



Solar Irradiance Prediction Using an Optimized Data Driven Machine Learning Models

Mantosh Kumar · Kumari Namrata ·
Nishant Kumar · Gaurav Saini

Received: 10 October 2022 / Accepted: 21 April 2023 / Published online: 24 May 2023
© Springer Nature B.V. 2023

Abstract For a higher degree of penetration of renewable energy into the controls of the existing power system, an accurate solar energy prediction is necessary. Data-driven algorithms may be used to enhance solar generation forecasts as data has now become readily accessible in large quantities. To address these predicting issues in this research article three machine learning models: Support Vector Regressor (SVR), Multilayer Perceptron (MLP) and Random Forest Regressor (RFR) have been incorporated to forecast the Global Horizontal Irradiance (GHI), Diffused Horizontal Irradiance (DHI), Diffused Normal Irradiance (DNI) based on the spatiotemporal factors. In order to improve the prediction accuracy, the parametric tuning of models has been carried out with the two met heuristic algorithms: Moth Flame

Optimization (MFO) and Grey Wolf Optimization (GWO) and also validated with the novel application of Evolve Class Topper Optimization (ECTO) method. Corresponding performance measures, including Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Max Error (ME), and Coefficient of Determination (R²), are employed to evaluate each model's performance. The results obtained through a comparative assessment of all machine learning models confirmed that the ECTO based models have outperformed others and the RFR-ECTO model is the best forecasting model having the highest R² scores of 0.9441, 0.9107 and 0.8882 and the lowest RMSE value of 75.8613 W/m², 40.8714 W/m², 94.8916 W/m² for GHI, DHI and DNI respectively which ensures that the designed predictive model can be implemented for prediction of solar energy.

M. Kumar · K. Namrata (✉)
Department of Electrical Engineering, National Institute of Technology, Jamsshedpur, Jharkhand, India
e-mail: namrata.ee@nitjsr.ac.in

M. Kumar
e-mail: mantosh.nith@gmail.com

N. Kumar
Electrical Engineering, B K Birla Institute of Engineering & Technology, Pilani, India
e-mail: krnishant125@gmail.com

G. Saini
Department of Mechanical Engineering, Harcourt Butler Technical University, Kanpur, Uttar Pradesh, India
e-mail: gaurav161990@gmail.com

Keywords SVR · MLP · RFR · MFO · GWO · ECTO · Solar Irradiance · Forecasting

1 Introduction

1.1 Motivation

Recent years have seen a tremendous expansion of renewable energy, fueled by favorable legislative developments and significant cost decreases for wind and solar photovoltaic in particular [1]. With the recent high rise of wind and solar photovoltaic,

adding to the already sizable contribution of hydropower, the energy sector continues to be the most promising for renewables [2]. However, just 5% of the world's energy consumption is from electricity, and the use of renewables in the transportation and heating industries is still essential to the energy transition. In the short and long run, renewable energy offers a huge potential to lower costs and dependency on fossil fuels. Natural gas, oil, and coal prices have surged considerably faster than those for new solar PV and wind installations, which have reversed a decade-long trend of cost decrease. As a result, renewable power is now more competitive than it was before.

As per a recent IEA analysis due to increasing commodity and freight prices, solar PV and wind costs are anticipated to remain higher in 2022 and 2023 than at pre-pandemic levels [3]. However, because natural gas and coal costs have increased considerably more suddenly, their competitiveness actually rises. In 2022, the capacity for renewable energy is predicted to grow by more than 8%, reaching over 320 GW. To make up for decreased hydropower and constant year-over-year wind additions, solar PV expansion is unable to fully offset growth in 2023 unless new rules are quickly put into place. The total generation capacity up to 2021 for the entire world has been shown in Fig. 1 as per IRENA's recent report, which shows the continuously incremented contribution of solar energy.

Despite the freely available nature of solar energy, the reliability of utilizing this renewable energy source is still challenging due to the dependency on meteorological parameters like latitude and longitude, temperature, wind speed and direction, seasonal variation etc. As a result the power generation totally proportional to

the amount of solar radiation received on a particular day and location with weather patterns associated [4]. In order to design and manage this variation of solar energy output on a daily basis and meet consumer demand and supply independent of weather conditions, researchers are continuously looking for models that handle these uncertainties[5]. Hence, it is extremely important to forecast short-term solar energy.

The recent introduction of advanced forecasting techniques serves as an example of how the two perspectives converge to merge physical and statistical methodologies as necessary to achieve better accuracy. Engineering and science fields have long realized the role of machine learning (ML) models, which are created using statistical techniques. It employs a process of gathering data, developing algorithms, creating machine learning models, making predictions, and then fine-tuning those models [6].

In the last few years of research, it has been found that the nominal machine learning models are facing some discrepancies in the determination of the right kernel functions in SVM, hidden layers and neurons in MLP and tree depth, leaf nodes and internal node splitting in RF regressor which are affecting the computational cost and accuracy of the model. ML models with fixed parametric features leave the scope for the improvement of prediction accuracy which can be only possible by selecting the best feature of an ML model. The tuning of parameters of the machine learning models with met heuristic methods has confirmed the improvement of the model prediction accuracy as per recent analysis [7]. Hence in order to improve the performance of the nominal ML models initially two metaheuristic algorithms have been applied for

Fig. 1 Overall Generation Capacity of Renewable Energy

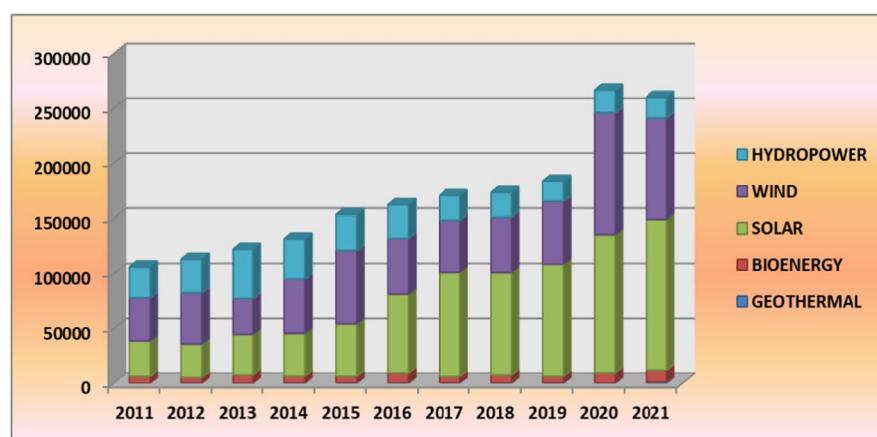


Table 1 Literature survey of latest solar forecasted methods adapted

Authors	Objective	Solution
Imane Jebli et.al. [12]	To forecast the short-term and real-time solar power, ML models with the Pearson coefficient were utilised	SVR, LR, MLP and RF
Soumia Benlebna, Nallapaneni Manoj Kumar and Ali Tahri [13]	Here, researchers at the University of Saida in Algeria have evaluated the real-time performance characteristics of PV modules under challenging outdoor situations and concluded that the real-time surrounding environment affects the module characteristics	PV Module
Muhammad Shahzad Nazi et.al. [14]	A hybrid model RF-GSA has been proposed for the multi-objective dispatch problem CCHP which improved the economy of the system by 7.13%	RF and GSA
JianxinTu [15]	Two supervised ML models i.e. RF and SVM have been applied for the prediction of acute kidney injury and compared with the Logistic Regression model (LGR)	RF,SVM and LGR
Aafaab Moosa et.al [16]	Machine learning models used to predict and forecast solar irradiance and performance evaluation parameters have been calculated	ANN, XGB and RF
Shahid Husain and Uzair Ali Khan[17]	ML algorithms were applied to predict the DSR on monthly basis for humid-subtropical atmospheres and were concluded that the ML models perform well in comparison with statistical methods	LR, SVR, KNN, RF, GPR and MLP
Nallapaneni Manoj Kumar and M.S.P. Subathra [18]	Random forest machine learning model has been used for the forecasting of 3 years ahead solar irradiance and quantify the degradation rate of the silicon-based PV system	RF
Junho Lee et.al. [19]	Rather of using individual ML models, an ensemble method is employed to forecast solar radiation over a short period of time	RF, Ensemble Method Bagged-Trees, SVM, Boosted-Trees and GPR
YousefAzimi et.al [20]	A prediction for the concentration of effluents in wastewater treatment plant under different scenarios was analysed with the combination of MLP and GA optimization methods	MLP and GA
Mohamed Massaoudi et.al. [21]	The grid load is forecasted using a stacking strategy that combines the three machine learning models	MLP-LGBM-XGB
Pratheeep Kumar P et.al. [22]	The parameters of the RF model have been tuned with Bayesian optimization for classification of breast cancer	RF and Bayesian Optimization
Junliang Fan et.al. [23]	The daily maize transpiration was predicted using ML models, and it's been found that the DNN model is more efficient for estimating daily maize T	ANN, XGB, SVM and DNN
Javad Seyedmohammadi et.al [24]	The forecasting of pistachio production based on the collected soil parameters was performed using the SVR, KNN and SVR-Fire Fly hybrid model, where the hybrid model has shown better results	SVR, KNN and SVR-Fire Fly optimization
Hoang D. Nguyen et.al. [25]	Using XGB and comparing it to other machine learning and empirical models allowed researchers to predict the punching shear resistance of R/C interior slabs	RF, ANN and XGB

Table 1 (continued)

Authors	Objective	Solution
Shuang Li [26]	Based on dynamic inertia weights, nonlinear control parameters and the grey wolf hierarchy, an IGWO with SVR model has been used for the prediction of spontaneous combustion temperature of coal on model is presented. The effectiveness of the improved grey wolf optimizer algorithm is confirmed by numerical experiments	SVR and IGWO
Min Yan Chia et.al. [27]	Two nature-inspired optimization techniques i.e. PSO and WOA have been applied to improve the accuracy of XGB for estimation of evapotranspiration	XGB, WOA and PSO
Ahmet Cevahir Cinar and Narayanan Natarajan [28]	In this research article the forecasting of tomato yield with a robust BMA model using the two-hybrid MLP and ANFIS machine learning model has been performed. As per the results obtained, BMA with various ANFIS and MLP models was efficient in forecasting tomato yield	BMA, ANFIS, MLP, MVO, PSO and FFA
Rui Liu et.al. [29]	Two machine learning models were optimized with GWO for estimation and prediction of the potential of groundwater	RF, SVM and GWO
Inas Bouzateur et.al. [30]	Energy band gap prediction in chalcopyrites has been carried out using the hybridisation of ANN and PSO. The recommended technique can greatly improve convergence and resolution, according to several computational tests	ANN and PSO
Siwei Li [31]	SVM parameters have been optimized using Manta ray foraging optimization to predict the load for the short-term and comparison have been done with other metaheuristic algorithms. The proposed model has shown an accuracy of 99.9% and 99.3% for training and testing dataset confirming the usability of the model	SVM, Manta ray foraging optimization, Fruit-fly optimization algorithm Slime Mould algorithm, Moth Flame optimization Tug of War optimization and Satin Bowerbird optimizer
Md Faisail Kabir et.al [32]	Three dimensionality reduction techniques like Autoencoder, PCA and PCA with kernel have been analysed using SVM and NN for the RNA sequencing data	PCA, Autoencoder, SVM and neural network
Jian Zhou et.aL[33]	RF with SSA has been applied for the prediction mean cutting force and validation has been done using four statistical error metrics	RF and SSA
SeyedHashem Samadi et.al. [34]	In this study, a fixed bed gasification prediction model was created using an MLP learning algorithm. Based on the physicochemical constitution of the biomass and the reactor operating circumstances, a number of MLP models were created to forecast the constitution of the generated gas and the reduced heating value	MLP

the optimization of the parameters associated with the particular model. In this research article the hyper-parameters tuning of three ML models (SVR, MLP and RFR) have been done using two met heuristics techniques (MFO and GWO) and the best hybrid

model has been chosen based on the various performance parameters for the improvement of the prediction accuracy. Further, to check the scope for the improvement in the accuracy of the hybrid model a novel application of the ECTO method has been

Table 2 Selected features for analysis of Solar Radiation

S.No	Input Parameters	Output Parameters
1	Solar Zenith Angle (SZA)	Global Horizontal Irradiance (GHI)
2	Perceptible Water (PW)	Diffused Horizontal Irradiance (DHI)
3	Temperature (T)	Direct Normal Irradiance (DNI)
4	Wind Speed (WS)	
5	Wind Direction (WD)	
6	Relative Humidity (RH)	
7	Pressure (P)	
8	Dew Point (DP)	
9	Surface Albedo (SA)	

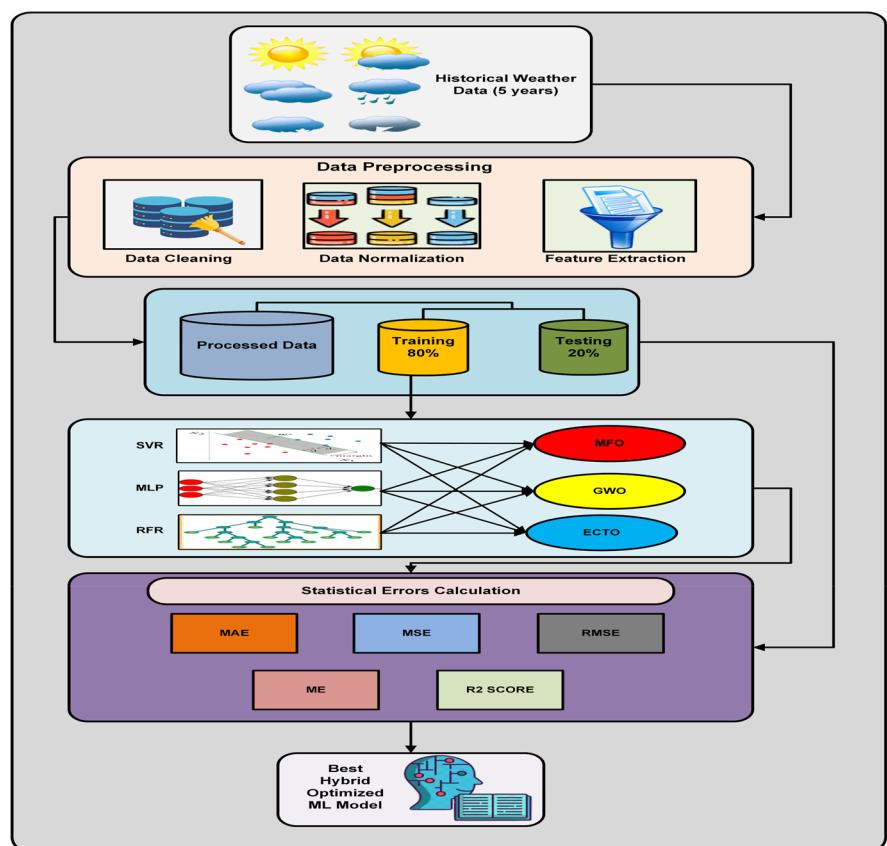
incorporated which has a good convergence rate and is efficient as compared to other optimization methods.

1.2 Status quo of Solar forecasting using AI

In recent years, the ML model is used in data-driven analysis to learn from historical records and establish the correlation between weather forecasts and PV power [8]. Artificial neural networks [9], support

vector regression [10], and other supervised learning models are only a few of the ML models that are employed for this purpose. It is challenging to determine which approaches are the most accurate, though, as the errors reported in various publications are hardly comparable because they depend on the variability and predictability of the solar resource, the accuracy of weather forecasts, as well as the design parameters of the PV systems [11]. Some recent

Fig. 2 Overall Work Flow Architecture



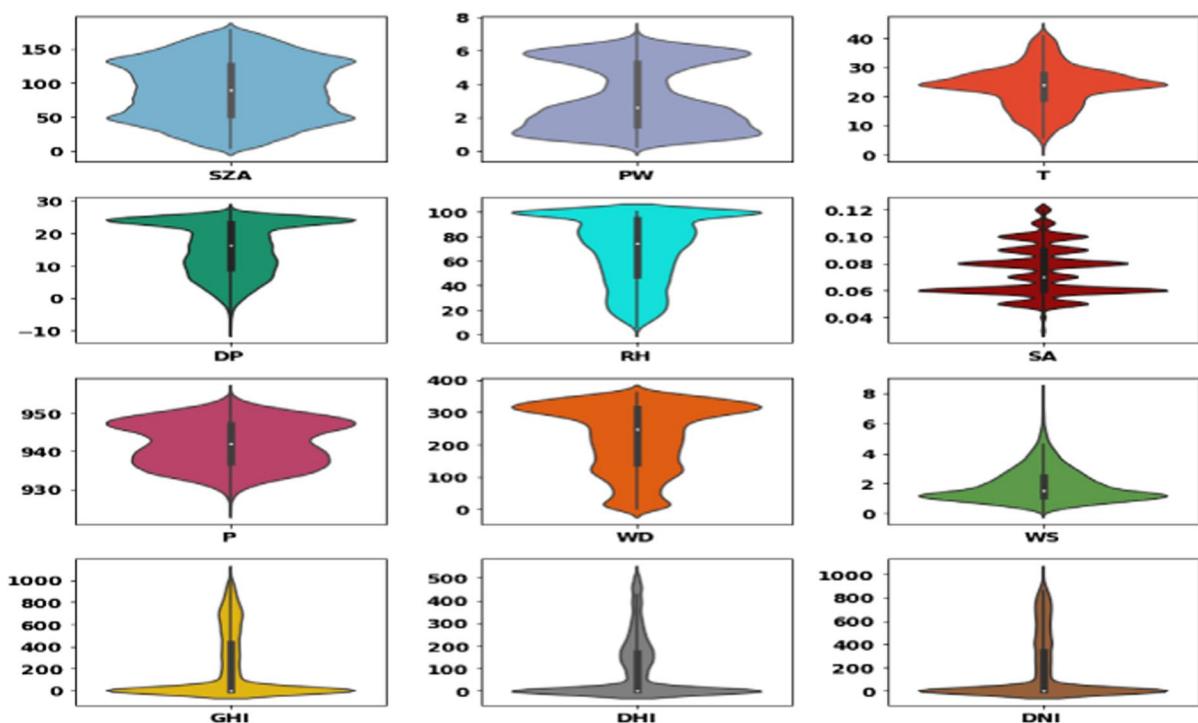


Fig. 3 Data Representation using Violin plot

works related to machine learning applications for prediction or forecasting have been put in Table 1.

As we can observe from the recent pieces of literature various met heuristic methods have been incorporated for the optimization of different objectives which motivated the implementation of these

techniques for the predictive analysis of the renewable energy system.

1.3 Contributions to the paper

The primary research goals of this study are:

Fig. 4 Fisher Test Score for the Chosen Dataset

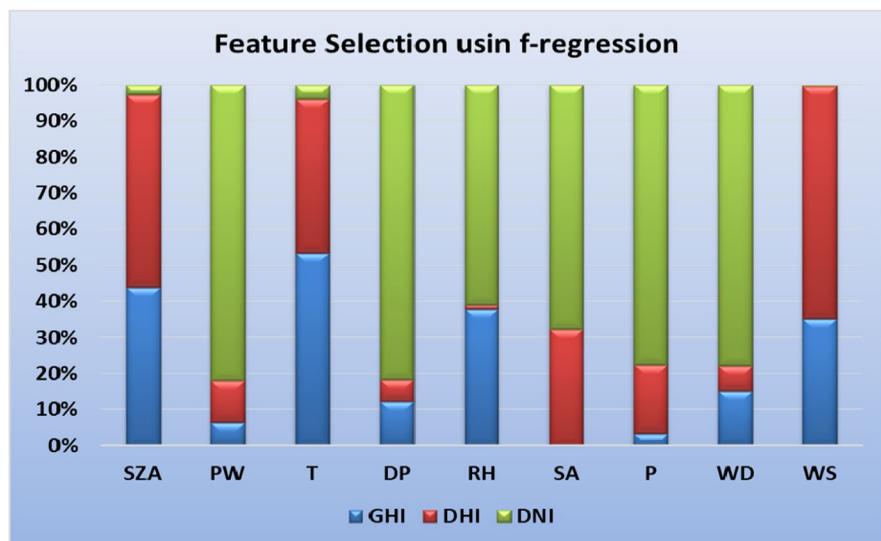
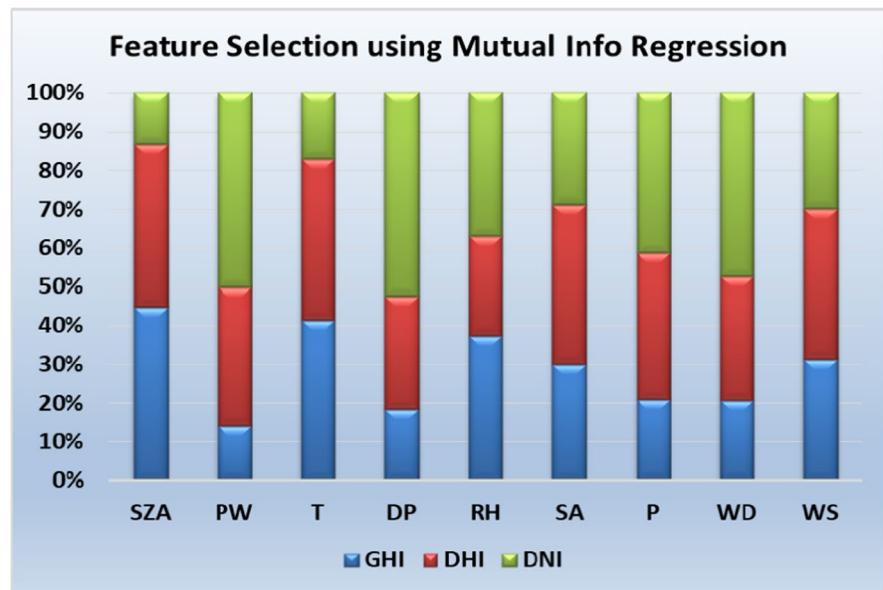


Fig. 5 Mutual Information Test Score for the Chosen Dataset



- (1) To design and evaluate the supervised machine learning models (SVR, MLP and RFR) for forecasting solar irradiance (GHI, DHI and DNI) utilizing large metrological data.
- (2) To implement MFO and GWO algorithms for hyperparameters tuning of all ML models.
- (3) To compare the accuracies of all the machine learning models and finds the best model among them for solar forecasting.

The remainder of the research article has been organized as follows: The methodology adopted for

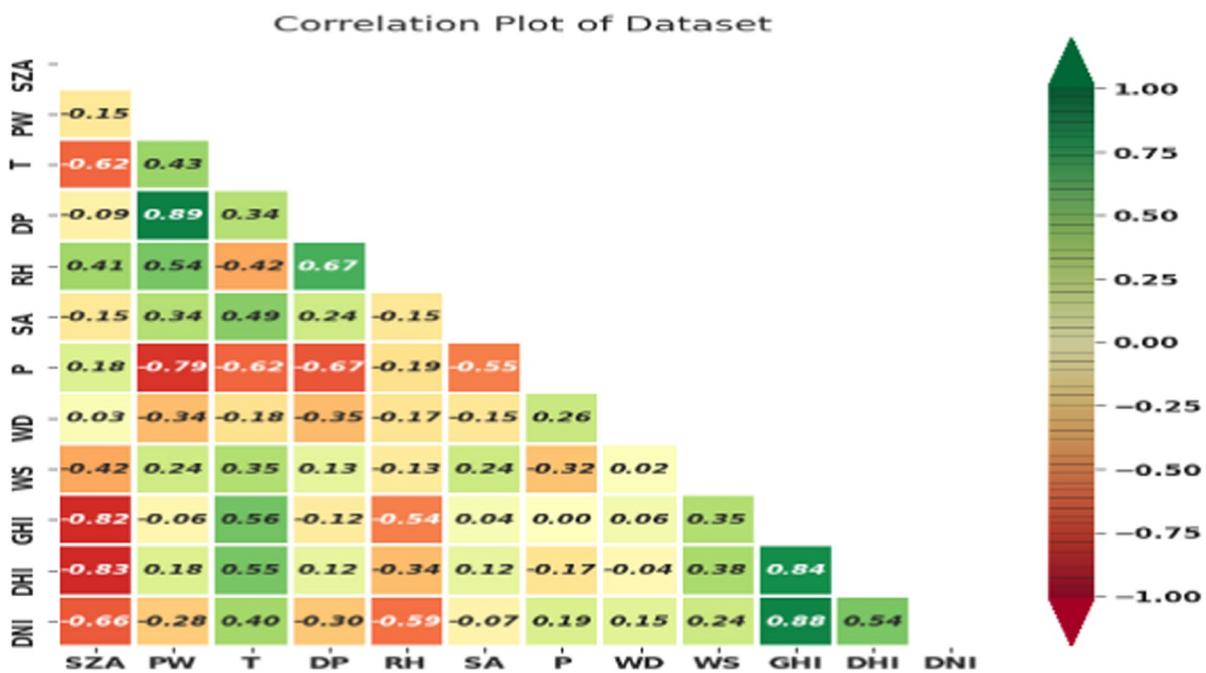


Fig. 6 Correlations of Dataset Variables

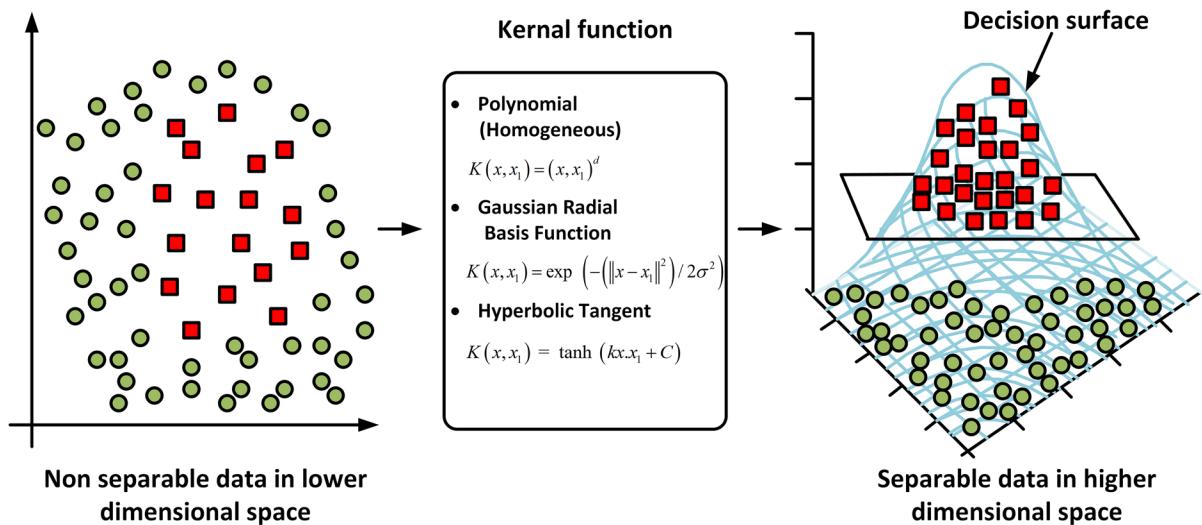


Fig. 7 SVR model

the analysis and the description of the dataset used for the research are discussed in Sect. 2. Section 3 has briefly explained the various algorithms applied and the performance parameters selected for the comparative analysis. The detailed experimental results and overall comparative analysis of all the machine learning models are presented in Sect. 4. Finally, the conclusion of the article has been quoted in Sect. 5.

2 Methodology and Data Description

2.1 Data Representation

The raw data has been taken from the National Solar Radiation Data Base (NSRDB)[35] and has been represented in Table 2 where only input and output variables have been tabulated.

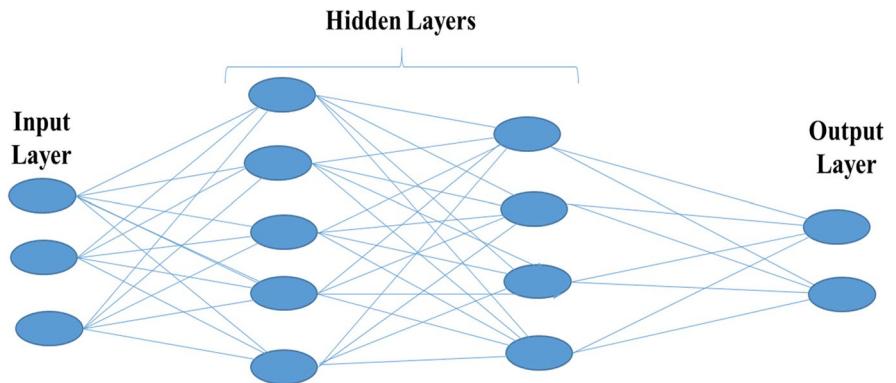
Algorithm 1 Implementation of SVR

Input: Dataset D with X (*Features*) and y (*Target*) with labelled training data, parameters (Kernel function, epsilon, C , tolerance etc.)

Output: Performance evaluation metrics

1. Initialization of regularization parameter, C
2. Put $\alpha_i = 0$ (untrained SVM) at initial phase
3. **while** (stopping criterion) **do**
4. **for** all C_i , $i = 0, i < D$, $i++$
5. **for** all $D = (x_i, y_i)$ **do**
6. Optimization of α_i with kernel function (rbf)
7. Store rate of success
8. **end for**
9. Updation of the best C if needed
10. **end for**
11. **end while**
12. Model fitting with training data
13. Model Validation with testing data
15. **return** performance evaluation parameters

Fig. 8 Architecture of single layer MLP



2.2 Methodology

The methodology applied in this research article has been briefly described through the following step-wise process and Fig 2 depicts the workflow for our experiment:

- 1) Collection of Dataset:—The dataset has been taken from the NREL at the interval of 30 min. The dataset contains 17 parameters out of which 9 have been taken as the input entities while 3 parameters have been chosen as the target entities which are already been cited in Table 2. The dataset has been represented in Fig 3 with the violin plot.

A violin plot generally uses kernel density curves to represent the distributions of numerical data for one or more groups. Each curve's thickness reflects the general frequency of data points in each location. Many of the summary data in violin plots and box plots are similar:

- The median is indicated by the white mark.
- The interquartile range is indicated by the broad grey bar in the middle.
- Except for points identified as "outliers" using a technique based on the interquartile range, the thin grey line reflects the remainder of the distribution.
- Preprocessing of Dataset: The meteorological data collected generally exists in its raw state and must

Algorithm 2 Implementation of MLP

Input: Dataset D with X (Features) and y (Target) with labelled training data, parameters
(number of hidden layers, alpha value, learning rate, maximum iteration etc.)

Output: Performance evaluation metrics

1. Initialization of weights (random value)
 2. **while** (stopping criterion) **do**
 3. **for** all network (N) and output nodes (j)
 4. Evaluate the activation (j)
 5. **for** all input nodes (i) to output nodes (j)
 6. $w_{ij} = w_{ij} - \eta \frac{\partial E}{\partial w_{ij}}$; η is learning rate
 7. **end for**
 8. **end for**
 9. update w_{ij}
 10. **end while**
 11. Model fitting with training data
 12. Model Validation with testing data
 13. **return** performance evaluation parameters
-

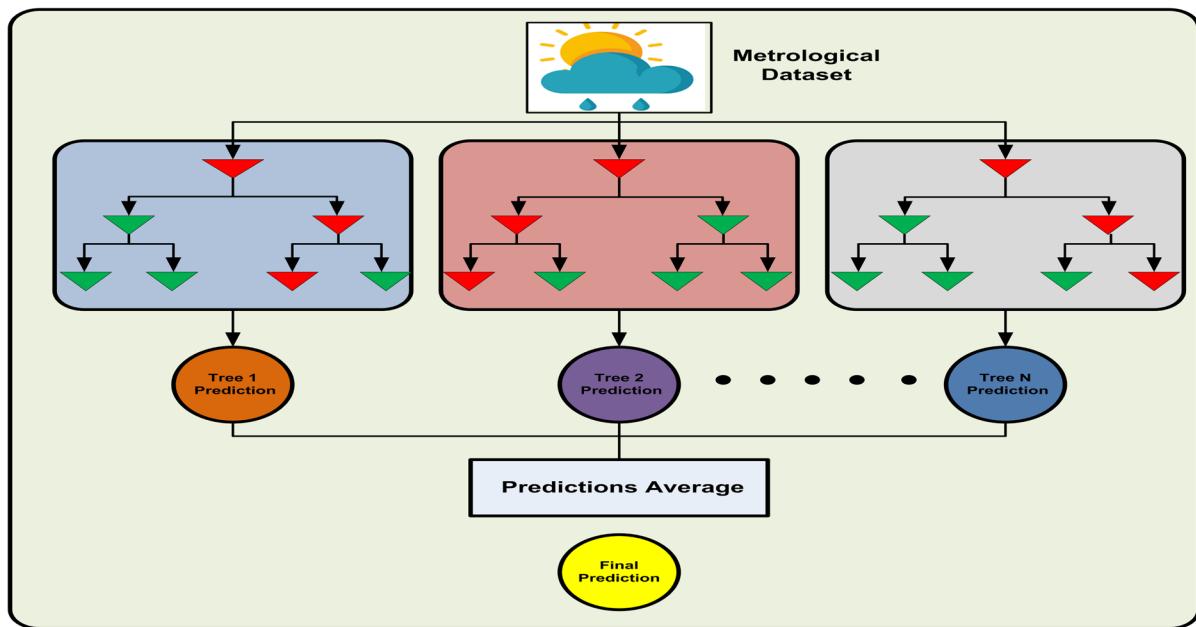


Fig. 9 RFR model

be pre-processed before it can be used for training purposes. It has been analyzed by various researchers that the preprocessed data improves the learning capacity of the ML models which results in the enhancement of the learning model accuracy. The

process adopted for the data preprocessing can be understood with the following steps:-

- Data Cleaning: Generally, the data are in raw forms which need to be cleaned by removing the unwanted and duplicate entities in the chosen dataset.

Algorithm 3 Implementation of RandomForest Regressor

Input: Dataset D , X (*Predictors*) and y (*Target*) loaded with training labelled data, parameters (estimators, maximum depth, minimum sample leaf etc.)

Output: Performance evaluation metrics

Bootstrapping Process

1. **for** $q = 0, q++$ (number of trees)
2. **for** $p = 0, p++$ (number of nodes)
3. Random selection of features ' f ' from D (with replacement)
4. **for** features = 1 to f
5. Evaluate the information gain
6. **end for**
7. Choose the best information gain value
8. **end for**
9. **end for**

Aggregation Process

10. Evaluate each decision tree for inputs X
11. Averaging the outcomes of all decision trees as per equation 2
14. Model fitting with training data
15. Model Validation with testing data
16. **return** performance evaluation parameters

- Data Normalization: Here, all data variables are normalized to a common interval, which is often between 0 and 1 or -1 and 1. This phase compares the values of numerous variables and speeds up the learning process of the ML models. The expression for data normalization has been represented in Eq. (1)

$$Data_{normalized} = \frac{Data_{actual} - Data_{minimum}}{Data_{maximum} - Data_{minimum}} \quad (1)$$

- iii) Feature Extraction: The Exploratory Data Analysis (EDA) process helps to identify the important features from the dataset which improves the learning algorithms efficiency and also reduces the size of the dataset. There are various methods available in the present scenario for feature extraction but here in this research article three methods have been incorporated for the analysis:
- iv) Fisher Test Score: Fisher score is an effective method for reducing data feature dimensions. Finding a feature subset that maximises the dis-

tances between data points in various classes while decreasing the distances between data points in the same class is the major objective of this approach [36]. The analysis has been presented in Fig 4.2. Mutual Information Score: Battiti (1994) developed the aggressive feature selection utilising the Mutual information approach, which can evaluate both linear and non-linear relationships between variables. This technique, which chooses one feature at a time from the initial collection of data, was developed to extract the most crucial subset of features [37]. The feature extraction scoring using this test has been shown in Fig 5.

- 3. Correlation Analysis: This analysis is the best way to select features from a large data set. The correlation plot has been shown in Fig 6 where the plot represents the relationship between the features having a range of 0 to 1. The features nearest 1 show a better correlation while others near zero represent the worst correlation.

All of the above mentioned feature extraction process shows that the features have different

Algorithm 4 Pseudo code of moth-flameoptimization

Input: Population (n), itr_{max} = max iteration
Output: Optimal solution (R_2)

1. **Fitness** Calculation of search agent
- F_{best} best optimal solution (search agent)
2. **while** ($itr < itr_{max}$) **do**
3. Update flame number according to equation 7
4. Check M_i boundary conditions
5. **if** k==1
4. F= sort (M_k)
5. OF= sort [(OM)_k]
6. else
7. F= sort (M_{k-1}, M_k)
8. OF= sort [(OM)_{k-1}, (OM)_k]
9. **end**
10. **Fitness:** best solution (search agent) as F_{best}
11. For i=1,2,...n
12. For j=1,2,...m
13. Update M_{ij} using eq 2
14. **Endfor**
15. **Endfor**
16. k=1+1; and $itr=itr+1$
17. **end while**
18. **Return** F_{best} (R_2)

dependencies with respect to the output variables which help to utilize only the features which are highly dependent on the target to improve the accuracy of the models. However, since the total features are fewer in this dataset so the impact of all the variables has been considered for the evaluation of the model.

- iv) Data Splitting: The preprocessed dataset has been segregated into testing and training sets where 80 percent for training is fed to ML models for learning and 20 percent for testing purposes which helps to obtain the performance parameters.

3 Forecasting Algorithm Used

3.1 Support Vector Regressor

The support vector machine was introduced by Cortes, C. and Vapnik, V in 1995, a kernel-based machine learning algorithm for classification problems which is now known as the support vector machine and also has applications in regression problems [38]. SVR approaches split data sets by projecting them from lower-dimensional to higher-dimensional space and making them linearly distinguishable; a hyper-plane is then drawn as the classification border to split data

points [39]. The working and implementation of SVR are described in Algorithm 1 and the schematic diagram for the SVM is shown in Fig. 7.

3.2 Multilayer Perceptron

The basic design of the Multilayer perceptron is based on deep learning, a classification of the feed-forward artificial neural network [40]. A typical MLP is a fully linked network with an input level, one or more hidden levels, and an output level. Each point in one layer is connected to every point in the following layer through a weighted connection. MLP modifies the weight values automatically while constructing the model using the feed-forward [41]. The working and implementation of MLP are shown in Algorithm 2 and the schematic diagram for the feed-forward MLP neural network is shown in Fig. 8.

3.3 Random Forest Regressor

Random forest regression falls under the category of supervised machine learning and was introduced first by L.Breiman which depends on the ensemble of various decision trees for predicting the target from an input features set [42]. RFR generally use the bagging method which is also known as bootstrapping where it creates different decision tree from the chosen training data with resampling and finally the predictive

Algorithm 5 Pseudo code of grey wolfoptimization

Input: Wolf Population (N), \vec{A} , \vec{a} and \vec{C}
Output: Optimal solution (R_2)

1. **Fitness** Calculation of search agent (i.e Grey wolves)
- \vec{X}_{GW}^a best optimal solution (search agent)
- \vec{X}_{GW}^b second best optimal solution (search agent)
- \vec{X}_{GW}^δ third best optimal solution (search agent)
2. **while** (itr < itr_{max}) **do**
3. **for** i=1, 2, 3, ..., N
4. **Update** current position using eq 14
5. **end for**
6. **Update** \vec{A} , \vec{a} and \vec{C}
7. **Fitness** Calculation of search agent
8. **Update** \vec{X}_{GW}^a , \vec{X}_{GW}^b , and \vec{X}_{GW}^δ
9. itr = itr+1
10. **end while**
11. **Return** \vec{X}_{GW}^a

Algorithm 6 Pseudo Code of ECTO

Step 1: Determine the higher range limit and lower limit of upper class deciding factors like solar panel numbers, battery capacity, Grid maximum value

Step 2: Initialize the learners or student's number

Step 3: for $p = 1, 2, \dots, a$ do

for $q = 1, 2, \dots, b$ do

Initialize the α_p^q between (0,1)

Evaluate the equation (17) and (18)

end for

Step 4: end for

Step 5: Initialization through OBLM

for $p = 1, 2, \dots, a$ do

for $q = 1, 2, \dots, b$ do

Evaluate, $s_{a,b} = s_{\min,a} + (s_{\max,n} - s_{a,b})$

Step 6: Compute the fitness value for every particle generated at random.

Step 7: Selecting Best-fitting students from the subgroups as the starting population.

Step 8: Check the G.T. and C.T

Step 9: While $ITR > M$ do

Step 10: Update of G.T.

if $G.T. \geq G.T_{ITR}^{-1}$ then

$G.T_{ITR} = G.T_{ITR}^{-1};$

Else

$G.T_{ITR} = G.T_{ITR};$

Step 11: Update C.T.

if $C.T_{ITR} \geq C.T_{ITR}^{-1}$ then

$C.T_{ITR} = C.T_{ITR}^{-1};$

else

$C.T_{ITR} = C.T_{ITR};$

end

Step 12: Calculate each student understanding factor

if $G.T_{ITR} = s_{a,b}$ then

$U_{a,b,c}^{ITR+1} = W * U_{a,b,c}^{ITR} + e_1 * rand * (C.T_{ITR} - G.T_{a,b,c}^{ITR})$

Else

$U_{a,b,c}^{ITR+1} = W * U_{a,b,c}^{ITR} + e_2 * rand * (G.T_{a,b}^{ITR} - G.T_{a,b,c}^{ITR})$

End

Step 13: Upgrade the student using $s_{a,b,c}^{ITR} = U_{a,b,c}^{ITR} + s_{a,b,c}^{ITR}$

Step 14: Student Mutation by equation 14

Step 15: $ITR = ITR + 1$

end while

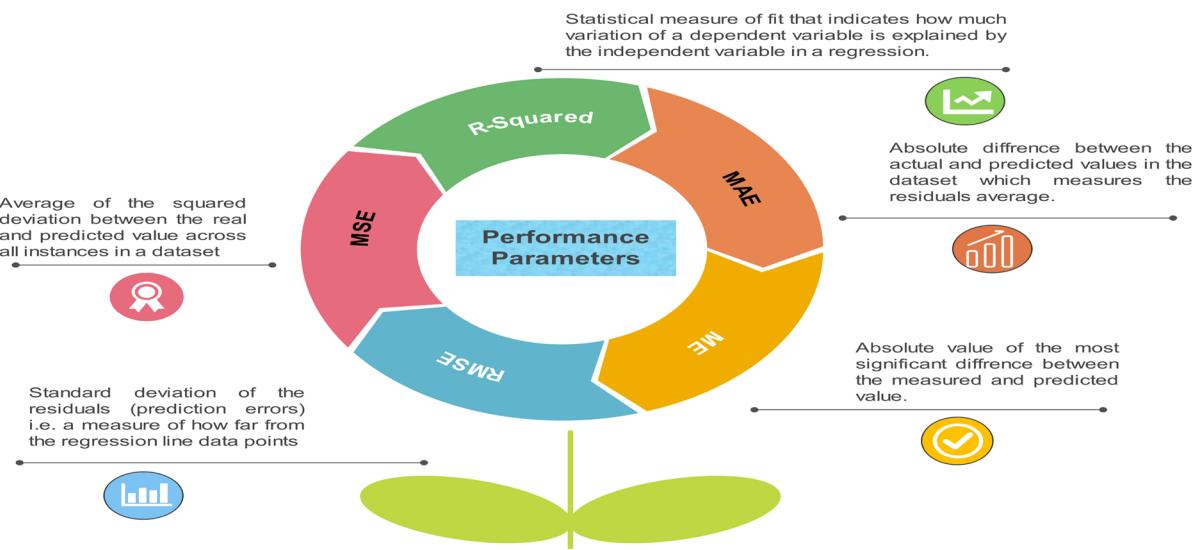


Fig. 10 Performance Evaluation parameter used in our work

outcome of each decision trees is averaged to determine the final prediction. The pseudo-code description for RFR has been described by Algorithm 3 and the schematic diagram for the RFR algorithm is shown in Fig. 9 where the red colour represents the workflow of the decision. The equation describing the final prediction of all the decision trees can be represented as:

$$P_{rf}(X) = \frac{1}{T} \sum_{t=1}^T P_t(x) \quad (2)$$

where T is the total number of the decision tree.

3.4 Moth Flame Optimization

Seyedali Mirjalili proposed the MFO algorithm, which was motivated by the mirroring behaviour of moths. Moths are the insects that resemble butterflies the most [43]. These moths employ a peculiar kind of nocturnal triangulation known as transverse orientation, which allows moths to hover in a straight line by remembering the stationary perspective parallel to the moon. Moths float in spiral patterns in the latency of an unreal source of light that is close to the moon by focusing upon the source of light.

$$AM = \begin{pmatrix} AM_1 \\ AM_2 \\ \vdots \\ AM_n \end{pmatrix} \quad (3)$$

Moths and flames are two significant components of the MFO structure. The moths that hover in a deeply engaged, d-dimensional plane act as search mediators. In the M matrix, the dwelling is reserved. The fitness value relevant to each month is subsequently stored in the array AM. The size of a moth and a flame are the same. Moth and flame both function as parts of the algorithmic solution. Flame, on the other hand, denotes the moth's ideal position, whereas the moth denotes the hunting agent. Moths release flames that serve as flags throughout the search process. As a result, both positions are being updated, decreasing the likelihood that one would

Table 3 Chosen values for SVR

S. No	Description	Values
1	Regularization Parameter (C)	0.01
2	Kernal function	RBF
3	Epsilon	0.1
4	Tolerance	0.001

Table 4 Performance evaluation parameters outcomes for SVR without Optimization

	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	MAX Error (W/m ²)	R ² Score	
					Train	Test
GHI	73.2215	13,138.95	114.6252	796.87	0.8514	0.8503
DHI	73.1946	13,124.34	114.5615	774.62	0.8523	0.8514
DNI	82.3159	18,149.01	134.7182	708.69	0.7506	0.7525

be lost. According to Eq. 4, the moth's location is updated.

$$M_j = SF(M_j, F_j) \quad (4)$$

M_j indicates the j^{th} moths, whereas F_j represents the j^{th} flames and SF is for the spiral function which is expressed in Eq. 5.

$$SF(M_j, F_j) = D_j * e^{bt} * \cos(2\pi t) + F_j \quad (5)$$

where b is the spiral constant, t is the arbitrary value (-1,1) and D_j is the j^{th} moth and j^{th} flame Euclidean distance. D_j is represented as Eq. 6

$$D_j = |F_j - M_j| \quad (6)$$

In the initial stage, flames and moths remain to be the exact number, which may reduce the potential of sophisticated solutions to be diverse due to moths' conscious choice of 'n' distinct locations in quest of room for updating. Equation 7 is used to update the flames.

$$F_{no} = \text{round}\left(F - j * \frac{F - 1}{itr_{\max}}\right) \quad (7)$$

The working of the algorithm in the form of the pseudo-code has been shown below:

3.5 Grey Wolf Optimization

Grey wolf optimization, modelled on the natural hunting tactics of grey wolves is a meta-heuristic optimization technique proposed by [44]. Every wolf in GWO symbolizes a search agent (potential solution). GWO classifies the wolves into four categories alpha (α), beta (β), gamma (δ) and omega (ω) by replicating the grey wolf population's hierarchy. The wolves in the first three grades correspond to the current three best solutions. The current three best solutions are embodied by the wolves in the first three categories (α, β, δ). The (ω) wolves follow the pack's strongest wolves.

3.5.1 Encircling

Grey wolves surround their prey as part of the hunting process. So, the initial phase of the mathematical modelling of the GWO is to surround the target, which may be expressed by the following equations [8, 9].

$$\vec{D}_{GW} = |\vec{C} \cdot \vec{X}_{GW}^p(t) - \vec{X}_{GW}(t)| \quad (8)$$

$$\vec{X}_{GW}(t+1) = |\vec{X}_{GW}^p(t) - \vec{A} \cdot \vec{D}| \quad (9)$$

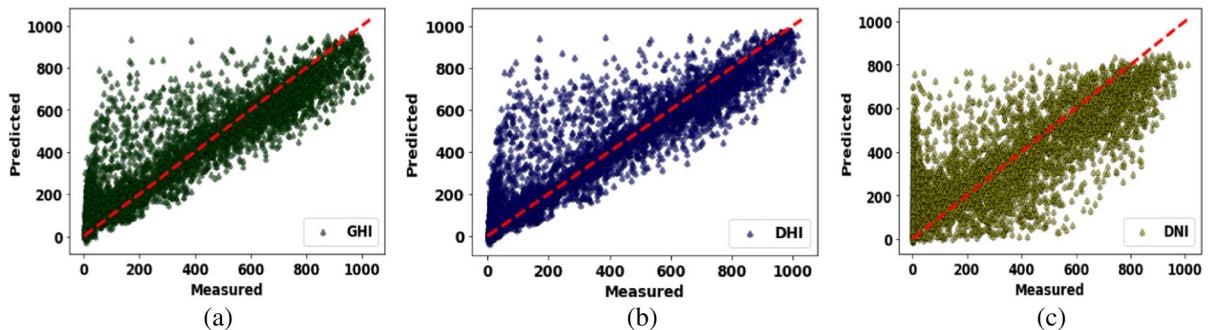


Fig. 11 Scatter plots of SVR model without optimization for (a) GHI (b) DHI (c) DNI

Table 5 Hyper-parameters range for SVR

Sl. No	Description	Lower Bound	Upper Bound
1	Regularization Parameter (C)	10	100
2	Epsilon	0.0001	0.1
3	Tolerance	0.001	0.1

where, \vec{A} and \vec{C} are noted as the coefficient vectors and “t” is symbolised as the current iterations. \vec{X}_{GW} signifies grey wolf position vector and Prey’s position vector is indicated by \vec{X}_{GW}^p , whereas, the \vec{D}_{GW} is the vector which depends on \vec{X}_{GW}^p . Computation for the coefficient vectors \vec{A} and \vec{C} are as follows:

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (10)$$

$$\vec{C} = \vec{2} \cdot \vec{r}_2 \quad (11)$$

$$\vec{a} = 2 - 2 \left(\frac{itr}{itr_{max}} \right) \quad (12)$$

where, \vec{r}_1 and \vec{r}_2 are random variables in the interval $[0, 1]$ and values of \vec{a} are linearly decreasing from 2

to 0 throughout the span of iterations. Concisely, \vec{r}_1 and \vec{r}_2 vectors enable wolves to extend to any location. Accordingly, Eq. (11) and (12) indicates that the grey wolf may update their position inside the search space (space circling prey) at any random point. The same approach could be employed in a search space with dimension "n", where the grey wolves will circle the best outcome thus far in hyper-cubes or hyper-spheres.

3.5.2 Hunting

Grey wolves are capable of identifying the position of prey and surrounds them. The α usually leads the hunt. The β and δ may occasionally engage in hunting as well. Nevertheless, we are unsure of where the optimal position is located in a vague search space (prey). We postulate that the alpha (best solution), beta, and delta have superior information about the probable location of prey in order to mathematically imitate the hunting behaviour of grey wolves. Therefore, we reserve the first three best responses we have so far identified. Thus, to compel the other searching agent, along with omegas, to upgrade their positions as per the top searching agent ‘s status. The

Table 6 Performance evaluation parameters outcomes for SVR-MFO

	Iterations	Population Size	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	R ² Score	
						Train	Test
GHI	100	50	46.7476	7787.72	88.2480	883.48	0.9109 0.9113
		100	50.0502	8174.96	90.4154	878.26	0.9065 0.9069
		150	46.8482	7803.03	88.3348	888.92	0.9107 0.9111
		200	46.0750	7712.37	87.8201	883.85	0.9118 0.9121
		500	45.8923	7694.22	87.7167	883.66	0.9120 0.9123
DHI	100	50	25.4888	2245.03	47.3817	423.6142	0.87334 0.87370
		100	26.0863	2287.53	47.8281	422.1221	0.87098 0.87131
		150	25.4859	2244.81	47.3795	423.6042	0.87335 0.87371
		200	25.7093	2259.19	47.5309	422.9880	0.87252 0.87290
		500	25.6826	2259.02	47.5291	423.2459	0.87255 0.87291
DNI	100	50	64.9512	13,471.83	116.0682	739.0360	0.81254 0.81632
		100	67.0448	13,906.95	117.9277	743.6099	0.80634 0.81039
		150	65.9984	13,676.05	116.9446	741.6379	0.80954 0.81354
		200	65.2329	13,510.32	116.2339	739.2310	0.81196 0.81579
		500	65.2318	13,515.73	116.2571	739.1579	0.81189 0.81572

Table 7 Performance evaluation parameters outcomes for SVR-GWO

	Iterations	Population Size	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	MAX Error (W/m ²)	R ² Score	
							Train	Test
GHI	100	50	46.4172	7754.20	88.0579	884.34	0.91135	0.91169
		100	46.7349	7788.43	88.2521	883.91	0.91094	0.91130
		150	46.4014	7752.90	88.0506	884.49	0.91136	0.91171
		200	45.9093	7696.13	87.7276	883.86	0.91203	0.91235
		500	45.8393	7689.83	87.6917	883.73	0.91211	0.91242
DHI	100	50	25.3464	2235.65	47.2826	423.59	0.87388	0.87422
		100	25.7797	2265.66	47.5989	423.42	0.87216	0.87254
		150	25.3434	2235.49	47.2810	423.77	0.87387	0.87423
		200	25.3600	2237.30	47.3001	423.78	0.87377	0.87413
		500	25.2275	2229.93	47.2221	424.04	0.87419	0.87455
DNI	100	50	64.8099	13,442.20	115.9405	738.16	0.81296	0.81672
		100	65.8535	13,441.10	115.9357	738.60	0.81297	0.81674
		150	64.7669	13,425.59	115.8688	738.12	0.81319	0.81695
		200	65.1328	13,505.26	116.2121	738.32	0.81205	0.81586
		500	65.1824	13,498.43	116.1827	739.15	0.81214	0.81596

Table 8 Performance evaluation parameters outcomes for SVR-ECTO

	Iterations	Population Size	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	MAX Error (W/m ²)	R ² Score	
							Train	Test
GHI	100	50	45.4612	7648.29	87.4540	883.17	0.92127	0.92113
		100	45.6479	7675.34	87.6090	882.88	0.91785	0.91372
		150	44.9843	7637.87	87.3949	883.36	0.92291	0.92243
		200	44.7567	7591.21	87.1275	882.72	0.92377	0.92351
		500	44.4318	7588.92	87.1144	882.65	0.92463	0.92394
DHI	100	50	24.7746	2123.73	46.0839	422.31	0.88365	0.88248
		100	24.9023	2159.71	46.4726	422.24	0.88228	0.88173
		150	24.7584	2142.57	46.2879	421.92	0.88363	0.88261
		200	23.2118	2147.44	46.3404	422.48	0.88342	0.88229
		500	23.1511	2106.95	45.9015	421.87	0.88471	0.88293
DNI	100	50	63.9741	12,879.17	113.4866	737.94	0.82122	0.82891
		100	64.4956	12,876.24	113.4735	738.03	0.82124	0.82928
		150	63.7241	12,814.71	113.2020	737.69	0.83097	0.83034
		200	64.2113	12,915.15	113.6448	737.97	0.82113	0.82819
		500	64.2094	12,901.19	113.5834	738.11	0.82191	0.81997

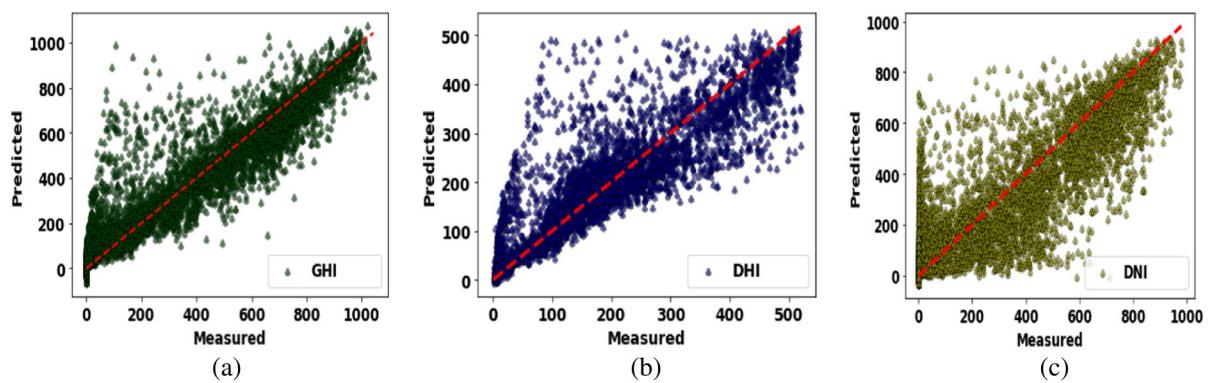


Fig. 12 Scatter plots of SVR model with MFO optimization for (a) GHI (b) DHI(c) DNI

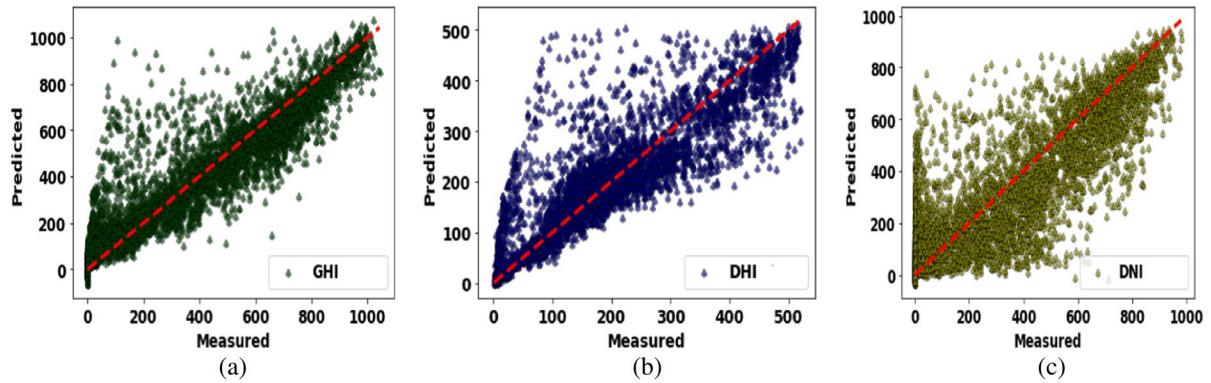


Fig. 13 Scatter plots of SVR model with GWO optimization for (a) GHI (b) DHI(c) DNI

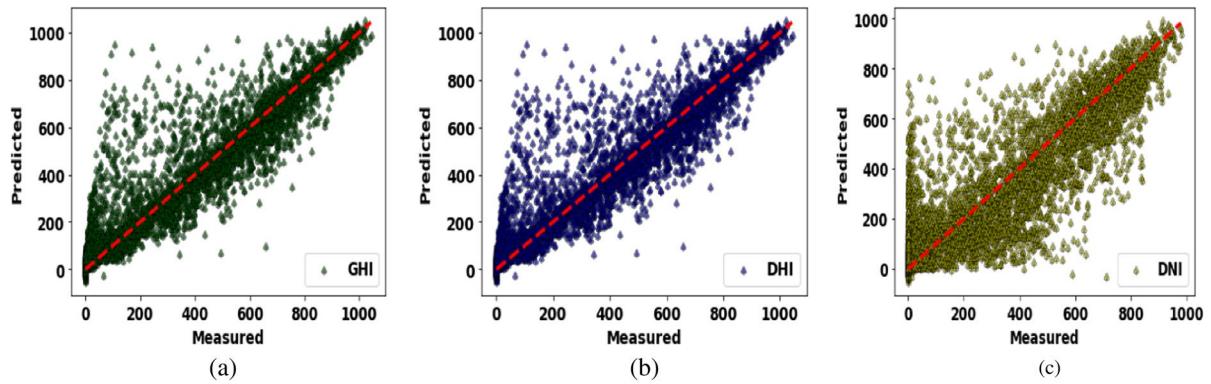


Fig. 14 Scatter plots of SVR model with ECTO optimization for (a) GHI (b) DHI(c) DNI

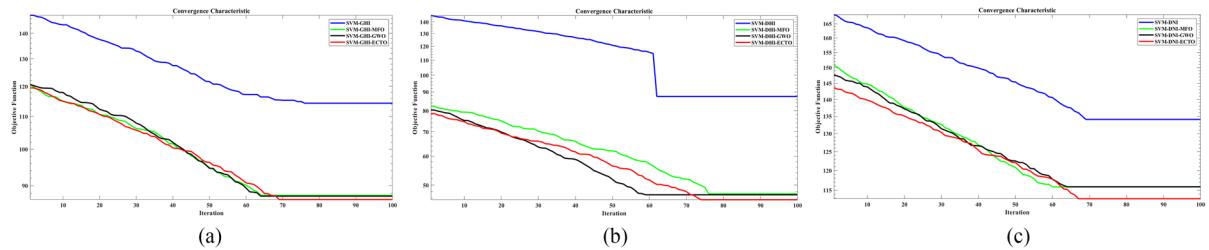


Fig. 15 Conversion curves for all SVR models for (a) GHI (b) DHI (c) DNI

below-mentioned Eqs. (13–16) are followed for above stated context:

$$\vec{X}_{GW}(t+1) = \frac{\vec{X}_{GW}^1 + \vec{X}_{GW}^2 + \vec{X}_{GW}^3}{3} \quad (13)$$

$$\vec{X}_{GW}^1 = \vec{X}_{GW}^\alpha - \vec{A}_1 \cdot \left(\left| \vec{C}_1 \vec{X}_{GW}^\alpha - \vec{X}_{GW} \right| \right) \quad (14)$$

$$\vec{X}_{GW}^2 = \vec{X}_{GW}^\beta - \vec{A}_2 \cdot \left(\left| \vec{C}_2 \vec{X}_{GW}^\beta - \vec{X}_{GW} \right| \right) \quad (15)$$

$$\vec{X}_{GW}^3 = \vec{X}_{GW}^\delta - \vec{A}_3 \cdot \left(\left| \vec{C}_3 \vec{X}_{GW}^\delta - \vec{X}_{GW} \right| \right) \quad (16)$$

3.5.3 Searching and Attacking Prey

Grey wolves primarily use the (α), (β), and (δ) positions to guide their search. They disperse from one another to explore for prey and then reassemble for the attack. In order to mathematically replicate divergence, we employ \vec{A} random values greater than 1 or lower than -1 to induce the local search to disperse from the prey. This facilitates exploration and makes it possible to search widely for the GWO algorithm.

As already mentioned, after the prey stops moving, the grey wolves attack it to end the hunt. We lower the value \vec{a} to mathematically simulate approaching the prey. Keeping in mind the reduction occurring in \vec{A} may also be decreased by \vec{a} . In other respects, \vec{a} decreases from 2 to 0 throughout the duration of iterations, and \vec{A} is a random number in the range [2a, 2a]. A search agent's future position may be anywhere between its present position and the prey's

position when random numbers \vec{A} are in the range [1, 1].

3.6 ECTO

E-CTO is the updated form of the conventional CTO algorithm developed by Das et.al. in 2015. The algorithm basically based on the learning capability of the students at the college or school level where students are categorized into groups or sections for interaction and sharing of the problem solution where the competition is among themselves to be the section or group topper (S.T. or G.T.) and finally the student among the topper of the groups giving best solution considered as the class topper (C.T.) or overall topper (O.T.). The E-CTO was developed to remove the drawbacks like the slow convergence rate and inefficient interaction among sections or groups of students of the previously developed algorithm. The algorithm generally focuses on two following steps:

1. Initialization of the Population:—Population generation is important in evolutionary algorithms because it affects convergence speed and the overall quality of the end solution. The population of students in each class is first generated

Table 9 Chosen values for MLP Hyper-parameters

S. No	Description	Values
1	No. of hidden layers	100
2	Activation function	Relu
3	Penalty parameter (alpha)	0.0001
4	Weight optimization (solver)	Adam
5	Learning rate	0.001

Table 10 Performance evaluation parameters outcomes for MLP without Optimization

	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	MAX Error (W/m ²)	R ² Score	
					Train	Test
GHI	39.7855	6512.17	80.6980	785.96	0.92456	0.92583
DHI	39.4676	6448.24	80.3009	770.76	0.92504	0.92656
DNI	50.4714	9513.09	97.5351	735.0.99	0.91899	0.85271

at random during the early stages of E-CTO depending on the finest students in the initial generation of the population. Chaos initiation is accomplished via the logistical mapping method. The logistic map's output is produced using the following expression:

$$\alpha_{p+1}^{q+1} = c * \alpha_p^q (1 - \alpha_p^q) \quad (17)$$

here, c denotes the control parameter whose value has been taken as four and α_p^q signifies the number between 1 and 0 chosen randomly, for $p=1,2,\dots$ section and $q=1,2,\dots$ student. It is clear that in a certain situation, fundamental knowledge may be affected. The student's earlier understanding influences the performance to some extent. The explosion method is used to create complicated activation in the end, which increases student individuality. Each student's initial value in each segment is calculated using (42).

$$s_{a,b} = s_{\min,a} + (s_{\max,n} - s_{a,b}) \quad (18)$$

where $s_{a,b}^c$ represents the a th learner from the b th group or section. Then a new population sample is created using an opposition based learning method (OBLM) which generally improves the initial solutions and convergence rate.

2. Mutation:—Students in this algorithm follow the topper or best student of the section and the class. Therefore, if the topper gets stuck in the local minima, all learners would follow and the solution will converge there which is not an acceptable situation. For addressing this problem, a stochastic mutation parameter is incorporated which provides more motivation to the section topper to handle the local minima. After the evolutionary phase of the E-CTO is finished, the subsequent mutation equation is implemented to the topper of section: where, k_c has a range of 0 to c and r_v denotes the homogeneously distributed random number between -1 and 1. s_u^c and s_l^c are the upper and lower range of features.

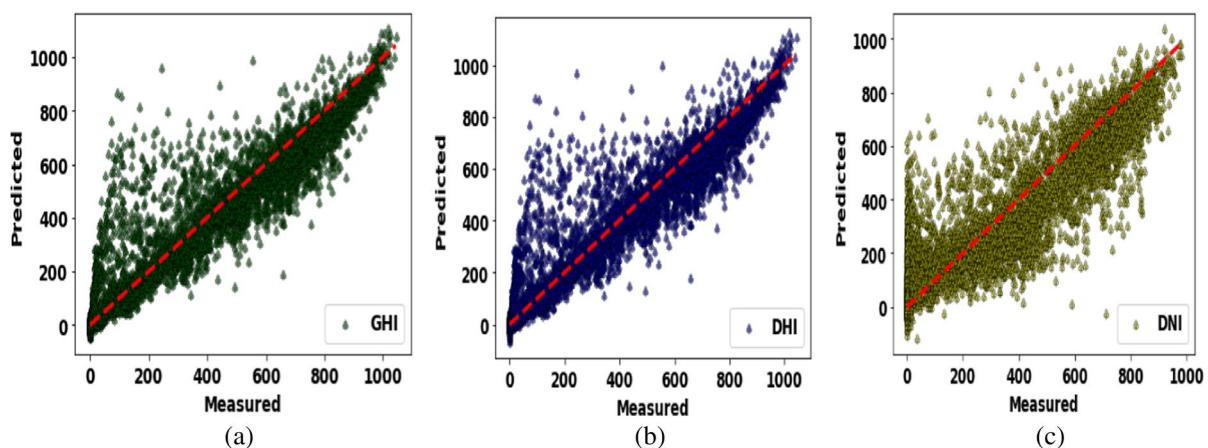


Fig. 16 Scatter plots of MLP model without optimization for (a) GHI (b) DHI(c) DNI

Table 11 Hyper-parameters range for MLP

Sl. No	Description	Lower Bound	Upper Bound
1	No. of hidden layers	10	100
2	Penalty parameter (alpha)	0.0001	0.1
3	Learning rate	0.001	0.1

$$s.t_{b,c}^{itr} = \begin{cases} s.t_{b,c} + r_v * (s_u^c - s_l^c), & \text{if } k = k_c \\ s.t_{b,c}, & \text{else} \end{cases} \quad (19)$$

3 . . . 7

Performance Parameters

To quantify the performance and their variation from the real value for the ML models, we provide several common statistical metrics. The difference between the estimated (or anticipated) and actual output parameter is known as the deviation sometimes referred to as the errors or residue. For example, the error for GHI can be expressed as:

$$\delta = GHI_{obs} - GHI_{pred} \quad (20)$$

These can be used to assess the degree of divergence and correlation between the predicted and actual data. Figure 10 depicts the expressions for forecasting the effectiveness of ML models in our research.

4 Results and Discussion

The major goal of this research is to extend the prediction accuracy of the solar irradiance of the three nominal machine learning models SVR, MLP and RFR. For which the 5 years of data have been taken from the NSRDB for analysis. As the initial data was in raw and improper form, the data pre-processing process has been carried out to feed the data to the learning model in the appropriate form. The processed data is split into two parts i.e. training and testing samples. The training sample carries 80 per cent and the testing sample carries 20 per cent of the data set. The trained model has been evaluated with the testing data where 6 performance parameters have been calculated to check the applicability of the model. As the nominal models have fixed parameters cannot assure the accuracy and reliability of the model. Therefore, in order to improve the performance of the three models two metaheuristic optimization methods MFO and GWO have used.

The population size of any optimization technique is an important parameter for the analysis as it has a larger effect on the computational time and convergence rate. A large population will make the running time much longer, making it hard to use the models to solve engineering problems. On the other hand, a

Table 12 Performance evaluation parameters outcomes for MLP-MFO

	Iterations	Population Size	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	MAX Error (W/m ²)	R ² Score	
							Train	Test
GHI	100	50	39.1977	6452.82	80.3294	777.15	0.92503	0.92648
		100	39.5914	6481.07	80.5051	768.95	0.92483	0.92619
		150	39.9944	6563.72	81.0168	784.26	0.92373	0.92525
		200	38.8147	6455.52	80.3462	775.38	0.92506	0.92651
		500	41.2802	6710.36	81.9168	795.85	0.92185	0.92358
DHI	100	50	23.9168	2041.34	45.1812	369.46	0.88476	0.88516
		100	23.4678	2009.16	44.8237	365.50	0.88580	0.88697
		150	24.2635	2070.66	45.5045	369.84	0.88286	0.88351
		200	23.9610	2089.30	45.7088	372.65	0.88167	0.88246
		500	23.4705	2044.36	45.2147	377.05	0.88458	0.88499
DNI	100	50	61.1930	9681.98	98.3971	712.67	0.84536	0.84854
		100	61.9136	9614.21	98.0521	726.47	0.84645	0.84953
		150	62.0220	9622.92	98.0965	732.62	0.84629	0.84940
		200	61.0919	9392.76	96.9163	710.43	0.84714	0.85021
		500	62.2068	9594.10	97.9495	738.62	0.84697	0.84982

Table 13 Performance evaluation parameters outcomes for MLP-GWO

	Iterations	Population Size	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	MAX Error (W/m ²)	R ² Score	
							Train	Test
GHI	100	50	42.2596	6996.44	83.6447	798.21	0.91793	0.92032
		100	39.4757	6447.21	80.2945	775.29	0.92522	0.92657
		150	39.8143	6456.55	80.3526	766.22	0.92513	0.92647
		200	39.6364	6556.21	80.9704	781.24	0.92384	0.92533
		500	40.4920	6642.97	81.5044	781.21	0.92259	0.92435
DHI	100	50	23.3568	2032.12	45.0790	370.15	0.88505	0.88567
		100	23.1382	2007.06	44.8002	367.81	0.88620	0.88708
		150	23.6088	2035.62	45.1178	376.50	0.88501	0.88548
		200	23.8901	2032.07	45.0785	369.69	0.88512	0.88568
		500	23.1514	2030.97	45.0633	376.95	0.88505	0.88574
DNI	100	50	62.19304	9593.55	97.9467	738.71	0.84773	0.84983
		100	62.10703	9603.68	97.9984	745.05	0.84700	0.84968
		150	62.23338	9411.21	97.0114	732.18	0.84661	0.84965
		200	62.05884	9191.01	95.8698	725.38	0.85067	0.85290
		500	62.05884	9359.03	96.7421	740.45	0.84737	0.85042

small population will cause fitness values to be unstable. In this study, five distinct population sizes—50, 100, 150, 200, and 500—were chosen to develop the hybrid models where iteration of the hybrid models has been taken 100.

The overall performance parameters have been calculated and summarized to conclude the model shows the least error and best accuracy. Further to validate the model and check the accuracy scope of the model if exist, the ECTO method has been applied and all

Table 14 Performance evaluation parameters outcomes for MLP-ECTO

	Iterations	Population Size	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	MAX Error (W/m ²)	R ² Score	
							Train	Test
GHI	100	50	40.1931	6838.17	82.6932	768.13	0.9299	0.9292
		100	38.9976	6386.11	79.9131	723.35	0.9342	0.9336
		150	39.1073	6425.45	80.1589	734.47	0.9335	0.9330
		200	39.1364	6434.25	80.2137	747.68	0.9326	0.9325
		500	40.1452	6573.87	81.0794	741.51	0.9323	0.9327
DHI	100	50	22.8612	2015.68	44.8963	369.06	0.8931	0.8939
		100	22.7291	1986.57	44.5709	366.74	0.8947	0.8944
		150	23.1126	2017.33	44.9146	371.16	0.8925	0.8924
		200	23.1284	2011.61	44.8509	367.81	0.8903	0.8836
		500	22.7533	2013.39	44.8708	370.22	0.8929	0.8938
DNI	100	50	61.6321	10,245.49	101.2200	736.65	0.8674	0.8685
		100	61.6309	10,378.20	101.8734	741.43	0.8679	0.8682
		150	61.7467	10,314.36	101.5596	730.21	0.8683	0.8677
		200	60.8944	9106.823	095.4296	722.17	0.8724	0.8711
		500	61.1135	10,584.72	102.8820	737.82	0.8701	0.8693

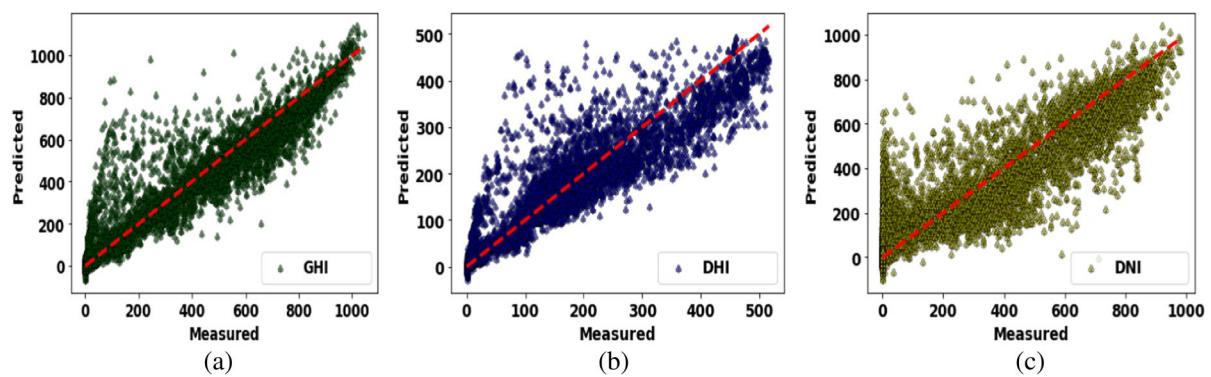


Fig. 17 Scatter plots of MLP model with MFO optimization for (a) GHI (b) DHI (c) DNI

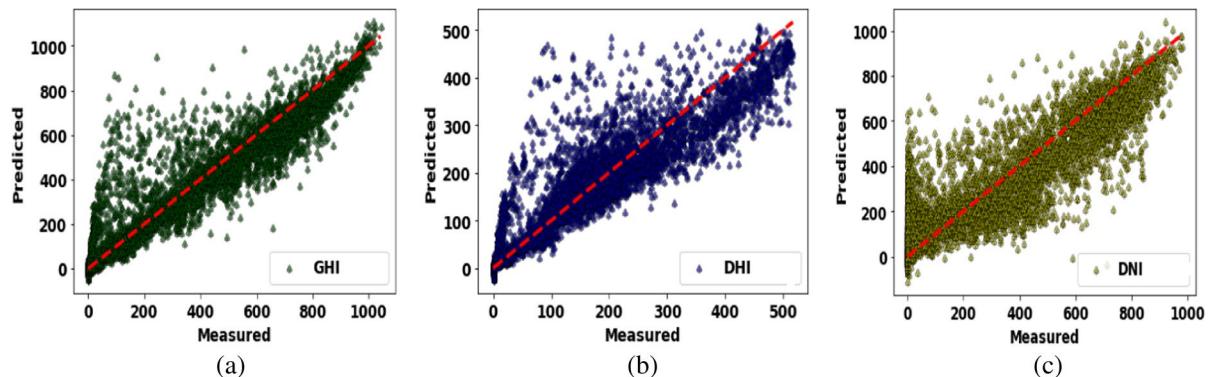


Fig. 18 Scatter plots of MLP model with GWO optimization for (a) GHI (b) DHI (c) DNI

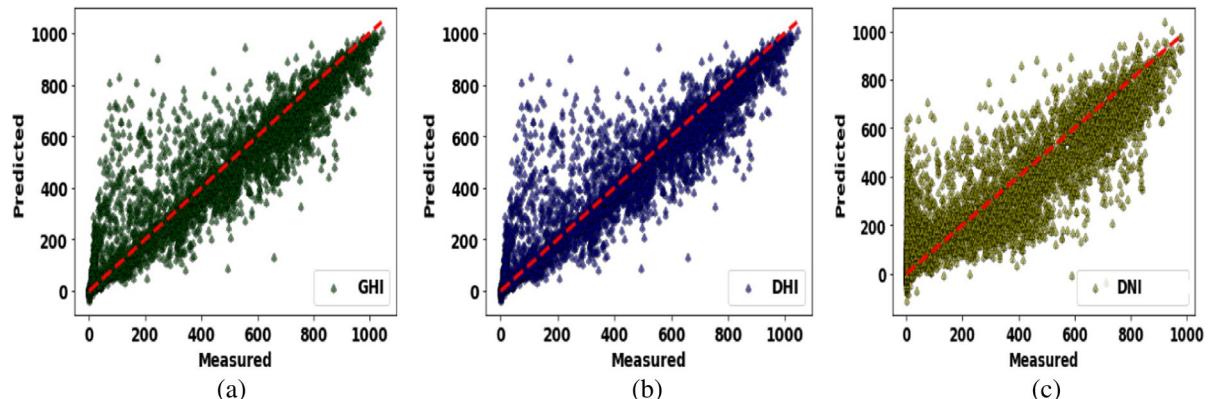


Fig. 19 Scatter plots of MLP model with ECTO optimization for (a) GHI (b) DHI (c) DNI

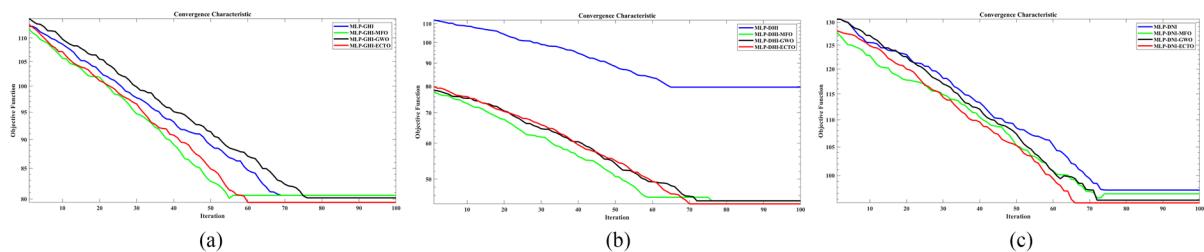


Fig. 20 Conversion curve plots of all MLP models for (a) GHI (b) DHI (c) DNI

the performance metrics have been analysed again for the ML models.

4.1 Performance analysis of SVR

SVR is most advantageous when used with non-linear datasets. In SVR, a variety of additional kernels are accessible, but to recent surveys, RBF is the most effective kernel. SVR maximises the margin in the vicinity of the separating hyperplanes. The radius of the RBF kernel functions acts as an effective structural regularizer on its own. The C parameter pays off proper training categorization versus optimization of the decision function's margins. Increasing C further is ineffective, most likely because there are no more training units in contravention (within the margin), or no better alternative can be identified because a very high C value often increases fitting time.

Lower C values, on the other hand, often result in additional support vectors, which might also increase prediction time. As a result, decreasing C entails an exchange between prediction time and fitting time. Tolerance denotes the stopping criteria. The parameter values for analysis of SVR have been shown in Table 3 while the performance evaluation parameters and their results for solar radiation data are tabulated in Table 4. The scatter plot between the true and predicted values of the SVR model without optimization has been shown in Fig. 11.

The results shown above can be improved with the proper selection of the parameters of support vector regressor. The hyperparameters range in order to optimize the SVR model has been shown in Table 5 and the performance parameters with respect to the hyperparameters range for MFO, GWO and ECTO methods have been represented in

Table 6, Table 7 and Table 8. The best values have been highlighted for each target variable.

The scatter plot for the best population size of both optimized models of SVR ie. SVR-MFO, SVR-GWO and SVR-ECTO for the target parameters have been shown in Fig. 12, Fig. 13 and Fig. 14.

The conversion curves for the RMSE loss has been shown in Fig. 15 where the ECTO method is giving the lowest error loss as compared to other two optimization methods.

4.2 Performance analysis of MLP

Multilayer perceptron can be considered as the neural network working in the forward direction. Layers that are deep and entirely linked allow it to transform any input dimension into the desired dimension. Multilayer perception refers to a neural network having several layers. The parameters chosen for the analysis have been represented in Table 9 and the results of the analysis as per the parameters selected for the MLP model as been shown in Table 10. The scatter plot between the true and predicted values of the MLP model without optimization has been shown in Fig. 16.

Table 15 Chosen values for RF Hyper-parameters

S. No	Description	Value
1	Maximum no. of trees	100
2	Maximum depth	5
3	Minimum samples leaf	1
4	Minimum samples split	2

Table 16 Performance evaluation parameters outcomes for RFR without Optimization

	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	MAX Error (W/m ²)	R ² Score	
					Train	Test
GHI	37.0127	6415.21	80.0950	774.63	0.97117	0.92881
DHI	35.2513	6331.42	79.5702	763.24	0.97172	0.92893
DNI	48.1427	9147.24	95.6412	683.27	0.91994	0.85743

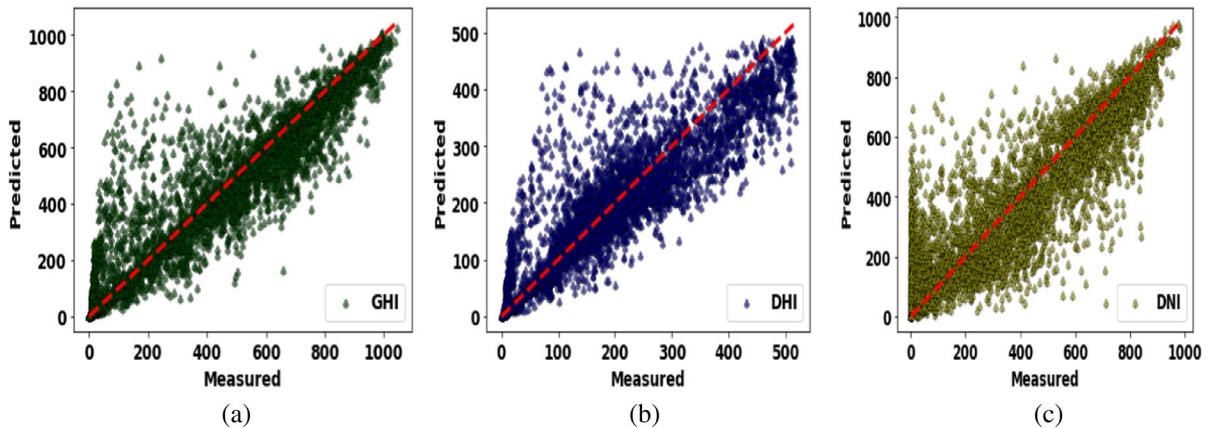


Fig. 21 Scatter plots of RFR model without optimization for (a) GHI (b) DHI(c) DNI

The speed and efficiency of the learning process are influenced by the learning rate, means the process of adjusting the weights. The hyperparameters must be tuned to get the optimum outputs. The respective hyperparameters range for the optimization of the MLP model with the other three optimization techniques has been tabulated under Table 11. The evaluation parameters for all three target variables has been shown with best-marked values in bold in Table 12, Table 13 and Table 14 for the MLP-MFO, MLP-GWO and MLP-ECTO for different population size.

The scatter plot for the best population size of both optimized models of MLP i.e. MLP-MFO and MLP-GWO for the target parameters has been shown in Fig. 17, Fig. 18 and Fig. 19.

The conversion of the RMSE loss for the MLP with the three optimization method has been presented in Fig. 20 where the RFR ECTO has shown the lowest error for all the target parameters.

4.3 Performance of RFR ML Model

The RF machine learning model can be utilized as a classification and regression of a tree. Generally,

subsampling happens once for every tree in RF. The parameters selected for the analysis of the RFR model without optimization have been shown in Table 15.

Increasing the depth of the tree makes the model more complicated and prone to overfitting. We may immediately obtain the new tree of the features after each boosting step and lowers the weights of the feature to make the boosting method more conservative. After using RFR over a solar radiation dataset of GHI, DHI, and DNI, their Performances and errors are tabulated under Table 16. The scatter plot between the true and predicted values of the RFR model without optimization has been shown in Fig. 21.

Table 17 Selected values for RFR Hyper-parameters

Sl. No	Description	Lower Bound	Upper Bound
1	Maximum no. of trees	100	1000
2	Maximum depth	5	50
3	Minimum samples leaf	1	5
4	Minimum samples split	2	10

Table 18 Performance evaluation parameters outcomes for RFR-MFO

	Iterations	Population Size	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	MAX Error (W/m ²)	R ² Score	
							Train	Test
GHI	100	50	33.7314	5938.27	77.0114	722.53	0.9621	0.9318
		100	32.4012	5889.32	76.8076	736.51	0.9727	0.9330
		150	32.8721	5932.84	77.0471	745.07	0.9749	0.9325
		200	33.2455	5978.27	77.3505	739.22	0.9588	0.9321
		500	33.2463	6017.65	77.6128	744.56	0.9517	0.9319
DHI	100	50	20.1511	1839.21	42.9153	358.77	0.9462	0.8963
		100	20.0917	1831.30	42.8274	361.35	0.9388	0.8968
		150	20.8929	1855.64	43.1260	359.68	0.9372	0.8953
		200	20.0223	1776.11	42.2097	345.74	0.9640	0.8997
		500	20.1209	1833.32	42.8827	362.23	0.9435	0.8965
DNI	100	50	48.7753	9253.27	96.2197	681.47	0.9609	0.8736
		100	49.0444	9271.49	96.3143	692.49	0.9407	0.8735
		150	48.5287	9065.02	95.2104	670.93	0.9506	0.8737
		200	49.3422	9372.85	96.8392	693.41	0.9345	0.8721
		500	49.6545	9442.31	97.1917	691.87	0.9322	0.8712

In order to optimize the RF model we have tabulated the range of parameters which have optimized using the two optimization methods i.e., MFO and GWO which has been shown in Table 17. The iteration for both methods has been fixed to 100 to analyse the effect of increasing population size.

The evaluation parameters for determining the effect of the MFO, GWO and ECTO optimization on

the RF model have been shown in Table 18, Table 19 and Table 20, where all the five parameters for the target variable GHI, DHI and DNI has been calculated for the various population size in order to check the best population of the nature-based algorithm with respect to the fixed iteration count. The best values for each parameter have been highlighted in bold.

Table 19 Performance evaluation parameters outcomes for RFR-GWO

	Iterations	Population Size	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	MAX Error (W/m ²)	R ² Score	
							Train	Test
GHI	100	50	33.3175	5819.83	76.3117	703.94	0.9826	0.9338
		100	32.3934	5813.79	76.2994	680.78	0.9889	0.9341
		150	33.7075	5901.93	76.8543	736.37	0.9659	0.9331
		200	32.7410	5937.29	77.0988	744.51	0.9618	0.9327
		500	32.9771	5973.33	77.3187	741.88	0.9611	0.9324
DHI	100	50	19.8533	1801.35	42.4886	360.35	0.9657	0.8988
		100	19.9671	1822.17	42.7679	361.34	0.9486	0.8981
		150	20.7486	1838.31	42.9342	352.57	0.9445	0.8966
		200	20.1607	1751.54	41.9088	350.08	0.9792	0.9015
		500	20.2066	1783.13	42.2493	355.49	0.9618	0.8999
DNI	100	50	48.5421	9052.73	95.1774	671.62	0.9678	0.8742
		100	48.9007	9236.14	96.1304	685.35	0.9481	0.8743
		150	48.3932	9052.49	95.1446	668.57	0.9616	0.8749
		200	48.9636	9239.91	96.1498	686.51	0.9439	0.8741
		500	49.0021	9385.62	96.9159	690.16	0.9313	0.8724

Table 20 Performance evaluation parameters outcomes for RFR-ECTO

	Iterations	Population Size	MAE (W/m ²)	MSE (W/m ²)	RMSE (W/m ²)	MAX Error (W/m ²)	R ² Score	
							Train	Test
GHI	100	50	32.5316	5758.18	75.8827	654.16	0.9792	0.9402
		100	31.8532	5754.93	75.8613	612.37	0.9813	0.9411
		150	32.7542	5812.88	76.2423	663.62	0.9776	0.9387
		200	31.2527	5793.12	76.1126	672.18	0.9754	0.9379
		500	31.9132	5806.91	76.2031	677.23	0.9737	0.9371
DHI	100	50	19.1876	1700.42	41.2362	355.32	0.9786	0.9088
		100	18.8447	1717.76	41.4459	351.25	0.9694	0.9081
		150	19.8018	1746.66	41.7931	350.19	0.9683	0.9079
		200	18.8391	1670.47	40.8714	346.73	0.9848	0.9107
		500	19.4275	1677.45	40.9567	349.69	0.9771	0.9093
DNI	100	50	47.7462	9049.50	95.1289	664.67	0.9718	0.8859
		100	47.7754	9006.33	94.9017	675.13	0.9683	0.8865
		150	46.9011	9004.41	94.8916	657.58	0.9727	0.8882
		200	47.8547	9058.73	95.1774	679.47	0.9665	0.8837
		500	48.1568	9070.40	95.2387	683.72	0.9574	0.8815

The graphical representation of the accuracy of the RFR-MFO, RFR-GWO and RFR-ECTO for various population sizes has been shown in Fig. 22, Fig. 23 and Fig. 24.

The comparative analysis of the three models i.e. RFR-MFO, RFR-GWO and RFR-ECTO has been

shown in Fig. 25, from which it can be observed that the value for R² score (validation/testing data) for GHI, DHI and DNI have higher values for the RF-ECTO and lesser statistical errors as compared to other two models.

The scatter plot between measured and predicted values of the output variables for the best population

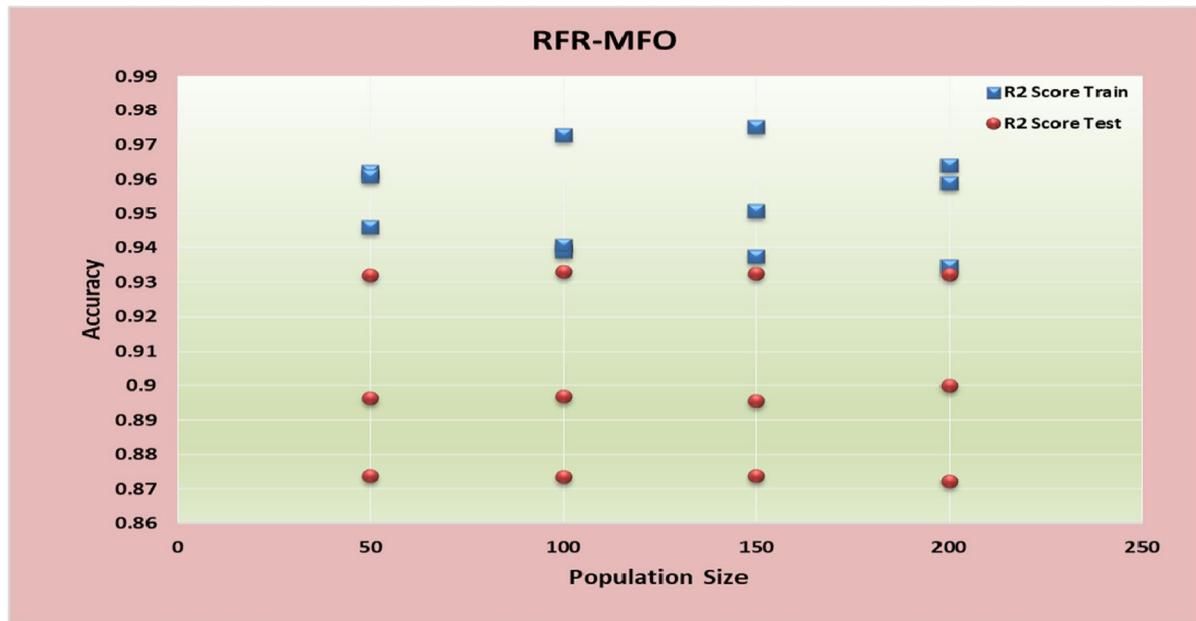
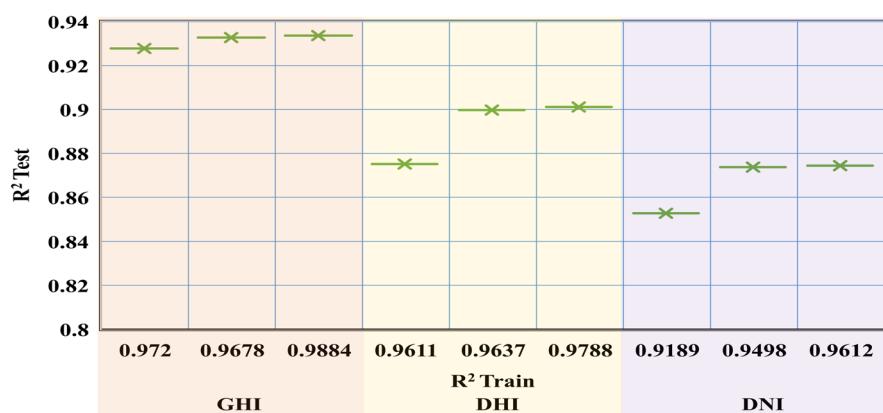
**Fig. 22** R² Score analysis using RFR-MFO

Fig. 23 R^2 Score analysis using RFR-GWO



size of both optimized models of RFR i.e. RFR-MFO, RFR-GWO and RFR-ECTO has been shown in Fig. 26, Fig. 27 and Fig. 28.

The above curves for the RMSE loss for the unoptimized and three optimized hybrid RFR models have been shown in Fig. 29 which confirms that the ECTO techniques has better capability to optimize the models as compared to other two techniques.

The above analysis confirms that the performance of the models can be improved by selection of the best parameters of the model. The best optimized model can be utilized for the real time prediction which can benefit the solar power operators to take necessary action within time frame for establishment of balance between demand and supply of

the energy which is needed to avoid the fluctuation in voltage or frequency of the grid connected PV system. The primary benefit of using this proposed model is the possibility of reducing the meteorological forecast error, which makes up the majority of the overall error. The challenge of handling integration issues grows as more PV power generating systems are incorporated into the primary energy grid by residential and commercial customers. Hence it is recommended that the further strategies or techniques like new optimization method and bi-level or multilevel approach can be done for the improvement of the machine learning models so that the forecasting error could be minimized which would definitely help the power plant operator to utilise

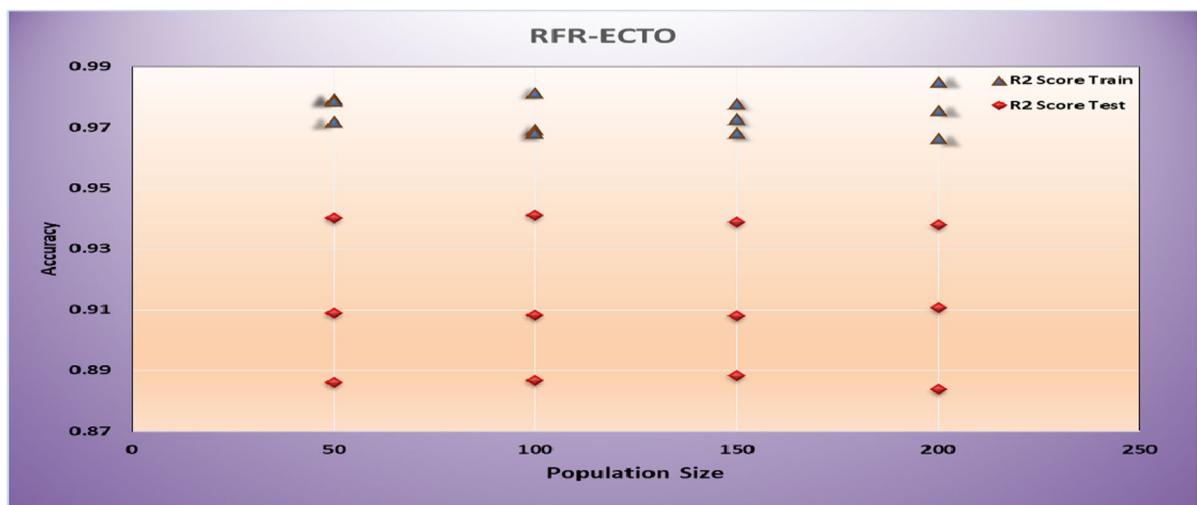


Fig. 24 R^2 Score analysis using RFR-ECTO

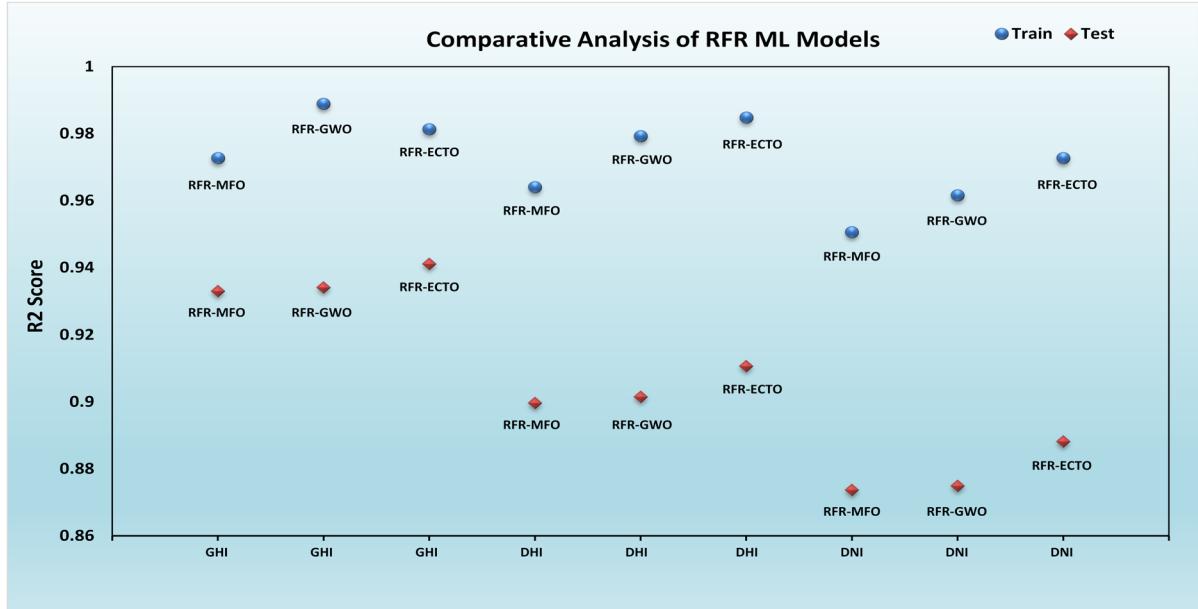


Fig. 25 Comparative analysis of all RFR models based on R^2 score

the accurate predicted power to plan, make decisions, and control distribution of electrical energy in any circumstances.

4.4 Comparative Analysis

In order to have the best machine-optimized ML models a comparative analysis of all 9 models has

been done and validated with the latest optimization method ECTO. The overall comparison of machine learning models has been shown in Table 21 where the best of the evaluated parameters of all models has been marked in bold, but to have the best model it requires to choose the best among all which has been done taking accuracy as the base performance criteria.

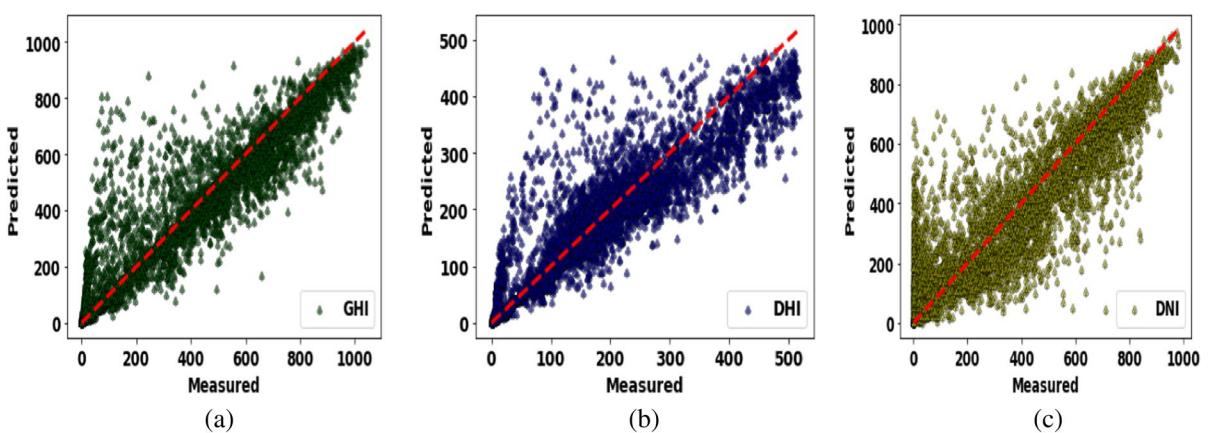


Fig. 26 Scatter plots of RFR model with MFO optimization for (a) GHI (b) DHI (c) DNI

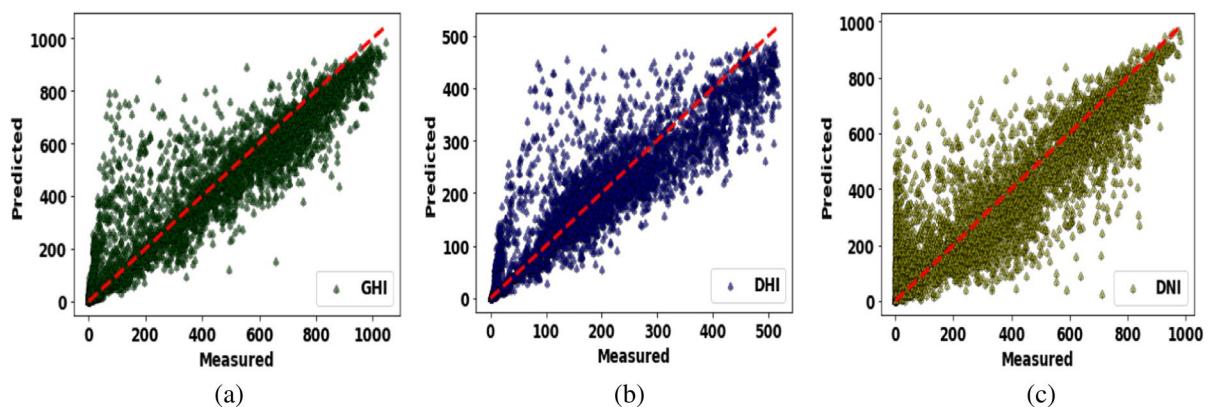


Fig. 27 Scatter plots of RFR model with GWO optimization for (a) GHI (b) DHI(c) DNI

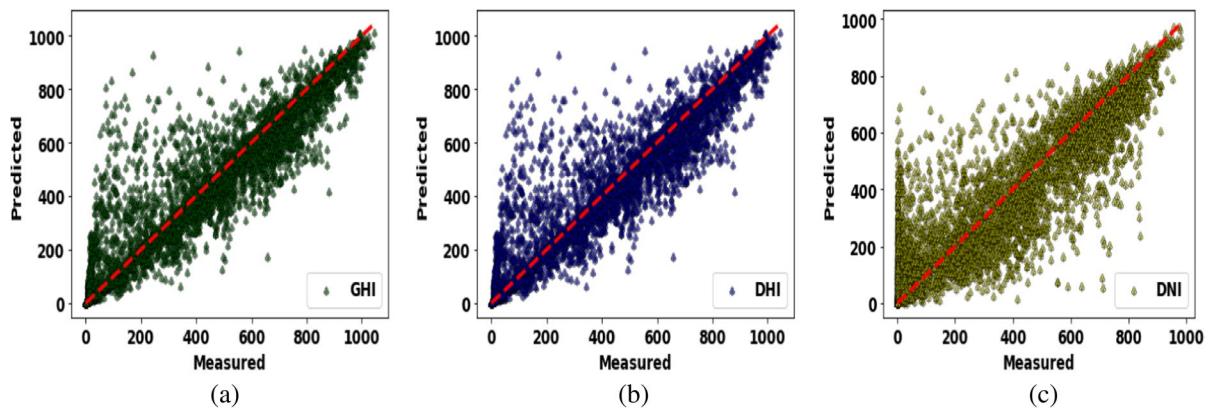


Fig. 28 Scatter plots of RFR model with ECTO optimization for (a) GHI (b) DHI(c) DNI

From the above table it can be observed that the RFR-GWO has performed better as compared to others ML models as having higher accuracy and lesser errors. The graphical comparative analysis of the most important determining performance parameter RMSE has been shown in Fig. 30 in the form of radar

plot where the RMSE score varying from the center to the outer diameter along the radius indicates that models having lower RMSE values lying near to origin with dark red color while yellow color showing the higher RMSE values. Hence we can conclude from the graph that the ML models tuned with the

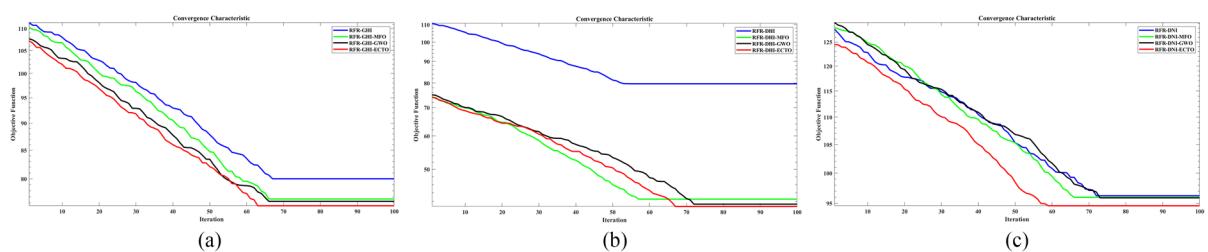
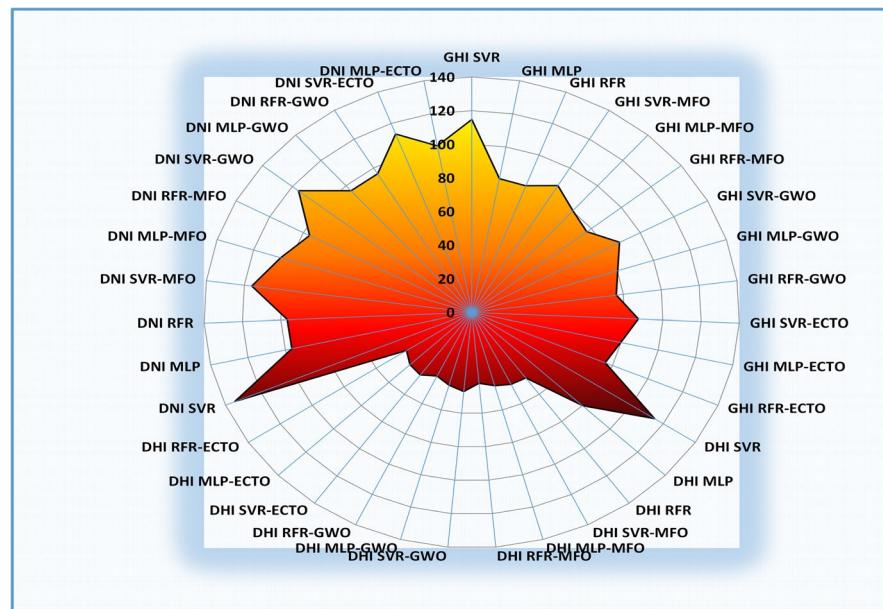


Fig. 29 Conversion curve plots of all RFR models for (a) GHI (b) DHI(c) DNI

Table 21 ML models comparative analysis

	ML Models	MAE	MSE	RMSE	ME	R2	
						Train	Test
GHI	SVR	73.2215	13,138.95	114.6252	796.87	0.8514	0.8503
	MLP	39.7855	6512.17	80.6980	785.96	0.9245	0.9258
	RF	37.0127	6415.21	80.0950	774.63	0.9711	0.9288
	SVR-MFO	45.8923	7694.22	87.7167	883.66	0.9120	0.9123
	MLP-MFO	38.8147	6455.52	80.3462	775.38	0.9250	0.9264
	RFR-MFO	32.4012	5889.32	76.8076	736.51	0.9727	0.9330
	SVR-GWO	45.8393	7689.83	87.6917	883.73	0.9121	0.9124
	MLP-GWO	39.4757	6447.21	80.2945	775.29	0.9252	0.9265
	RFR-GWO	32.3934	5813.79	76.2994	680.78	0.9889	0.9341
	SVR-ECTO	44.4318	7588.92	87.1144	882.65	0.92463	0.92394
	MLP-ECTO	38.9976	6386.11	79.9131	723.35	0.9342	0.9336
	RFR-ECTO	31.8532	5754.93	75.8613	612.37	0.9813	0.9411
DHI	SVR	73.1946	13,124.34	114.5615	774.62	0.8523	0.8514
	MLP	39.4676	6448.24	80.3009	770.76	0.9250	0.9265
	RFR	35.2513	6331.42	79.5702	763.24	0.9717	0.9289
	SVR-MFO	25.4859	2244.81	47.3795	423.60	0.8733	0.8737
	MLP-MFO	23.4678	2009.16	44.8237	365.50	0.8858	0.8869
	RFR-MFO	20.0223	1776.11	42.2097	345.74	0.9640	0.8997
	SVR-GWO	25.2275	2229.93	47.2221	424.04	0.8741	0.8745
	MLP-GWO	23.1382	2007.06	44.8002	367.81	0.8862	0.8870
	RFR-GWO	20.1607	1751.54	41.9088	350.08	0.9792	0.9015
	SVR-ECTO	23.1511	2106.95	45.9015	421.87	0.8847	0.8829
	MLP-ECTO	22.7291	1986.57	44.5709	366.74	0.8947	0.8944
	RFR-ECTO	18.8391	1670.47	40.8714	346.73	0.9848	0.9107
DNI	SVR	82.3159	18,149.01	134.7182	708.69	0.7506	0.7525
	MLP	50.4714	9513.09	97.5351	735.99	0.9189	0.8527
	RFR	48.1427	9147.24	95.6412	683.27	0.9199	0.8574
	SVR-MFO	64.9512	13,471.83	116.0682	739.03	0.8125	0.8163
	MLP-MFO	61.0919	9392.76	96.9163	710.43	0.8471	0.8502
	RFR-MFO	48.5287	9065.02	95.2103	670.93	0.9506	0.8737
	SVR-GWO	64.7669	13,425.59	115.8688	738.12	0.8131	0.8169
	MLP-GWO	62.0588	9191.01	95.8698	725.38	0.8506	0.8529
	RFR-GWO	48.3932	9052.49	95.1446	668.57	0.9616	0.8749
	SVR-ECTO	63.7241	12,814.71	113.2020	737.69	0.8309	0.8303
	MLP-ECTO	60.8944	9106.82	95.4296	722.17	0.8724	0.8711
	RFR-ECTO	46.9011	9004.41	94.8916	657.58	0.9727	0.8882

Fig. 30 Comparative analysis of all models based on RMSE



ECTO method have lowest value as compared to the unoptimized models and other two 6 hybrid models.

The R^2 score of the all the nominal and hybrid models for the target variable GHI has been shown in Fig. 31 with the bar plot where values are varying from bottom to top of the cylindrical bar with red to yellow color. The values of the R^2 score has been placed on the on the top which shows that the

RFR-ECTO hybrid model has greater accuracy among all ML models.

The R^2 score of each ML model for the second output parameter, DHI, is displayed in Fig. 32 as a pie plot with distinct coloured segments that indicate the values. The R^2 score values have likewise been positioned at the top, highlighting that the ECTO approach has optimised the three ML models the best.

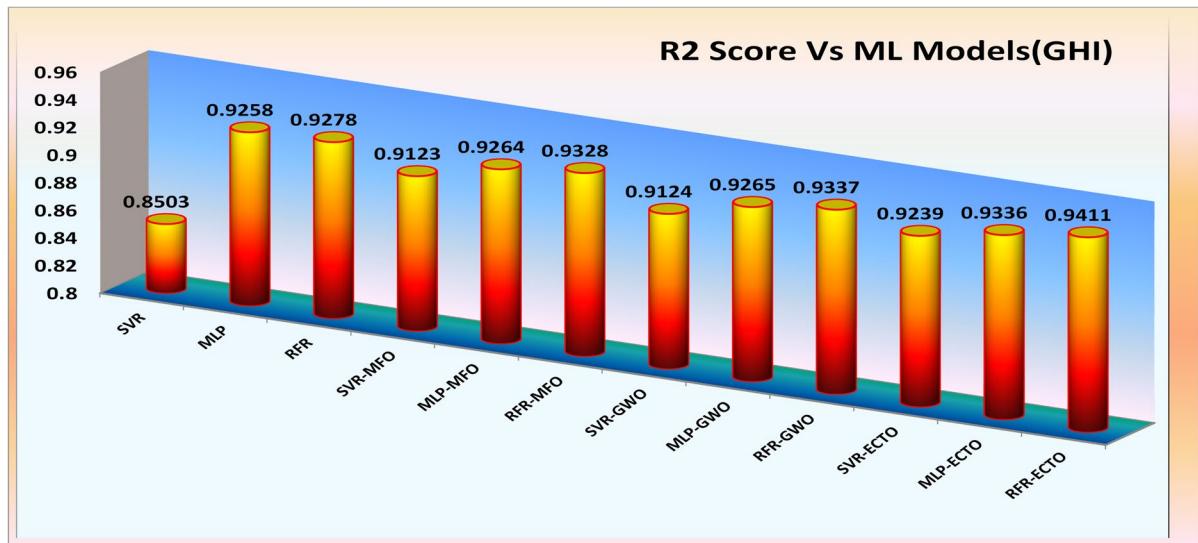
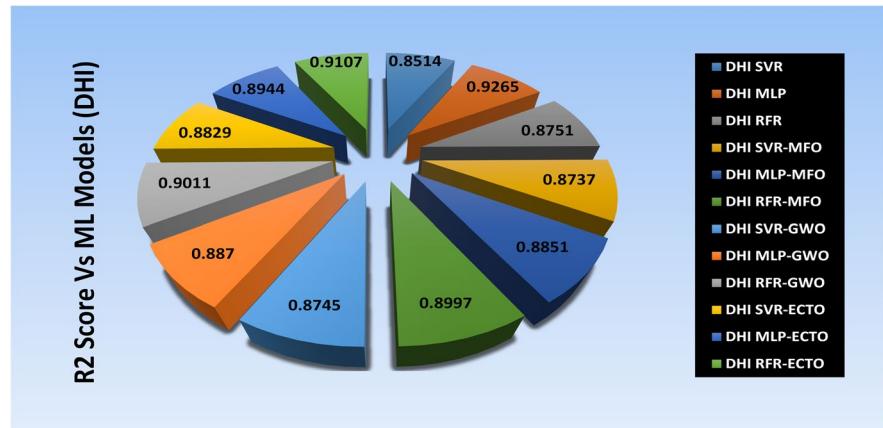


Fig. 31 Comparative analysis of all models based on R^2 score for GHI

Fig. 32 Comparative analysis of all models based on R^2 score for DHI



Finally the R^2 score of the third target parameter DNI has been shown in Fig. 33 with the help of the radar plot where values are varying from the origin of the radar to its circumferences. The values at the circumference are higher confirming the highest accuracy of the model while the values near the origin are showing the lowest accuracy among the above discussed models. The DNI radar plot again confirms that the ECTO has outperformed the other two optimization methods the ECTO as it has best optimized the chosen ML models.

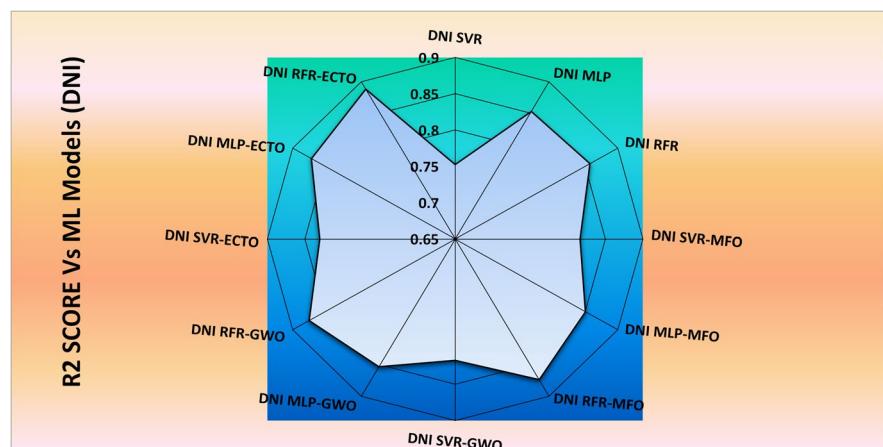
5 Conclusions

The performance of the learning algorithms and the data are both needed for a reliable machine learning model. Advanced learning algorithms must be designed and trained to utilize the available real-world

data, i.e. the true value of the data and the knowledge associated to the desired application before the system can assist in intelligent decision-making. The prediction of accurate solar irradiance is very useful for the forecasting of solar energy.

The main goal of this research is to optimize the supervised machine learning models with hyper-parameter tuning using the most prominent optimization algorithms, such as GWO, and MFO, in an effort to increase the prediction accuracy which can be useful for engineering practice. As a result, in this study, the 6 hybrid models (SVR-MFO, SVR-GWO, MLP-MFO, SVR-GWO, RFR-MFO and RFR-GWO) have been designed and analysed for the prediction purpose. Further to check the optimized machine learning model the ECTO technique has been applied for validation purposes. Five statistical parameters were chosen to assess the consistency between the true and the forecasted value in order to

Fig. 33 Comparative analysis of all models based on R^2 score for DNI



examine the performance of each hybrid model fully. An unoptimized supervised model was also developed, verified, and trained using the NREL historical datasets in order to assess the performance of the aforementioned two optimization techniques. The experimental findings show that both in the training stage and the test stage, all RFR-based hybrid models performed much better than the other machine learning models. The three RFR-based hybrid models' prediction accuracy exceeded 0.9 during testing, particularly the RFR-ECTO model has outperformed all other optimization techniques. The R^2 score or accuracy of the best hybrid supervised learning model for GHI, DHI and DNI are 94.11%, 91.07% and 88.82% respectively which confirms the model's excellent prediction ability and application to solve real-world problems.

Abbreviations IEA: International Energy Agency; IRENA: International Renewable Energy Agency; PV: Photovoltaic; GW: Gigawatt; LR: Linear Regression; ANN: Artificial Neural Network; XGB: Extreme Gradient Boosting; KNN: K-Nearest Neighbours; GPR: Gaussian Process Regression; DNN: Deep Neural Network; PSO: Particle Swarm Optimization; SVM: Support Vector Machine; LGBM: Light Gradient Boosting Machine; NREL: National Renewable Energy Laboratory; BMA: Bayesian Model Averaging; ANFIS: Adaptive Neuro Fuzzy Interface System; MVO: Multiverse Optimization; FFA: Firefly Algorithm; IGWO: Improved Grey Wolf Optimization; SSA: Sarp Swarm Algorithm; WOA: Whale Optimization Algorithm

Authors' Contributions Mantosh Kumar has collected and analysed the data along with computation and mathematical modelling for the methodology adopted. Kumari Namrata supervised the project and formatted the manuscript. Nishant Kumar and Gaurav Saini have worked in optimization technique formulation and also assisted in editing and formulating the manuscript.

Data Availability On request, the data utilised in this work may be made public.

Declarations

Ethics Approval and Consent to participate Not applicable.

Consent for Publication Not applicable.

Competing Interests Authors have no competing interests associated with this research work.

References

- Guo, Z., Zhou, K., Zhang, C., Lu, X., Chen, W., Yang, S.: Residential electricity consumption behavior: Influencing factors, related theories and intervention strategies. *Renew. Sustain. Energy Rev.* **81**, 399–412 (2018). <https://doi.org/10.1016/j.rser.2017.07.046>
- Zhang, Y., He, C.Q., Tang, B.J., Wei, Y.M.: China's energy consumption in the building sector: A life cycle approach. *Energy Build.* **94**, 240–251 (2015). <https://doi.org/10.1016/J.ENBUILD.2015.03.011>
- IEA: Global Energy Review: CO₂ Emissions in 2021.: Global Emissionrebound sharply to highest ever level. (2021). <https://www.iea.org/reports/global-energy-review-co2-emissions-in-2021-2>. Accessed 22 Sept 2022
- Samadhiya, A., Namrata, K., Kumar, N.: An Experimental Performance Evaluation and Management of a Dual Energy Storage System in a Solar Based Hybrid Microgrid. *Arab. J. Sci. Eng.* **2022**, 1–24 (2022). <https://doi.org/10.1007/S13369-022-07023-W>
- Nourani, V., Elkiran, G., Abdullahi, J., Tahsin, A.: Multi-region Modeling of Daily Global Solar Radiation with Artificial Intelligence Ensemble. *Nat. Resour. Res.* **28**, 1217–1238 (2019). <https://doi.org/10.1007/S11053-018-09450-9>
- Ahmad, T., Zhang, D., Huang, C.: Methodological framework for short-and medium-term energy, solar and wind power forecasting with stochastic-based machine learning approach to monetary and energy policy applications. *Energy* **231**, 120911 (2021). <https://doi.org/10.1016/J.ENERGY.2021.120911>
- Kumar, N., Namrata, K., Samadhiya, A.: Bi-level decision making in techno-economic planning and probabilistic analysis of community based sector-coupled energy system. *Appl. Intell.* **2022**, 1–25 (2022). <https://doi.org/10.1007/S10489-022-03794-9>
- Badescu, V.: A new kind of cloudy sky model to compute instantaneous values of diffuse and global solar irradiance. *Theor. Appl. Climatol.* **72**, 127–136 (2002). <https://doi.org/10.1007/s007040200017>
- Yadav, A.K., Chandel, S.S.: Solar radiation prediction using Artificial Neural Network techniques: A review. *Renew. Sustain. Energy Rev.* **33**, 772–781 (2014). <https://doi.org/10.1016/J.RSER.2013.08.055>
- Chen, J.L., Li, G.S., Xiao, B.B., Wen, Z.F., Lv, M.Q., Chen, C.D., Jiang, Y., Wang, X.X., Wu, S.J.: Assessing the transferability of support vector machine model for estimation of global solar radiation from air temperature. *Energy Convers. Manag.* **89**, 318–329 (2015). <https://doi.org/10.1016/J.ENCONMAN.2014.10.004>
- Mayer, M.J.: Benefits of physical and machine learning hybridization for photovoltaic power forecasting. *Renew. Sustain. Energy Rev.* **168**, 112772 (2022). <https://doi.org/10.1016/J.RSER.2022.112772>

12. Jebli, I., Belouadha, F.Z., Kabbaj, M.I., Tilioua, A.: Prediction of solar energy guided by pearson correlation using machine learning. *Energy* **224**, 120109 (2021). <https://doi.org/10.1016/J.ENERGY.2021.120109>
13. Benlebna, S., Kumar, N.M., Tahri, A.: Realtime monitoring on the HIT photovoltaic module characteristic parameters at STC, high and low irradiance conditions in Algeria. *Procedia Comput. Sci.* **132**, 1238–1242 (2018). <https://doi.org/10.1016/J.PROCS.2018.05.039>
14. Nazir, M.S., Almasoudi, F.M., Abdalla, A.N., Zhu, C., Alatawi, K.S.S.: Multi-objective optimal dispatching of combined cooling, heating and power using hybrid gravitational search algorithm and random forest regression: Towards the microgrid orientation. *Energy Rep.* **9**, 1926–1936 (2023). <https://doi.org/10.1016/J.EGYR.2023.01.012>
15. Tu, J., Hu, L., Mohammed, K.J., Le, B.N., Chen, P., Ali, E., Ali, H.E., Sun, L.: Application of logistic regression, support vector machine and random forest on the effects of titanium dioxide nanoparticles using macroalgae in treatment of certain risk factors associated with kidney injuries. *Environ. Res.* **220**, 115167 (2023). <https://doi.org/10.1016/J.ENVRES.2022.115167>
16. Moosa, A., Shabir, H., Ali, H., Darwade, R., Gite, B.: Predicting solar radiation using machine learning techniques. In: Proceedings of the 2nd International Conference on Intelligent Computing and Control Systems, ICICCS 2018, pp. 1693–1699. Institute of Electrical and Electronics Engineers Inc. (2019)
17. Husain, S., Khan, U.A.: Machine learning models to predict diffuse solar radiation based on diffuse fraction and diffusion coefficient models for humid-subtropical climatic zone of India. *Clean. Eng. Technol.* **5**, 100262 (2021). <https://doi.org/10.1016/J.CLET.2021.100262>
18. Manoj Kumar, N., Subathra, M.S.P.: Three years ahead solar irradiance forecasting to quantify degradation influenced energy potentials from thin film (a-Si) photovoltaic system. *Results Phys.* **12**, 701–703 (2019). <https://doi.org/10.1016/J.RINP.2018.12.027>
19. Lee, J., Wang, W., Harrou, F., Sun, Y.: Reliable solar irradiance prediction using ensemble learning-based models: A comparative study. *Energy Convers. Manag.* **208**, 112582 (2020). <https://doi.org/10.1016/J.ENCONMAN.2020.112582>
20. Azimi, Y., Talaeeian, M., Sarkheil, H., Hashemi, R., Shiradm, R.: Developing an evolving multi-layer perceptron network by genetic algorithm to predict full-scale municipal wastewater treatment plant effluent. *J. Environ. Chem. Eng.* **10**, 108398 (2022). <https://doi.org/10.1016/J.JECE.2022.108398>
21. Massaoudi, M., Refaat, S.S., Chihi, I., Trabelsi, M., Oueslati, F.S., Abu-Rub, H.: A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for Short-Term Load Forecasting. *Energy* **214**, 118874 (2021). <https://doi.org/10.1016/J.ENERGY.2020.118874>
22. Kumar, P., Nair, G.G.: An efficient classification framework for breast cancer using hyper parameter tuned Random Decision Forest Classifier and Bayesian Optimization. *Biomed. Signal Process. Control* **68**, 102682 (2021). <https://doi.org/10.1016/J.BSPC.2021.102682>
23. Fan, J., Zheng, J., Wu, L., Zhang, F.: Estimation of daily maize transpiration using support vector machines, extreme gradient boosting, artificial and deep neural networks models. *Agric. Water Manag.* **245**, 106547 (2021). <https://doi.org/10.1016/J.AGWAT.2020.106547>
24. Seyedmohammadi, J., Zeinadini, A., Navidi, M.N., McDowell, R.W.: A new robust hybrid model based on support vector machine and firefly meta-heuristic algorithm to predict pistachio yields and select effective soil variables. *Ecol. Inform.* **74**, 102002 (2023). <https://doi.org/10.1016/J.ECOINF.2023.102002>
25. Nguyen, H.D., Truong, G.T., Shin, M.: Development of extreme gradient boosting model for prediction of punching shear resistance of r/c interior slabs. *Eng. Struct.* **235**, 112067 (2021). <https://doi.org/10.1016/J.ENGSTRUCT.2021.112067>
26. Li, S., Xu, K., Xue, G., Liu, J., Xu, Z.: Prediction of coal spontaneous combustion temperature based on improved grey wolf optimizer algorithm and support vector regression. *Fuel* **324**, 124670 (2022). <https://doi.org/10.1016/J.FUEL.2022.124670>
27. Chia, M.Y., Huang, Y.F., Koo, C.H.: Swarm-based optimization as stochastic training strategy for estimation of reference evapotranspiration using extreme learning machine. *Agric. Water Manag.* **243**, 106447 (2021). <https://doi.org/10.1016/J.AGWAT.2020.106447>
28. Bazrafshan, O., Ehteram, M., Dashti Latif, S., Feng Huang, Y., Yenn Teo, F., Najah Ahmed, A., El-Shafie, A.: Predicting crop yields using a new robust Bayesian averaging model based on multiple hybrid ANFIS and MLP models. *Ain Shams Eng. J.* **13**, 101724 (2022). <https://doi.org/10.1016/J.ASEJ.2022.101724>
29. Liu, R., Li, G., Wei, L., Xu, Y., Gou, X., Luo, S., Yang, X.: Spatial prediction of groundwater potentiality using machine learning methods with Grey Wolf and Sparrow Search Algorithms. *J. Hydrol.* **610**, 127977 (2022). <https://doi.org/10.1016/J.JHYDROL.2022.127977>
30. Bouzateur, I., Bennacer, H., Ouali, M.A., Ziane, M.I., Hadjab, M., Ladjal, M.: A new ANN-PSO framework to chalcopyrite's energy band gaps prediction. *Mater. Today Commun.* **34**, 105311 (2023). <https://doi.org/10.1016/J.MTCOMM.2023.105311>
31. Li, S., Kong, X., Yue, L., Liu, C., Ahmad, M.: Short-term electrical load forecasting using hybrid model of manta ray foraging optimization and support vector regression. *J. Clean. Prod.* **388**, 135856 (2023). <https://doi.org/10.1016/j.jclepro.2023.135856>
32. Kabir, M.F., Chen, T., Ludwig, S.A.: A performance analysis of dimensionality reduction algorithms in machine learning models for cancer prediction. *Healthc. Anal.* **3**, 100125 (2023). <https://doi.org/10.1016/J.HEALTH.2022.100125>
33. Zhou, J., Dai, Y., Tao, M., Khandelwal, M., Zhao, M., Li, Q.: Estimating the mean cutting force of conical picks using random forest with salp swarm algorithm. *Results Eng.* **17**, 100892 (2023). <https://doi.org/10.1016/J.RINENG.2023.100892>
34. Hashem Samadi, S., Ghobadian, B., Nosrati, M., Rezaei, M.: Investigation of factors affecting performance of a downdraft fixed bed gasifier using optimized MLP neural

- networks approach. *Fuel* **333**, 126 (2023). <https://doi.org/10.1016/J.FUEL.2022.126249>
35. NSRDB: National Solar Radiation Database (NSRDB:NREL). <https://nsrdb.nrel.gov/data-viewer/query/query-tools>. Accessed 10 Sept 2022
36. Sun, L., Wang, T., Ding, W., Xu, J., Lin, Y.: Feature selection using Fisher score and multilabel neighborhood rough sets for multilabel classification. *Inf. Sci. (Ny)* **578**, 887–912 (2021). <https://doi.org/10.1016/J.INS.2021.08.032>
37. Salem, O.A.M., Liu, F., Chen, Y.P.P., Hamed, A., Chen, X.: Fuzzy joint mutual information feature selection based on ideal vector. *Expert Syst. Appl.* **193**, 116453 (2022). <https://doi.org/10.1016/J.ESWA.2021.116453>
38. Cortes, C., Vapnik, V., Saitta, L.: Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995). <https://doi.org/10.1007/BF00994018>
39. Deo, R.C., Wen, X., Qi, F.: A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset. *Appl. Energy* **168**, 568–593 (2016). <https://doi.org/10.1016/J.APENERGY.2016.01.130>
40. Lima, M.A.F.B., Carvalho, P.C.M., Fernández-Ramírez, L.M., Braga, A.P.S.: Improving solar forecasting using Deep Learning and Portfolio Theory integration. *Energy* **195**, 117016 (2020). <https://doi.org/10.1016/J.ENERGY.2020.117016>
41. Mellit, A., Pavan, A.M.: A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Sol. Energy* **84**, 807–821 (2010). <https://doi.org/10.1016/J.SOLENER.2010.02.006>
42. Breiman, L.: Random Forests. *Mach. Learn.* **45**, 5–32 (2001). <https://doi.org/10.1023/A:1010933404324>
43. Mirjalili, S.: Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowledge-Based Syst.* **89**, 228–249 (2015). <https://doi.org/10.1016/J.KNOSYS.2015.07.006>
44. Mirjalili, S., Mirjalili, S.M., Lewis, A.: Grey Wolf Optimizer. *Adv. Eng. Softw.* **69**, 46–61 (2014). <https://doi.org/10.1016/J.ADVENGSOFT.2013.12.007>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.