International Journal of Ambient Energy

ISSN: 0143-0750 (Print) 2162-8246 (Online) Journal homepage: http://www.tandfonline.com/loi/taen20

Hybrid Models for Global Solar Radiation Predection: case of study

Abdelaziz Rabehi, Mawloud Guermoui & Djemoui Lalmi

To cite this article: Abdelaziz Rabehi, Mawloud Guermoui & Djemoui Lalmi (2018): Hybrid Models for Global Solar Radiation Predection: case of study, International Journal of Ambient Energy, DOI: 10.1080/01430750.2018.1443498

To link to this article: https://doi.org/10.1080/01430750.2018.1443498

Accepted author version posted online: 20 Feb 2018.

Publisher: Taylor & Francis & Informa UK Limited, trading as Taylor & Francis Group

Journal: *International Journal of Ambient Energy*

DOI: 10.1080/01430750.2018.1443498

Hybrid Models for Global Solar Radiation Predection: case of study

Abdelaziz Rabehi, Mawloud Guermoui, Djemoui Lalmi

Unité de Recherche Appliquée en Energies Renouvelables, URAER, Centre de Développement des Energies Renouvelables, CDER, 47133, Ghardaïa, Algeria

Corresponding author's contact information: Phone: +213660444461

Phax:

Email: rab_ehi@hotmail.fr

Hybrid Models for Global Solar Radiation Predection: case of study

Abstract

This paper, present a comparative study between different predection models for solar radiation application . The current study assess the performance of Multi-layer perceptron (MLP), Boosted decision tree (BDT) and a new combinition of these models with linear regression (LR) was employed for daily global solar irradiation (DGSR) prediction. The performance of the studied models was vaalidated using a real dataset mesured at the Applied Research Unit for Renewable Energies, (URAER), situated in the south of Algeria. Differents inputs combination have been analyzed in order to select the relevants inputs parameters for DGSR prediction. The acheived results show that the MLP-model perfom better than the others models in terms of statistical indicators: nRMSE (0.033) and R^2 (97.7%).

Keywords

Global solar radiation, Prediction, Multi-layer perceptron, Boosted decision tree, Artificial neural networks.

Nomenclature

ANN	Artificial Neural Networks
DGSR	Daily global solar irradiation
H_0	extraterrestrial global solar radiation
S_0	sunshine duration
T_{max}	max air temperature (°C)
T_{mean}	mean air temperature (°C)
T_{min}	min air temperature (°C)
a–b	regression coefficients
MAE	Mean absolute error
nMAE	normalised Mean absolute error
RMSE	root mean square error
nRMSE	normalised root mean square error
rRMSE	relative root mean square error
R^2	Coefficient of determination
RBF	Radial basis function
MLP	Multi-layer perceptron
BDT	Boosted decision tree
LR	Linear regression

1. Introduction

Energy plays the most vital role in the economic growth, progress, and development, as well as poverty eradication and security of any nation. Uninterrupted energy supply is a vital issue for all countries today. Future economic growth crucially depends on the long-term availability of energy from sources that are affordable, accessible, and environmentally friendly.

Solar energy is a clean, affordable and environmentally friendly. As it reduces the need for other energy that drains the earth's resources, such as oil and coal. Using solar energy instead of other fuels, we maintain these limited resources and also stop the destruction of forests, oil drilling and other harmful practices.

An optimal use of solar energy needs an accurate knowledge of solar radiation at a particular geographical location. Nevertheless, these data are not always dispensable particularly in isolated areas. In this respect several approaches have been developed in the literature for modeling and predicting solar radiation components such as: empirical models and soft computing using artificial neural network [1].

The empirical models can be mainly classified into four following categories based on the employed meteorological parameters:

(1) Sunshine-based models.

(2) Cloud-based models.

(3) Temperature-based models.

(4) Other meteorological parameter-based models

One of the most widely used correlation for estimating daily global solar radiation was proposed by Angstrom –Prescott [2], who derived a linear relationship between the ratio of average daily global radiation to the corresponding value on a completely clear day (H/H_C) at a given location and the ratio of average daily sunshine duration to the maximum possible sunshine duration according to equation 1:

$$\frac{H}{H_0} = a + b\left(\frac{S}{S_0}\right) \quad (1)$$

Despite its wide use, the empirical model has not met the desirable accuracy required from a reliable prediction model. It does not wholly capture nonlinear characteristics exhibited by solar radiation [3].

In the recent years, soft computing techniques (neural network, fuzzy logic, genetic programming Support Vector Machine, Artificial Neuro-fuzzy inference system and Extreme Learning Machine (ELM)) and particularly Artificial Neural Networks (ANN) are widely applied for solar radiation prediction [4].

Guermoui *et al* [1], used multi layer neural network for global solar radiation prediction in a semi arid climate using three years of measurements. The obtained results show that MLP-model based on sunshine duration and mean air temperature gives high results in term of Mean Absolute Bias Error, Root Mean Square Error, Relative Square Error and Correlation Coefficient. The obtained values of these indicators are 0.67 MJ/m², 1.28 MJ/m², 6.12%and 98.18%.

Rabehi *et al* [5], proposed a simple model based on Radial Basis Function neural network for global solar radiation estimation in a semi arid climate. In their study, they used a limited set of meteorological data as inputs for the developed RBF model. The obtained results show that the Radial Basis Function neural network model can predicts the daily global solar radiation of clear and perturbed days with a good accuracy.

Benkaciali *et al* [6], applied multilayer perceptron for solar radiation assessment. Many experiments were carried out using different weather inputs parameters. The obtained results between the measured and the predicted data generated by the proposed model (MLP) in terms of normalized mean bias error (nMBE), normalized root mean square error (nRMSE) , and coefficient of determination (R^2) are about 0.048%, 4.422%, and 98.00%.

In [8], the authors proposed many SVM-models with different inputs parameters for solar radiation forecasting. The obtained results shows that SVM-models based only on combinations of temperatures in their inputs give lower precision (NRMSE=20.095%, R=74%) compared to those based on temperature, extraterrestrial global solar radiation and sunshine duration (NRMSE=13.163%, R=89.6%).

Banghanem *et al* [8] used Radial Basis Functions (RBF) neural network to estimate DGSR at Saudi Arabia. Four RBF models were developed using different meteorological data as input, it was found that the RBF model that uses sunshine duration and air temperature as input gives high performance (r = 98.8%).

Mellit *et al* [9] proposed a simplified hybrid model for generating sequences of DGSR, which combines a neural network and Markov chain. This model is called ANN-MTM (Markov Transition Matrix). The input parameters of the proposed model are the geographical coordinates while the outputs are the DGSR. The obtained results between the measured and the predicted data generated by the proposed model in term of root mean square error (RMSE) is below 8% and the correlation coefficient between 90 and 92%.

Behrang *et al* [10] used Radial Basis Functions (RBF) and Multi-Layer Perceptron (MLP) networks based on meteorological parameters such as daily mean air temperature, relative humidity, sunshine hours, evaporation, and wind speed. It was found that the MLP approach with the input parameters wind speed, daily mean air temperature, and relative humidity had the better results than the other proposed models. The mean absolute percentage error (MAPE) for MLP model is 5.21%.

Sözen *et al* [11-12] used meteorological and geographical data as input variables in the ANN model for solar radiation estimation in Turkey. With MAPE of 6.73% for MLP network.

Linares-Rodriguez *et al* [13] used the MLP model for estimating solar radiation over Spain from satellite-obtained irradiances. The input layer has 12 inputs (11 Meteosat channels and clear sky solar radiation). The RMSE is 6.74%. The model performs well in cloudy and clear sky condition.

Moreover, Using Multilayer Perceptron (MLP) neural network and considering different inputs, Asl et al. have predicted the amount of daily global solar radiation of Dezful city in Iran with Mean Absolute Percentage Error of 6.08%. [14]. we have summarized in the Table 1 the results of some published works where DGSR prediction models have been developed using The Multi-Layer Perceptron (MLP) networks.

Mellit *et al* [15] have developed a prediction model of daily global solar radiation from the mean sunshine duration and air temperature using an adaptive neuro-fuzzy inference scheme (ANFIS).

Gani *et al* [16] used nonlinear autoregressive (NAR) neural network model to predict the daily global radiation received on a horizontal surface for seven cities in Iran, employing day of the year as unique input to the network. The validated model was compared with the adaptive neuro-fuzzy inference scheme (ANFIS) and the achieved result shows lower values of statistical performances of NAR model against ANFIS.

Gairaa *et al* [17] used a new combined ARMA–ANN approach for estimating daily global solar radiation. It was found an improvement of the combined model over ARMA and ANN models in term of mean absolute error (MPE) of about 18.1% and 2.7%.

Al-Hajj *et al* [18], applied genetic programming approach for solar irradiance estimation. The proposed method was validated on dataset (2004-2007) provided by the National Center of meteorology and Seismology (NCMS) in Abu-Dhabi—UAE. The meteorological data used are: air temperature, wind speed, relative humidity, and sunshine duration. The achieved results in term of root mean square error RMSE and Mean Bias Error MBE are: 0.21 and 0.052, respectively.

Another interesting work proposed by Al-Hajj et al [19]. They proposed to use MLP and RBF model for global horizontal irradiation estimation over three cities in UAE (United Arab Emirates), namely Abu Dhabi, Dubai and Al-Ain. They proved that MLP model achieve high performance compared to the RBF model using maximum temperature, wind speed, sunshine hours, and relative humidity as input parameters. The obtained results in terms of mean bias error (MBE), Root Mean Square Error (RMSE),Correlation coefficient (R^2) ,Nash-Sutcliffe model Efficiency coefficient ,(NSE) t-statistic (T_{stat}) using MLP models are:0.096,0.421,0.98,0.81 and 0.776 ,respectively for Abu Dhabi and -0.13,0.383,0.98,0.86 and 1.202 for Al Ain region and for Dubai city the achieved results are 0.153,0.414,0.99,0.82 and 1.314 ,respectively.

In [20], the others proposed different mathematical formula for global solar radiation using sunshine duration as input from tow weather station in the United Arab Emirates (UAE). The obtained results showed that all mathematical models perform well in both cities of Abu Dhabi and Al Ain,with slight improvement of Linear Angstron-Perscot model for estimating the average global solar radiation for Abu Dhabi city (MAPE=1.89% and R^2=94%).

In [21], the team of researchers at UAE proposed short term forecast weather models of 10 years of measurements. They employed various models such as classical models, artificial neural network (ANN) models and time series regression models. The models was validated on three years (2005-2007), yielding deterministic coefficient of 92.6% and 99.98% for mean daily and monthly data, respectively.

Maitha et al [22], proposed to use ANN model for solar radiation assessment on 13 year of measurements. The meteorological data between 1995 and 2004 are used for training the ANN and data between 2004 and 2007 are used for testing the predicted values. The achieved values of RMSE, MBE, MAPE, and R^2 are found to be, respectively, 35%, 0.307%, 3.88%, and 92%.

Table 1 summarizes the recent works done in this field using MLP model and the accomplished performances.

Please Insert Table 1.

As mentioned before, this paper presents a comparative study between different models namely, Multi-layer perceptron neural network (MLP), Boosted decision tree (BDT), and Linear regression model. A case of hybrid model was also studied (LR-MLP and LR-BDT), to assess the performance of these model for solar radiation application. The experimental dataset includes the extraterrestrial solar radiation, daily minimum and maximum and mean temperatures, and sunshine duration ratio.

The remainder of the paper is organized as follows. Section 2, describes site location and data collection. In Section 3, we present the theory of LR, MLP and BDT and the proposed hybrid models (LR-MLP and LR-BDT). Sections 4 represent model validation. Results and discussion are given in Section 5. Finally, Section 6 concludes the paper and suggests future work.

2. Material and Methods

2.1 Site and Dataset

Algeria is the largest country in Africa with a total area of 2, 381.741km^2 where the desert comprises 86% of the total area of the country. Algeria holds one of the highest solar potentials in the world which is estimated at 13.9 TWh per year. The country receives annual sunshine exposure equivalent to 2,500 KWh/m^2. Daily solar energy potential varies from 4.66 kWh/m^2 in the north to 7.26 kWh/m^2 in the south.

The experimental dataset used in this work is available by the Applied Research Unit for Renewable Energies, (URAER) situated in the south of Algeria about 20 km from Ghardaïa city with (latitude = +32.37, longitude = +3.77 and altitude = 450 m), Figure 1 shows the site location. The database contains DGSR, extraterrestrial global solar radiation, air (mean, min and max) temperature and sunshine duration. The data are recorded every 5 min with a high precision by a radiometric station installed at the rooftop of the URAER, as shown on Figure 2.

Figs. 3 and 4 shows the evolution of daily global solar irradiation, extraterrestrial global solar radiation and sunshine duration received along three years (2014, 2015 and 2016) on horizontal surface.

Fig. 1. Location of the studied site.

Fig. 2. Instrumentation station.

Fig. 3. Daily evolution of global solar irradiation and extraterrestrial global solar radiation at Ghardaïa city, Algeria.

Fig. 4. Daily evolution of sunshine duration at Ghardaïa city, Algeria.

Daily solar radiation data are useful for evaluating the performance of solar electric and solar thermal systems. Figure 5 presents the minimum, maximum and average DGSR. It can be seen that daily global solar irradiation varied between a minimum of 580.47 Wh/m^2 in December to a maximum of 9073.40 Wh/m^2 in June. The average GSR is about 5852,77 Wh/m^2. The maximum GSR observed in July is mainly due to the

high solar altitude and long day length in summer. The below average GSR recorded from December to April are due mainly to the low solar altitude in winter and unstable weather conditions in spring.

We can also note that the maximum and average values of solar radiation for the three years are very close, which is the most important feature of this site. For the minimum monthly values, we note that their distribution is very broad and not close, possibly due to several factors, such as the appearance of sudden sand storms.

Fig. 5. Maximum, mean and minimum of monthly DGSR (2014, 2015 and 2016).

3. Prediction Model

3.1 Linear regression

Linear regression model **(**LR) is a simple method which intend to finds the relation between the dependent variable and the independent variable [34]. This method depend on the assumption of the relationship of this variables are polynomial as in Eq. 1.

$$y_i = \beta_1 x_{i_1} + \cdots + \beta_p x_{i_p} + \varepsilon_i \quad (1)$$

Where, equation 1 can also be shown as a format of Eq. 2.

$$y = X\beta + \varepsilon \quad (2)$$

Where y is dependent variable, X is independent variable, β is parameter vector and ε is error term shown below.

$$y = \begin{bmatrix} y_1 \\ y_1 \\ \vdots \\ y_p \end{bmatrix}, X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & .. & x_{np} \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix} \quad (3)$$

3.2 Non-linear process:

- **Multi-layer perceptron (MLP)**

Neural networks can be classified into dynamic (e.g. Elman) and static (e.g. MLP) categories. MLP is the most commonly used static networks [35], in which the input is presented to the network along with the desired output, and the weights are adjusted so that the network attempts to produce the desired output. The MLP network has three layers of neurons (nodes) an input layer, a hidden layer and an output layer.

The Neurons are the fundamental elements in each layer, and every neuron in one layer is associated and interacts with other layers. The outputs of every neuron in the hidden and output layers are determined by the previous output ($\Sigma\, w_{ij}\, x_j$, x_j is the input signals), activation function ($f(\Sigma\, w_{ij}\, x_j)$), and weighting coefficients (w_{ij}) [36]. The MLP network is processed as follows: the information flow is input into the

input layer and passes through the hidden and output layers to achieve the output information. The Tan-sigmoid activation function is used in the neurons of hidden and output layers in this work. In detail, neurons sum the weight-controlled input production to apply the nonlinear activation function as follows:

$$a_i = \sum_{j=1}^{n} w_{ij}\, x_j \quad (4)$$

$$y_i = f(a_i) = \frac{2}{1+e^{-2a_i}} - 1 \quad (5)$$

Where w_{ij} is the weight coefficient, x_j is the input signals, a_i is the summation of the weighted inputs, and y_i is the output at neuron i of the output layer.

- **Boosted decision tree (BDT)**

In the Boost Decision Tree (BDT) method, decision trees are created using a classification and regression tree method. This method can be used to predict the quantitative consequences (regression tree) or classified (classification tree). Using a set of logical conditions, if then rather than a linear relationship, the regression tree method predicts or classifies the result under study [37]. By combining the capabilities of two algorithms, including regression trees. BDT models describe response to predictors through optimized and optimized binary separation and use an adaptive method to combine a large number of simple models to obtain adequate performance. Among the advantages of this method is a lower sensitivity to excessive adjustment [38], a high-speed analysis of high-volume data, and a lack of statistical assumptions regarding the distribution of data. The improved performance of this algorithm is obtained by adjusting the options related to the reinforced shafts and the shrinkage rate. The selection of the best shrinkage rate avoids exceeding the forecasts. Empirical studies have shown that a shrinkage rate of 1.0 or less usually generates appropriate models [39].

3.3 Combined models

We assume that the daily solar radiation series is composed by linear and nonlinear components [40]. A general model describing a particular time series, thus, can be formulated as the following:

$$o_t = Linear(t) + Nonlinear(t) + e_t \quad (6)$$

Where, o_t denotes the time series observation at time *t*, *Linear (t)* and *Nonlinear (t)* represent the linear and nonlinear components of the time series, respectively, and e_t denotes the random error term.

In order to yield a more general and more accurate hybrid linear and nonlinear model than the artificial neural network model, generalized hybrid model has been proposed. The process of building the model consists of three phases comprises the following sequential steps:

Step 1:

Creating a linear regression model based on raw data as a linear component. We can then calculate the model residual using of the following relation:

$$r_t = o_t - \widehat{Linear}\,(t) \quad (7)$$

Since daily solar radiation data have nonlinear pattern, these means the residual series should contain nonlinear component only. That needs to be captured by nonlinear models.

Step 2:

In second stage, the main aim is nonlinear modeling; therefore, a multilayer perceptron (MLP) and Boosted decision tree (BDT) are used in order to model the nonlinear that may be remained in residuals of linear modeling.

Step 3:

The MLP and BDR models developed in Phase 2 which accounts for the nonlinear residuals will then be combined with the linear model developed in step 1 for produce the final model. The hybrid model can, therefore, be represented as follows:

$$y_t = \widehat{Linear}\,(t) + \widehat{Nonlinear}\,(t) \quad (8)$$

The flowchart of the combined model is summarized in Fig. 6.

Fig. 6. Combined approach flowchart.

4 Statistical evaluation

In this work the performance of the proposed models is assessed by comparing the measured values with the estimated values using different statistical indicators such coefficient of correlation R^2, root mean square error (RMSE) and its normalised value (nMBE), relative root mean square error (rRMSE), mean absolute error (MAE) and its normalised value (nMAE). These error indices are defined as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right)^2} \quad (9)$$

$$\mathrm{rRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right)^2}}{\frac{1}{N}\sum_{i=1}^{n} x_i} \times 100 \quad (10)$$

The ranges of rRMSE define the model performance as:

Excellent if: rRMSE < 10%

Good if: 10 % < rRMSE < 20%

Fair if: 12 % < rRMSE < 30%

Poor if: rRMSE > 30%

$$MAE = \frac{1}{N}\sum_{i=1}^{n} |y_i - x_i| \quad (11)$$

$$R^2 = \frac{\sum_{i=1}^{n} (y_i - x_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2} \quad (12)$$

Where: y_i and x_i are the estimated and the measured values, $\bar{y}_i$ is the average of estimated values and N is data number.

5 Results and discussion

As mentioned before, the main objective of this study is to assess the performance of different models presented in section 3 for solar radiation prediction .The proposed models were now used to predict the global solar radiation for Ghardaia region, using the available meteorological parameters (e.g. air temperature, sunshine ratio, extraterrestrial global solar radiation). In this context, several models are developed based on different inputs parameters. In the current study, the inputs data are divided mainly into four sets: {(T), (T, H), (T, S_0), (T, H, S_0)} to select the most relevant inputs data for global solar radiation prediction (See Table 2 and Table 3).

The collected data are normalize to [-1, 1] then divided into two subsets the first part (70% values) are used for training the sets models, while the remaining (30% values) are used for testing the performance of the studied models.

The achieved statistical results for MLP and BDT regression is reported in Tables 2 and 3, respectively.

Table 2.The studied MLP models with different input attributes

From table .2 , it can be noted that, the first group of models based only on temperatures as input parameter (MLP_1...MLP_7) gives low performance compared to those based on sunshine duration (MLP_{15}...MLP_{21}). We can say that the calculated maximum sunshine duration improves significantly the performances of the models, since the relationship between DGSR and S_0 is relatively good.

Another important remark, is that the performance of the model does not improve significantly when combining sunshine duration and air temperature with the extraterrestrial solar radiation,

The best prediction output (model MLP_{27}) is shown in the figure .7. We can see that the gap between the measured data and the predicted values are very low, except for a few days where the solar radiation is strong, probably due to a few gusts of sand wind.

For the second sets, using the BDT regression model, as can be seen from table 2. that BDT_{26} model with combination {T_{mean}, T_{min}, H_0, S_0} achieve good results compared to other combinations (nRMSE =0.0415, R^2= 0.96). Figure 10. Show the comparison between the predicted and the measured values.

Fig 7. Comparisons (a) and correlations (b) between measured and predicted global solar irradiation for MLP model.

Table 3.The studied BDT models with different input attributes.

Fig 8. Comparisons (a) and correlations (b) between measured and predicted global solar irradiation for BDL model.

Lastly, for the combined model we have merged the linear regression with the MLP and BDL models, Figures (9.10) shows the comparison between the predicted and the measured values. The summary for all three models is shown in Table 3.

Fig 9. Comparisons (a) and correlations (b) between measured and predicted global solar irradiation for LR-MLP model.

Table .4 Results summary

Table 4, represent the results obtained by using hybrid and single models (MLP, BDT). As can be seen from table 3, MLP model achieve the highest performance compared with others models in terms of nRMSE (0.033) and R^2 (97.7%). Another important remark is that combined the MLP model with linear regression don't boost the performance of the predictive values.

For the third case, the combination of the BDT model with linear regression boost the performance with R^2=1% and REDUCE the nRMSE by 1.3 compared to the BDT model.

Fig 10. Comparisons (a) and correlations (b) between measured and predicted global solar irradiation for LR-BDL model.

6 Conclusion

In this work, three models (MLP, BDT, and LR) were applied for solar radiation components prediction for semi arid climate. In this study, different inputs data was used for estimating the daily solar radiation (extraterrestrial solar radiation, daily minimum and maximum and mean temperatures, and sunshine duration ratio). The experimental results showed that MLP model achieve high precision accuracy compared to the LR,BDT and hybrid LR-MLP,LR-BDT model. This has been validated using data measured in weather station located in Ghardaia city, Algeria.

The attained results in this study proved that MLP model is a suitable approach for solar radiation components prediction. Accordingly, it can be used also in regions that have a similar climate. Furthermore, other cases with different climate conditions will be considered in the future investigations

References

1. Guermoui M., Rabehi A., Benkaciali S., & Djafer D.,. Daily global solar radiation modelling using multi-layer perceptron neural networks in semi-arid region. Leonardo Electronic Journal of Practices and Technologies, 2016; 28: 35-46.
2. Angstrom A., Solar and terrestrial radiation. Quarterly Journal of Royal ¨ Meteorological Society 1924; 5 0:121–5.
3. Mubiru J., Predicting total solar irradiation values using artificial neural networks, Renewable Energy ,2008) ;33:2329–2332.
4. Kalogirou SA. Applications of artificial neural-networks for energy systems. Appl Energy 2000;67:17–35.
5. RABEHI, Abdelaziz, et al. "Radial basis function neural networks model to estimate global solar radiation in semi-arid area." Leonardo Electronic Journal of Practices and Technologies 27 (2015): 177-184.
6. Benkaciali, S., et al. "Evaluation of the global solar irradiation from the artificial neural network technique." Revue des Energies Renouvelables 19.4 (2016): 617-631.
7. Belaid, S., and A. Mellit. "Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate." Energy Conversion and Management 118 (2016): 105-118.
8. Benghanem M., Mellit A., Radial basis function network-based prediction of global solar radiation data: application for sizing of a stand-alone photovoltaic system at Al-Madinah, Saudi Arabia, Energy, 2010; 35(9): p. 3751-3762.
9. Mellit A, Benghanem M, Hadj Arab A, Guessoum A. A simplified model for generating sequences of global radiation data for isolated sites: using artificial neural network and a library of Markov transition Matrices. Solar Energy, 2005;79:468e82.
10. Behrang MA, Assareh E, Ghanbarzadeh A, Noghrehabadi AR. The potential of different artificial neural network (ANN) techniques in daily global solar radiation modelling based on meteorological data. Solar Energy 2010;84:1468e80.
11. Sözen A, Arcaklioğlu E, Özalp M, Kanit EG. Use of artificial neural net works for mapping of solar potential in Turkey. Appl Energy 2004;77:273–86.
12. Sözen A, Arcaklioğlu E, Özalp M. Estimation of solar potential in Turkey by artificial neural net works using meteorological and geographical data. Energy Convers Manag 2004;45:3033–52.
13. Linares-Rodriguez A, Ruiz-Arias JA, Pozo-Vazquez D, Tovar-Pescador J. An artificial neural net work ensemble model for estimating global solar radiation from Meteosat satellite images. Energy 2013; 61: 636–45.
14. Asl SFZ, Karami A, Ashari G, Assareh A, Hedayat N. Daily global solar radiation modelling using multi-layer perceptron (MLP) neural networks, World Academy of Science, 2011, vol. 79, p. 740-742.
15. Mellit A, Hadj Arab A, Khorissi N, Salhi H. An ANFIS-based forecasting for solar radiation data from sunshine duration and ambient temperature. Power Engineering Society General Meeting. IEEE, 2007. p. 1-6.
16. Gani, A., Mohammadi, K., Shamshirband, S., Khorasanizadeh, H., Danesh, A. S., Piri, J., ... & Zamani, M. Day of the year-based prediction of horizontal global solar radiation by a neural network auto-regressive model. Theoretical and applied climatology, 2016: 125(3-4); 679-689.
17. Gairaa, K., Khellaf, A., Messlem, Y., & Chellali, F. Estimation of the daily global solar radiation based on Box–Jenkins and ANN models: A combined approach. Renewable and Sustainable Energy Reviews, 2016: 57; 238-249.

18. Al-Hajj, Rami, and Ali Assi. "Estimating solar irradiance using genetic programming technique and meteorological records." (2017).
19. Hejase, Hassan AN, Maitha H. Al-Shamisi, and Ali H. Assi. "Modeling of global horizontal irradiance in the United Arab Emirates with artificial neural networks." Energy 77 (2014): 542-552.
20. Assi, Ali, Mohammed Jama, and Maitha Al-Shamisi. "Prediction of global solar radiation in abu dhabi." ISRN Renewable Energy2012 (2012).
21. Hejase, Hassan AN, and Ali H. Assi. "Time-series regression model for prediction of mean daily global solar radiation in Al-Ain, UAE." ISRN Renewable Energy 2012 (2012).
22. Al-Shamisi, Maitha H., Ali H. Assi, and Hassan AN Hejase. "Artificial neural networks for predicting global solar radiation in al ain city-uae." International journal of green energy 10.5 (2013): 443-456.
23. Linares-Rodríguez A, Ruiz-Arias JA, Pozo-Vázquez D, Tovar-Pescador J. Generation of synthetic daily global solar radiation data base don ERA-Interim reanalysis and artificial neural networks. Energy 2011; 36: 5356–65.
24. Koca A, Oztop HF, Varol Y, Koca GO. Estimation of solar radiation using artificial neural networks with different input parameters for Mediterranean region of Anatolia in Turkey. Expert Syst Appl 2011; 38:8756–62.
25. Khatib T , Mohamed A, Sopian K, Mahmoud M. Solar energy prediction for Malaysia using artificial neural networks. Int J Photoenergy 2012:1–16.
26. Alam S, Kaushik SC, Garg SN. Computation of beam solar radiation at normal incidence using artificial neural network. Renew Energy 2006; 31:1483–91.
27. Rumbayan M, Abudureyimu A, Nagasaka K. Mapping of solar energy potential in Indonesia using artificial neural network and geographical information system. Renew Sustain Energy Rev2012; 16:1437–49.
28. Behrang MA, Assareh E, Ghanbarzadeh A, Noghrehabadi AR. The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. Sol Energy 2010; 84:1468–80.
29. Moghaddamnia A, Remesan R, Kashani MH, Mohammadi M, Han D, Piri J. Comparison of LLR, MLP, Elman, NNARX and ANFIS Models—with a case study in solar radiation estimation. J Atmos Solar Terr Phys 2009; 71:975–82.
30. Ramedani Z, Omid M, Keyhani A. Modelling Solar Energy Potential in a Tehran Province Using Artificial Neural Networks, International Journal of Green Energy, 2013;10:427-441.
31. Kumar, S., & Kaur, T. Development of ANN Based Model for Solar Potential Assessment Using Various Meteorological Parameters. Energy Procedia; 2016: 90, 587-592.
32. Al-Alawi, S. M., & Al-Hinai, H. A. An ANN-based approach for predicting global radiation in locations with no direct measurement instrumentation. Renewable Energy; 1998:14(1-4), 199-204.
33. Mohandes, M., Rehman, S., & Halawani, T. O. Estimation of global solar radiation using artificial neural networks. Renewable Energy;1998:14(1-4), 179-184.
34. Almorox J Hontoria C, Global Solar Radiation estimating using sunshine duration in Spain 2004; 45 pp: 1529-1535.
35. Moghaddamnia, A., Remesan, R., Kashani, M. H., Mohammadi, M., Han, D., & Piri, J. Comparison of LLR,

MLP, Elman, NNARX and ANFIS Models—with a case study in solar radiation estimation. Journal of Atmospheric and Solar-Terrestrial Physics; 2009: 71(8), 975-982.

36. Hassanpour Kashani, M., Flood estimation at ungauged sites using a new hybrid model. Journal of Applied Sciences; 2008: 9, 1744‑1749.
37. Westreich, D., Lessler, J., & Funk, M. J. Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. Journal of clinical epidemiology; 2010:63(8), 826-833.
38. Zounemat-Kermani, M., Rajaee, T., Ramezani-Charmahineh, A., & Adamowski, J. F. Estimating the aeration coefficient and air demand in bottom outlet conduits of dams using GEP and decision tree methods. Flow Measurement and Instrumentation; 2017: 54, 9-19.
39. Shataeea, S., Weinaker, H., & Babanejad, M. Plot-level forest volume estimation using airborne laser scanner and TM data, comparison of boosting and random forest tree regression algorithms. Procedia Environmental Sciences; 2011: 7, 68-73.
40. Zheng, G.P., Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing; 2003: 50, 159–175.

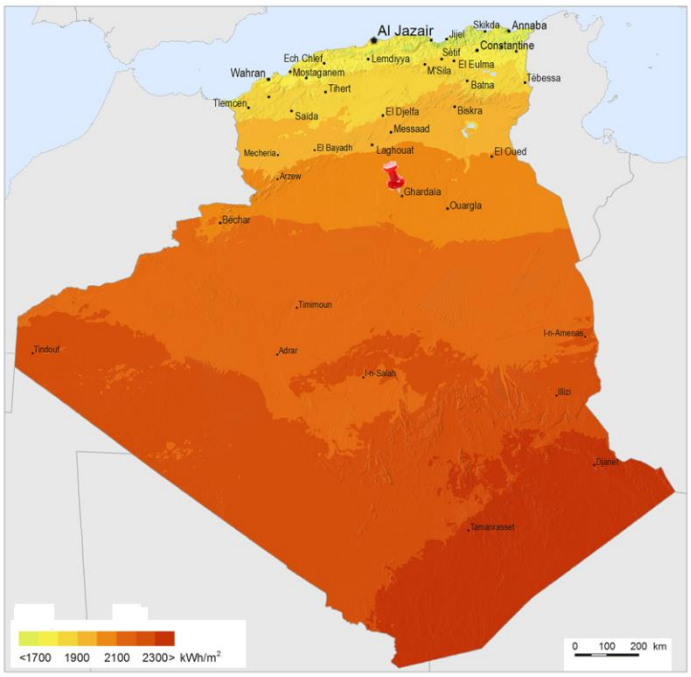


Fig. 1. Location of the studied site.

Fig. 2. Instrumentation station.

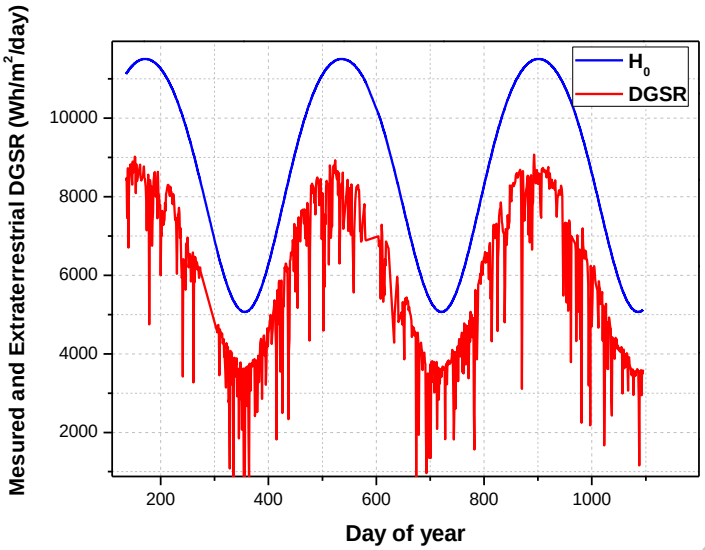


Fig. 3. Daily evolution of global solar irradiation and extraterrestrial global solar radiation at Ghardaïa city, Algeria.

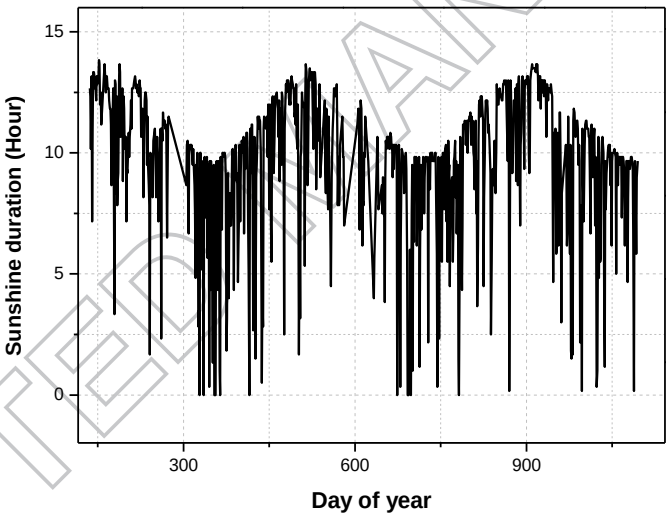


Fig. 4. Daily evolution of sunshine duration at Ghardaïa city, Algeria.

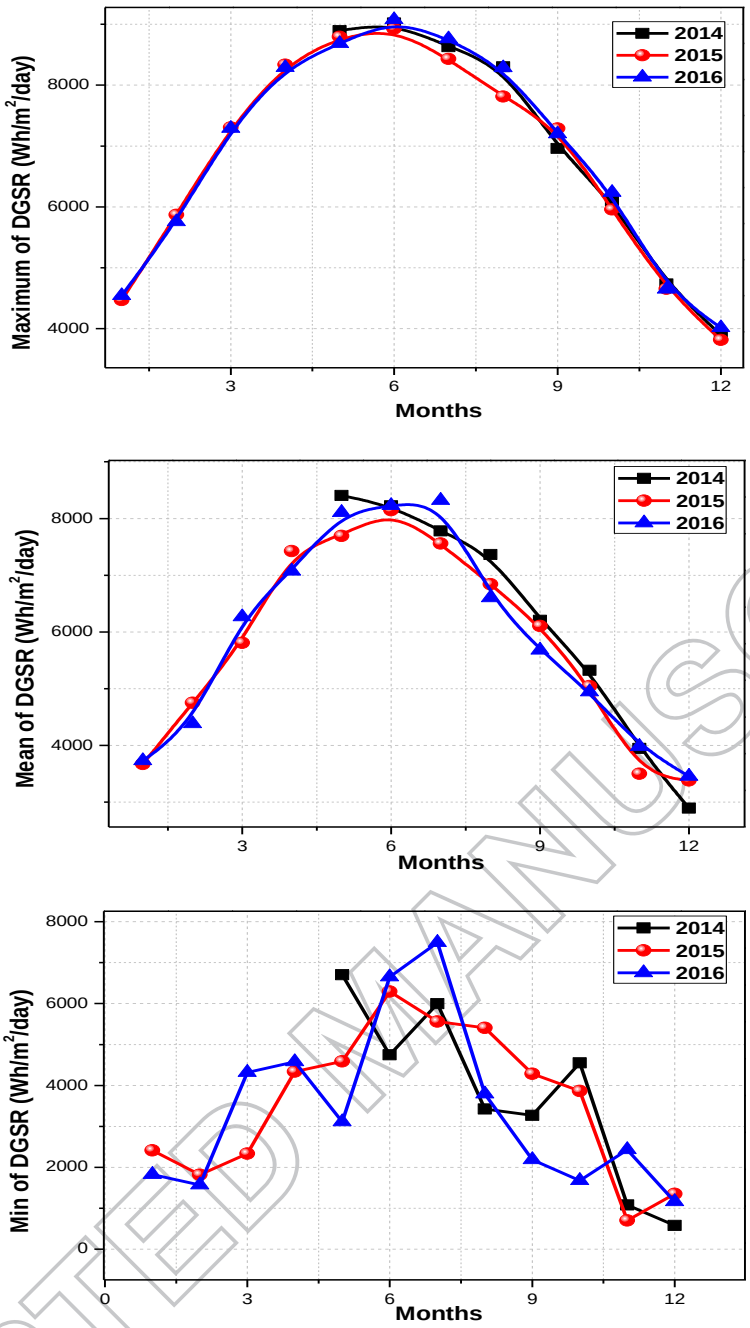
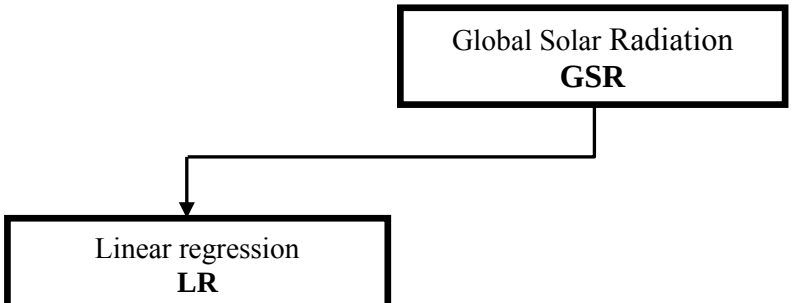
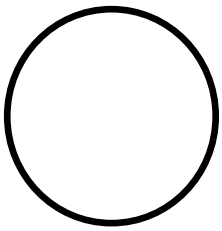


Fig. 5. Maximum, mean and minimum of DGSR for each month (2014, 2015 and 2016).



$o_t - \widehat{Linear}(t)$



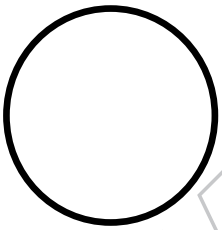
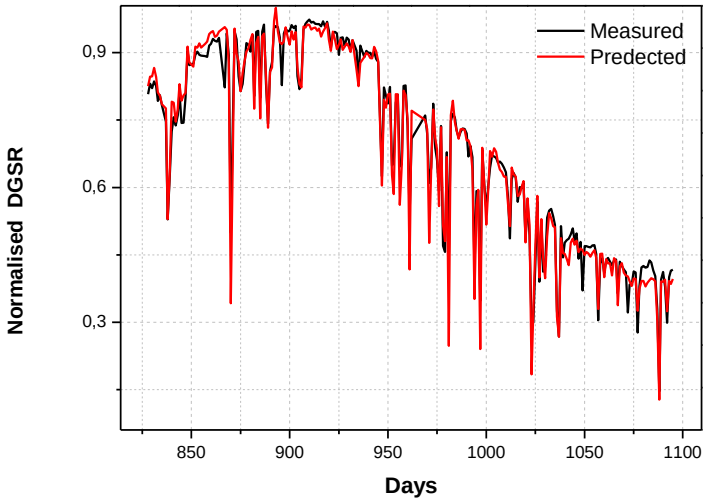
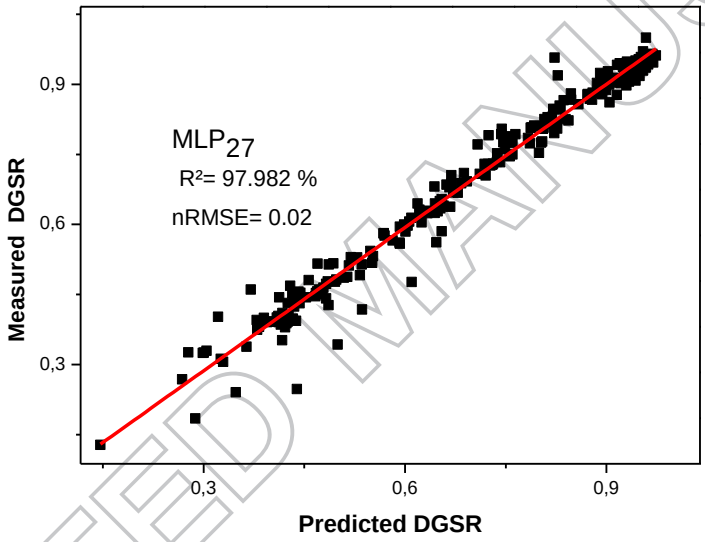


Fig. 6. Combined approach flowchart.

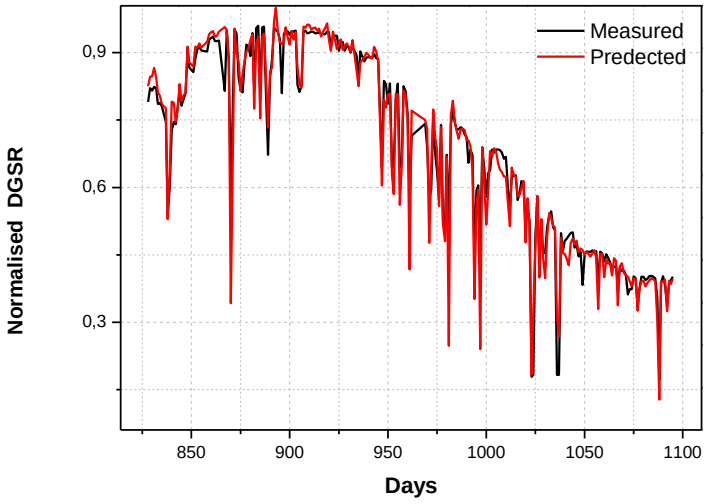


(a)

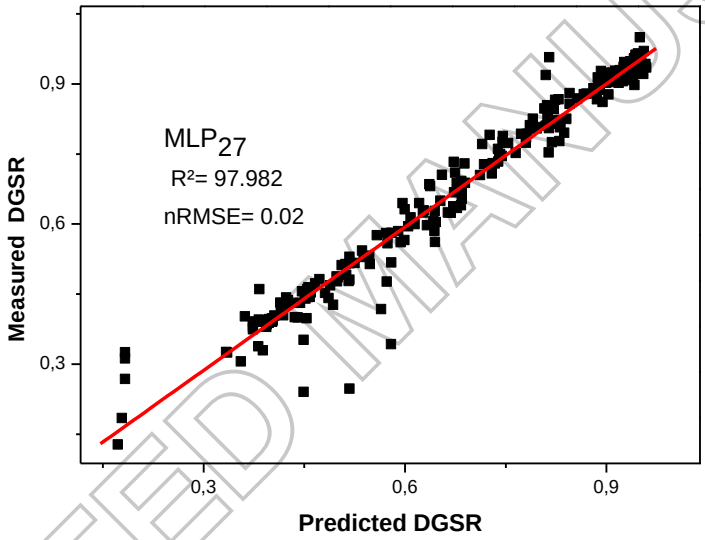


(b)

Fig 7. Comparisons (a) and correlations (b) between measured and predicted global solar irradiation for MLP model.

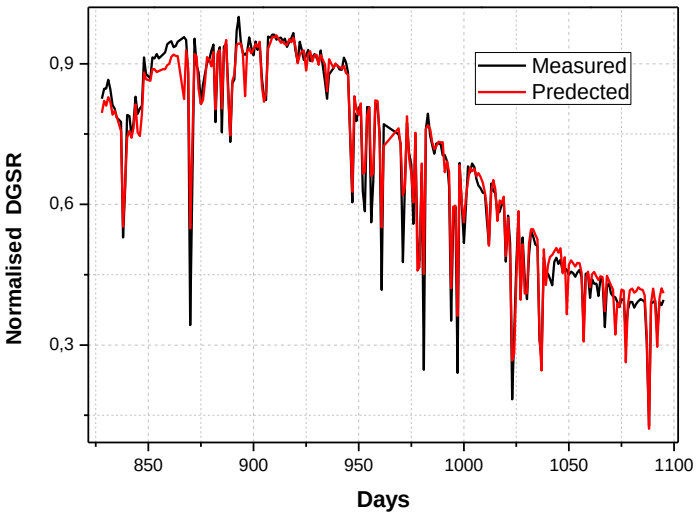


(a)

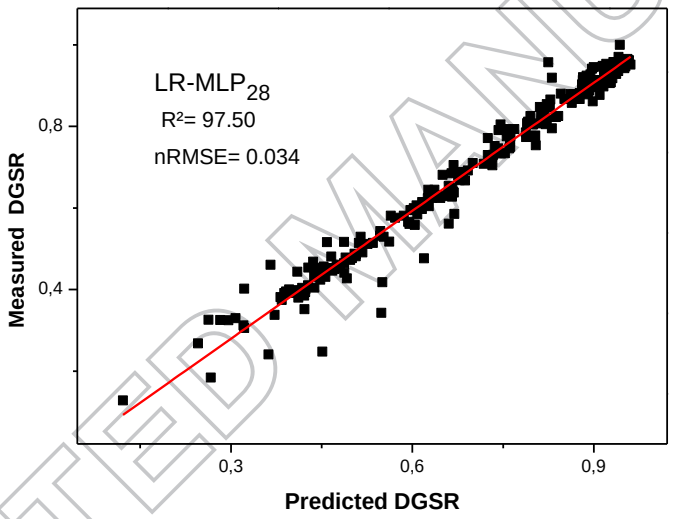


(b)

Fig 8. Comparisons (a) and correlations (b) between measured and predicted global solar irradiation for BDT model.



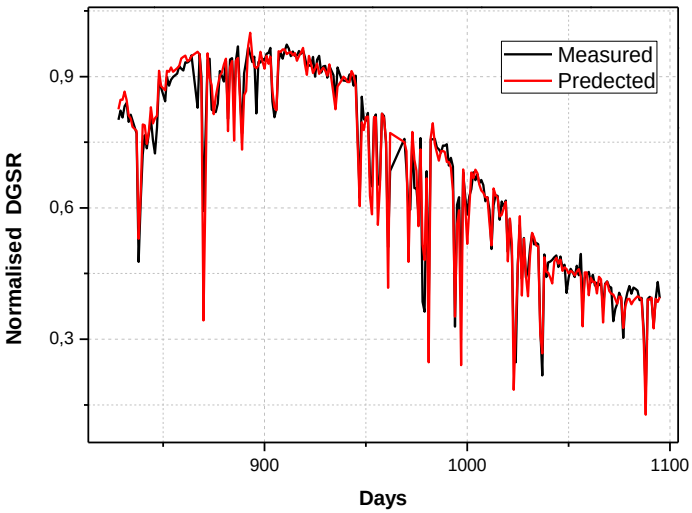
(a)



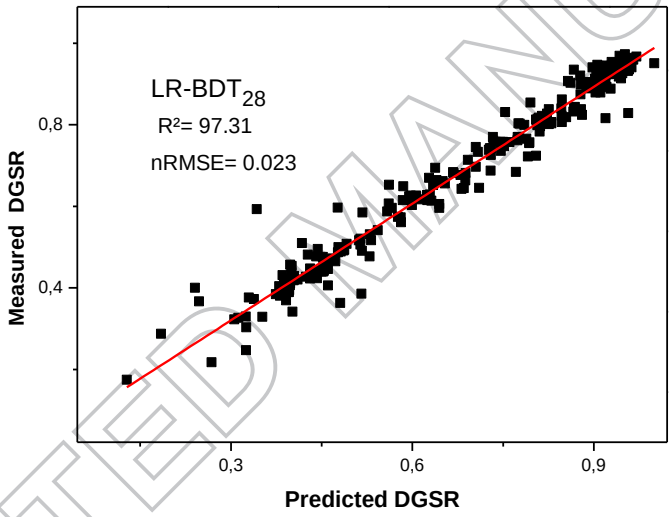
(b)

Fig 9. Comparisons (a) and correlations (b) between measured and predicted global solar irradiation for LR-MLP model.

.



(a)



(b)

Fig 10. Comparisons (a) and correlations (b) between measured and predicted global solar irradiation for LR-BDT model.

Table 1 Some DGSR prediction models.

Reference	Input variables to MLP Model	MLP Model prediction accuracy	Location
[23]	Latitude, longitude, day of the year, daily clear sky global radiation, cloud cover, total column ozone and water vapor	RMSE is 14.20%	Spain
[24]	Latitude, longitude, altitude, months, average temperature, average cloudiness, average wind velocity and sunshine duration	RMSE is 6.9%	Anatolia, Turkey
[25]	Latitude, longitude, day's number and sunshine ratio	MAPE is 7.96%	Malaysia
[26]	Latitude, longitude, altitude, month of the year, mean duration of sunshine per hour, rainfall ratio, relative humidity	RMSE varies from 1.65 to 2.79%	India
[27]	Month, latitude, wind speed, precipitation, sunshine duration, humidity and temperature	MAPE is 3.4%	Indonesia
[28]	day of the year, minimum air temperature, sunshine duration, relative humidity, vapor pressure, wind speed	MAPE varies from 5.21–12.86 %	Iran
[29]	maximum air temperature, mean air temperature, extraterrestrial global solar radiation, wind speed, precipitation	NRMSE is 19.10 %	Iran
[30]	maximum daily temperature, relative humidity, sunshine duration and the amount of precipitation	RMSE is 3.09	Dezful city in Iran
[31]	temperature, humidity, barometric pressure, rainfall, sun shine hours	MAPE is 16.45 %	Mumbai, India
[32]	pressure, temperature, Month, Latitude	MAPE is 5.43%	Sultanate Oman
[33]	Longitude, Altitude, Sunshine duration	MAPE is 12.61%	Saudi Arabia

Table 2. The studied MLP models with different input attributes

Model	Inputs	MAE	nRMSE	rRMSE	R^2
MLP_1	T_{max}	0.064408	0.094974	13.85167	81.0752

MLP_2	T_{min}	0.065878	0.096787	14.11609	80.3458
MLP_3	T_{mean}	0.06593	0.095704	13.95814	80.7834
MLP_4	T_{max}, T_{min}	0.065075	0.090993	13.27106	82.6287
MLP_5	**T_{mean},T_{min}**	**0.063729**	**0.089746**	13.08919	**83.1016**
MLP_6	T_{mean},T_{max}	0.064204	0.091122	13.28987	82.5793
MLP_7	T_{mean},T_{min},T_{max}	0.063489	0.091384	13.32808	82.479
MLP_8	T_{max}, H_0	0.064425	0.097994	14.29213	79.8526
MLP_9	T_{min}, H_0	0.06855	0.097933	14.28323	79.8779
MLP_{10}	T_{mean}, H_0	0.06803	0.097829	14.26807	79.9205
MLP_{11}	T_{max}, T_{min}, H_0	0.061578	0.088576	12.91854	83.5393
MLP_{12}	T_{mean}, T_{min}, H_0	0.065564	0.094248	13.74579	81.3636
MLP_{13}	T_{mean},T_{max}, H_0	0.063414	0.090207	13.15642	82.9273
MLP_{14}	**T_{mean},T_{min},T_{max}, H_0**	**0.057662**	**0.086468**	12.6111	**84.3133**
MLP_{15}	T_{max}, S_0	0.022062	0.03129	4.56355	97.9458
MLP_{16}	T_{min}, S_0	0.023062	0.032603	4.75505	97.7698
MLP_{17}	T_{mean}, S_0	0.022998	0.033326	4.8605	97.6699
MLP_{18}	T_{max}, T_{min}, S_0	0.022887	0.032119	4.68446	97.8355
MLP_{19}	T_{mean}, T_{min}, S_0	0.023066	0.032251	4.70371	97.8177
MLP_{20}	**T_{mean},T_{max}, S_0**	**0.022453**	**0.031773**	4.634	**97.882**
MLP_{21}	T_{mean},T_{min},T_{max}, S_0	0.023381	0.032569	4.75009	97.7745
MLP_{22}	T_{max}, H_0, S_0	0.021356	0.033322	4.85991	97.6705
MLP_{23}	T_{min}, H_0, S_0	0.021502	0.033484	4.88354	97.6478
MLP_{24}	T_{mean}, H_0, S_0	0.021873	0.033685	4.91286	97.6193
MLP_{25}	T_{max}, T_{min}, H_0, S_0	0.020903	0.032877	4.79501	97.7322
MLP_{26}	T_{mean}, T_{min}, H_0, S_0	0.021122	0.032859	4.79239	97.7347
MLP_{27}	**T_{mean},T_{max}, H_0, S_0**	**0.020052**	**0.031014**	4.5233	**97.982**
MLP_{28}	T_{mean},T_{min},T_{max}, H_0, S_0	0.021585	0.033105	4.82827	97.7006

Table 3. The studied BDT models with different input attributes

Model	Inputs	MAE	nRMSE	rRMSE %	R^2
BDT_1	T_{max}	0.072738	0.102844	13.85167	77.8091
BDT_2	T_{min}	0.072405	0.101587	14.11609	78.3482
BDT_3	T_{mean}	0.074331	0.10571	13.95814	76.5552
BDT_4	T_{max}, T_{min}	0.071235	0.100842	13.27106	78.6644
BDT_5	T_{mean},T_{min}	0.07134	0.100641	13.08919	78.7497

BDT_6	T_{mean}, T_{max}	0.072385	0.102438	13.28987	77.984
BDT_7	**$T_{mean}, T_{min}, T_{max}$**	**0.069109**	**0.099571**	14.99949	**79.1988**
BDT_8	T_{max}, H_0	0.070106	0.102461	14.81616	77.974
BDT_9	T_{min}, H_0	0.068723	0.099554	15.41749	79.2061
BDT_{10}	T_{mean}, H_0	0.071538	0.102853	14.7075	77.8053
BDT_{11}	T_{max}, T_{min}, H_0	0.065564	0.096491	14.67819	804661
BDT_{12}	T_{mean}, T_{min}, H_0	0.065964	0.097076	14.94028	80.2284
BDT_{13}	T_{mean}, T_{max}, H_0	0.070593	0.101643	14.52213	78.3242
BDT_{14}	**$T_{mean}, T_{min}, T_{max}$, H_0**	**0.064278**	**0.095688**	14.94363	**80.7895**
BDT_{15}	T_{max}, S_0	0.027325	0.044714	14.51965	95.8053
BDT_{16}	T_{min}, S_0	0.028563	0.047329	15.0008	95.3003
BDT_{17}	T_{mean}, S_0	0.027413	0.046595	14.07292	95.445
BDT_{18}	T_{max}, T_{min}, S_0	0.027512	0.044935	14.15824	95.7637
BDT_{19}	T_{mean}, T_{min}, S_0	0.027757	0.047069	14.82433	95.3517
BDT_{20}	**T_{mean}, T_{max}, S_0**	**0.027601**	**0.044196**	13.95581	**95.9018**
BDT_{21}	$T_{mean}, T_{min}, T_{max}$, S_0	0.027684	0.044564	6.5214	95.8334
BDT_{22}	T_{max}, H_0, S_0	0.02433	0.042157	6.90279	96.2713
BDT_{23}	T_{min}, H_0, S_0	0.024247	0.041813	6.79574	96.3319
BDT_{24}	T_{mean}, H_0, S_0	0.024073	0.041801	6.55364	96.3341
BDT_{25}	T_{max}, T_{min}, H_0, S_0	0.024045	0.041741	6.86487	96.3446
BDT_{26}	**T_{mean}, T_{min}, H_0, S_0**	**0.02411**	**0.041545**	6.44585	**96.3787**
BDT_{27}	T_{mean}, T_{max}, H_0, S_0	0.024417	0.042266	6.49953	96.252
BDT_{28}	$T_{mean}, T_{min}, T_{max}$, H_0, S_0	0.02437	0.041928	6.14847	96.3116

Table .4 Results summary

Model	Inputs	MAE	nRMSE	rRMSE %	R^2 %
MLP_{28}	$T_{mean}, T_{min}, T_{max}$, H_0, S_0	0.021585	0.033105	4.828	97.70
LR- MLP_{28}	$T_{mean}, T_{min}, T_{max}$, H_0, S_0	0.02322	0.03469	5.059	97.50
BDT_{28}	$T_{mean}, T_{min}, T_{max}$, H_0, S_0	0.02437	0.041928	6.115	96.31
LR- BDT_{28}	$T_{mean}, T_{min}, T_{max}$, H_0, S_0	0.02371	0.03537	5.158	97.31