. Mehrabi, M., Scaioni, M. & Previtali, M. Forecasting Air Quality in Kiev During 2022 Military Conflict Using Sentinel 5P and Optimized Machine Learning. *IEEE Trans. Geosci. Remote Sens.* (2023).

Acknowledgements

Author contributions

All authors wrote the main manuscript text. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.