Rev. Roum. Sci. Techn.– Électrotechn. et Énerg.
Vol. **68**, *1*, pp. 7–11, Bucarest, 2022

NONLINEAR CLUSTERED ADAPTIVE-NETWORK-BASED FUZZY INFERENCE SYSTEM MODEL FOR HOURLY SOLAR IRRADIATION ESTIMATION

KHALIL BENMOUIZA [1]

Keywords: Solar irradiation; Adaptive-network-based fuzzy inference system (ANFIS); Grid partitioning; Fuzzy C-means; Subtractive clustering.

Solar energy occupies an important place among the various sources of renewable energy. A precise knowledge of the distribution of solar irradiation in a specified location is needed before any solar irradiation system installation. This paper introduces a nonlinear clustering, adaptive-network-based fuzzy inference system (ANFIS) model to estimate the hourly solar irradiation data using meteorological inputs and clustering algorithms: grid partitioning, subtractive clustering, and fuzzy c-means. Comparing these clustering algorithms is investigated to classify the inputs into clusters, which helps the solar irradiation estimation model build better. This method's advantage is understanding and simplifying the nonlinearity presented in the input's datasets. Moreover, the FCM algorithm gives the best results from comparing the testing data; the RMSE is 43.2274 W/m^2, and MSE equals 2001.34 W/m^2 with an R2 equal to 0.9893.

1. INTRODUCTION

The depletion of fossil energy resources is inevitable due to high global energy consumption. Hence, it is essential to find new sources of energy. The primary renewable energy sources are the sun, the wind, biomass, tides, the waves of the seas and oceans, and the heat from the depths of the earth. The benefits of renewable energy are many; durability, availability, and cleanliness, but the major drawback remain the very high initial cost of installing the conversion equipment [1].

Solar energy is used widely because it will always stay supplied and available globally. Indeed, the earth receives from the sun an energy equivalent to 5000 times the world's needs; it is, therefore, a strong potential and a practical solution in front of the increasingly growing challenges in the energy and environmental field. This is an urgent and common primary goal of all humanity to preserve a holy environment for future generations [2,3].

Many studies have been made on the solar radiation received on the ground to develop conversion techniques and minimize the cost of equipment for optimal exploitation. The development of solar energy applications requires long-term data on solar irradiation. The availability and quality of these data are obtained in regions equipped with radiometric stations, which measure meteorological parameters, sunshine duration, and solar irradiation. These measurements use several models to estimate the radiation on planes at different inclinations and azimuth angles.

Several models are proposed in the literature for hourly solar irradiation estimation. They use measured meteorological parameters such as temperature, humidity, wind, extraterrestrial solar irradiation, and sunshine hours [4–7].

The first main category is the physical model, which uses the radiative transfer equations [8–10]. These equations rely on complex analytical solutions and therefore require a very specific mathematical tool to simultaneously solve an extensive data set.

The second group is the parametric model [11,12], which has the same physical principles as the previous group with a simplified set of parameters. The parametric model improved the expression of transmittance for different attenuation processes in the atmosphere, which can be used to estimate the direct component of incoming solar radiation. Diffuse radiation calculated from a few approximations reduces the complexity of the scattering process. Finally, global radiation was produced by combining direct and diffuse radiation.

The third category is Statistical models [13,14], developed based on available data. The duration of insolation was the first data exploited to facilitate its availability, and the first models are based on the duration of insolation type equation of Angstrom [15]. Global radiation is the essential input for solar system designs. Models predicting global radiation at different positions use almost the same equations. The main difference lies in the estimation of diffuse radiation.

The fourth category is the artificial intelligence models [16,17], represented by the development of the artificial neural network (ANN), which provides a mechanism to estimate one or more outputs of some input data, an activation function used, and several layers.

The last category is the estimation model by processing satellite images collected by geostationary satellites [18,19].

Artificial intelligence techniques already have well-developed uses, among which we find the identification of models, classification, function approximation, and automatic control. We can only note an increasing use of this artificial intelligence for data analysis because they offer an effective alternative to more traditional techniques in many scientific fields.

Artificial neural network models have been used widely for solar irradiation estimation, which has been used to overcome the nonlinearity presented in this series [20,21]. A hybrid ANN model with fuzzy logic (FL), which is called the adaptive-network-based fuzzy inference system (ANFIS) model, is proposed to take advantage of both ANN and FL for better estimation results and to enhance precision [17].

ANFIS is an adaptive technique to solve non-linear time series problems [22]. Improved estimation of solar irradiation data has been achieved using an ANN/ANFIS hybrid. Model performance is enhanced, and robust modeling is enabled to obtain better estimation results.

Using an ANFIS-based approach, you may analyze very tiny datasets. This is the first report on non-linear regression modeling of solar irradiation using hybrid ANFIS and ANFIS. Parsimony, interpretability, and prediction are the critical advantages of non-linear models [22].

[1] Laboratory of Materials, Energetic Systems, Renewable Energies and Energy Management, Amar Teledji University of Laghouat, Algeria, Email: k.benmouiza@lagh-univ.dz

Because of this, the primary goal of this research is to determine how well various FIS generation methods for ANFIS modeling perform. To this end, we have combined the basic ANFIS model with clustering methods to enhance the estimation results. Hence, we have used grid partitioning, subtractive clustering, and the FCM algorithms for clustering purposes. ANFIS model can be applied to any data set size. However, more research must be done on applying the nonlinear regression using the ANFIS model for hourly solar irradiation estimation.

Hence, a nonlinear ANFIS model is proposed in this paper to estimate the hourly solar radiation data. Considering the power of nonlinear modeling to overcome the nonlinearity of the ample solar radiation time series data and the clustered ANFIS for estimation purposes. The use of Non-linear modeling of solar radiation time series because of their several advantages, among them the time series estimation. In addition, for simulation purposes, several inputs were selected: the temperature, the wind, the humidity, the pressure and the clear sky, and the top of the atmosphere solar irradiation time series. A comparison between each method is also viewed for clear and cloudy skies.

The rest of the paper is organized as follows; in section 2, the adopted methodology is given, and a theoretical background of the ANFIS, grid partitioning, subtractive clustering, and FCM is shown. In section 3, the data used for the simulation is illustrated. Section 4 is devoted to the simulation results and discussions; the last section is to the conclusion.

2. **METHODOLOGY**

We aim to estimate solar irradiation using the ANFIS clustering model. Hence, several inputs were selected to train the model. The data are selected and filtered then Fuzzy grid partitioning, subtractive clustering, and FCM are used for the estimation. In what follows, we will give a detailed theoretical overview of each used model.

2.1. ANFIS MODEL

ANFIS consists of neuro-fuzzy adaptative systems which use five multi-layer perception neural network models. Each layer has its functions to construct the fuzzy system. The ANFIS contains five layers [22,23], as shown in Fig. 1.

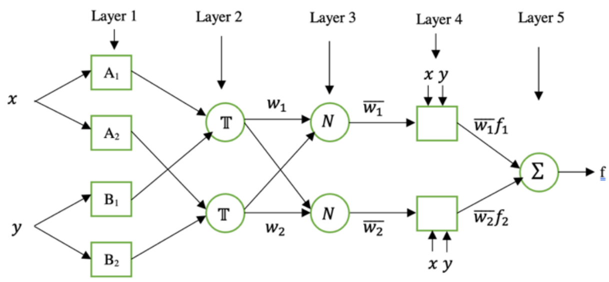

Fig. 1 – ANFIS architecture model.

We suppose that this system has two inputs x and y and one output f . The construction of each layer is given as follows.

- Layer 1

This layer contains neurons, which their number is equal to the fuzzy subsets. These neurons used a transfer function as given in eq. (1) (generally Gaussian function) to calculate the truth degree of a given fuzzy subset

$$f_i^1 = \mu_{Ai}(x), \tag{1}$$

where x is the input to neuron i, A_i is a fuzzy subset of x, f_1 is the membership function, and $\mu_{Ai}(x)$ is a function that varies between 0 and 1.

- Layer 2

The inputs of this hidden layer are the outputs from layer 1. They calculate the degree of activation function, which depends on the AND or OR operators. The activation function of the ith neurons from this layer is given by eq. (2)

$$W_k = \mu_{Ai}(x) \times \mu_{Bj}(x) \tag{2}$$

where i and j are the partition numbers of x and y, respectively, and k is the rule number.

- Layer 3

called the normalization layer; it is used to normalize the weights by calculating the ratio between the neuron weight by the sum of all rule weights as given in eq. (3);

$$\overline{W_k} = \frac{W_k}{\sum W_i} \tag{3}$$

- Layer 4

This layer is used to calculate the result parts of the rules, as given by

$$f_k^4 = \overline{W_k} \times f_k = \overline{W_k} \times (p_k x + q_k y + r_k) \tag{4}$$

(p_k, q_k, r_k) is a set of consequent parameters and W_k third layer output.

- Layer 5

This layer contains one output from the 4th layer, calculated using eq. (5).

$$f^5 = \sum_k \overline{W_k} \times f_k^4 \tag{5}$$

2.2. FUZZY C-MEANS

Fuzzy C-means (FCM) [24,25] is a clustering algorithm derived from the C-means algorithm. It consists of assigning data points into groups (clusters) with a certain degree based on the cluster center of gravity. The FCM algorithm can be summarized as follows.

- Step 1: set the following input parameters

The inputs $X = x_j\,(j = 1 \ldots N)$ vector set, the number of clusters C, the threshold representing the convergence error data ϵ, and the fuzzy degree m.

- Step 2: Initialization of matrix U (membership degree matrix) between [0,1].
- Step 3: update the matrix B (prototype matrix) using eq. (6) and eq. (7).

$$b_j \leftarrow \frac{\sum_{k=1}^{n} u_{ik}^m x_k}{\sum_{k=1}^{n} u_{ik}^m} \tag{6}$$

$$j^{old} \leftarrow \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m d^2(x_j, b_j) \tag{7}$$

- Step 4: update the membership degree matrix by eq. (8) and eq. (9)

$$u_{ij} \leftarrow \left[\sum_{k=1}^{c}\left(\frac{d^2(x_j,b_j)}{d^2(x_j,b_k)}\right)^{2/(m-1)}\right]^{-1} \quad (8)$$

$$j^{new} \leftarrow \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m} d^2(x_j,b_j) \quad (9)$$

- Step 5: repeat steps 3 and 4 until this criterion of eq. (10) is achieved.

$$\|j^{new} - j^{old}\| \leq \epsilon. \quad (10)$$

The outputs are the membership degree matrix U and the centroid of the cluster B.

2.3. SUBTRACTIVE CLUSTERING

Subtractive clustering [26] is used where the number of centroids for the data is unknown. It consists of the assumption that each data point has a protentional to belong to a cluster. This protentional is calculated by measuring the nearest points' data density. After that, a cluster is constructed by grouping some data according to how far each point is from the radius of this cluster. The algorithm repeats this mechanism each time until all clusters are constructed. The summary of the subtractive clustering algorithm is given as follows.

- Step 1: Select the first data group with the highest protentional to be a center.
- Step 2: Clear all points which are out of the 1st center of the cluster (determined by the radius).
- Step 3: Determine the next center.
- Step 4: Repeat until all data is within a cluster's radius.

2.4 GRID PARTITIONING

This algorithm is employed when the number of input variables is limited or minor. It allows dividing space into a grid-like structure to avoid overlapping parts in the input space [22]. There are several ways in which grid partitioning can be used to generate the specific partitioned areas, such as those that contain the fuzzy rule.

Generally, without considering the physical interpretation of the data density, the fuzzy rules are generated using the grid partitioning technique based only on the training sets (input-output). This mechanism helps to optimize the calculation time. However, it depends strongly on the input size and grid size; more precisely, a finer grid gives high performance.

3. **DATA**

For simulating and testing the proposed nonlinear ANFIS model. A solar radiation time series must be selected. To this end, we have chosen the city of Ghardaia, Algeria (Lat. : 32.4908° N, Long.: 3.6728° E). It is in the middle North of Algeria, precisely in the Sahara. It has more than 3000 sunny hours per year and an annual average global solar radiation of more than 6 kWh/m^2, contributing to the use of solar energy in various ways.

From Ghardaia's National Meteorological Office in Algeria, we've chosen two years solar irradiation time series and the clear sky, top of the atmosphere solar irradiation, the temperature, wind, humidity, and pressure time series between the period of 2019 until 2021 to test the proposed models. There is an error of 2 percent over the year in the global horizontal data measurement, measured using CM 11 pyranometer (Capderou 1986) at 1-hour intervals. These data are split into two groups: one for training the model and another for validating it.

4. **RESULTS AND DISCUSSIONS**

Our objective is to estimate the hourly solar irradiation using a nonlinear ANFIS model based on measured parameters such as temperature, wind, humidity, pressure, clear sky, and top of the atmosphere solar irradiation time series. Hence, the data were divided into a training set to train the proposed model based on grid partitioning, subtractive clustering, and FCM algorithms and a testing set to evaluate the goodness of the results. In addition, different error metrics such as RMSE, MSE, and error mean are selected to evaluate the accuracy of the estimation model. Moreover, a comparison is made at the end of this section to evaluate the best for similar cases.

4.1. RESULTS

In what follows, we have tested the three-clustering method to estimate the hourly solar radiation data. For each method, we have divided the data into training and testing data sets ranging from sunny to cloudy days. Moreover, we plotted the R2 values for each method for both the testing and training phases. In addition, a global comparison table has also been done to choose the best method for our purpose.

4.1.1. GRID PARTITIONING

The first method to test for estimation is grid partitioning. Several partitions have been tested, and we have chosen the best ones. The training and testing data results are shown in Fig. 2.

These figures show an RMSE equal to 50.1337 W/m^2 and an MSE equal to 2513.38 W/m^2 with an R^2 equal to 0.98681 for the testing set. Generally, the estimation results are good, but this method takes a high computation time, which is determined by the size of each input's membership function it is difficult to do in the case of numerous inputs for hourly solar radiation calculation.

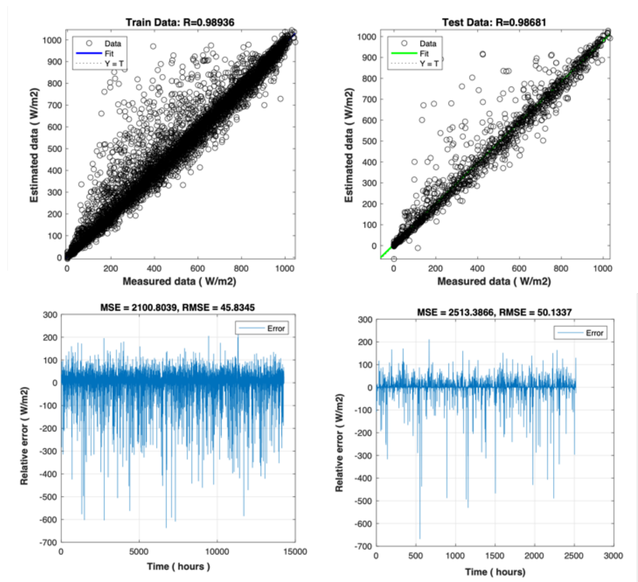

Fig. 2 – Results of RMSE, MSE, training, and testing of estimated hourly solar radiation data versus measured data using the grid partitioning algorithm.

4.1.2. SUBTRACTIVE CLUSTERING

This method depends strongly on the radius values. Low values can affect the ANFIS model mapping. However, high values increase the overfitting problem. The results of the R^2 between testing and training sets are shown in Fig. 3.

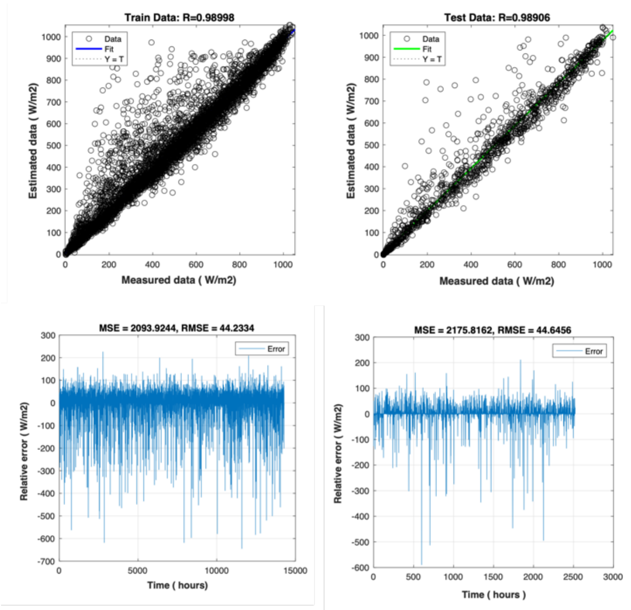

Fig. 3 – Results of RMSE, MSE, training, and testing of estimated hourly solar radiation data versus measured data using the subtractive clustering algorithm.

4.1.3. FCM

The FCM method seems the best one that gives the results – according to the R^2 values of the testing data sets and training ones, shown in Fig. 4.

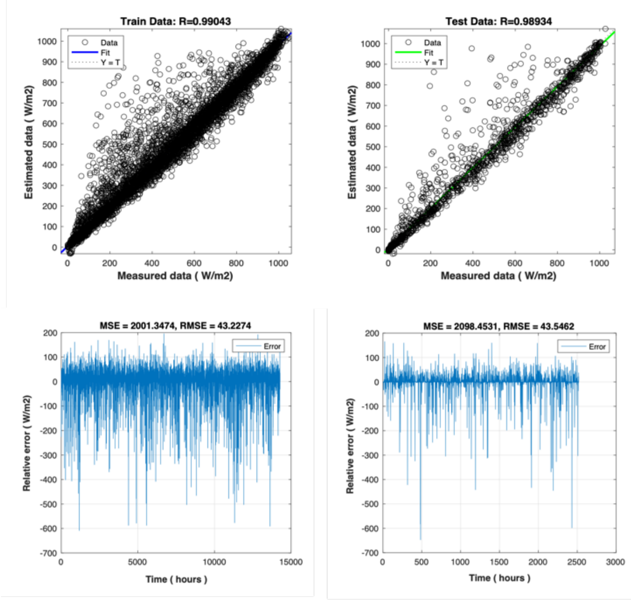

Fig. 4 – Results of RMSE, MSE, training, and testing of estimated hourly solar radiation data versus measured data using the FCM algorithm.

We can see a low RMSE and MSE. This is because of the influence of the number of nodes. Low nodes can affect the results because of the non-good partitioning of the inputs. On the other hand, a high number of nodes which is done by increasing the number of clusters, can. It gives good results but has more computational time compared to low nodes.

4.2. COMPARISON

Two random testing days were selected to evaluate the best model between grid partitioning, subtractive clustering, and FCM. A clear sky day (5th July 2021) and a cloudy day (3rd February 2021). The results of the three methods are shown in Fig. 5 and 6. The measured days are given in blue lines; another method is in other colors. Moreover, Table 1 shows the comparison results between these three methods.

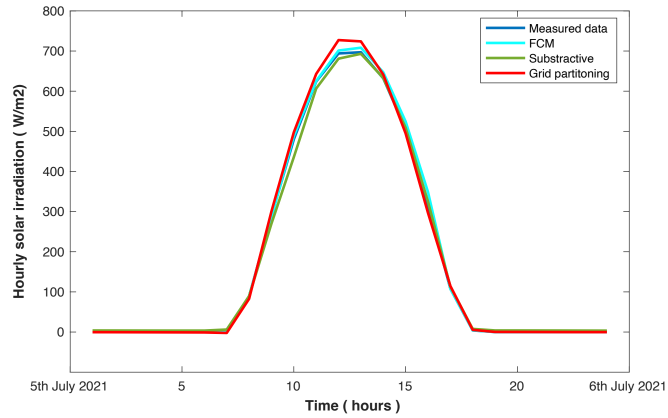

Fig. 5 – Comparison results between measured hourly solar radiation data and the estimated ones using FCM, subtractive clustering, and grid partitioning for the 6th July 2021 for Ghardaia, Algeria.

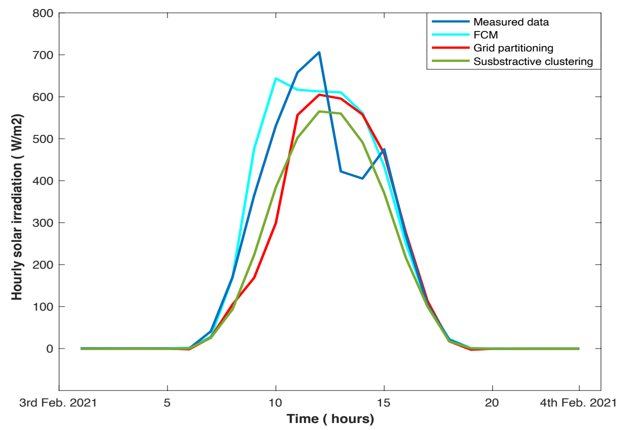

Fig. 6 Comparison results between measured hourly solar radiation data and the estimated ones using FCM, subtractive clustering, and grid partitioning for the day of the 4th of February 2021 for Ghardaia, Algeria.

Table 1.

RMSE, MSE, and R^2 of the comparison results between measured hourly solar radiation data and estimated using FCM using FCM, subtractive clustering, and grid partitioning for clear and cloudy days.

	Clear sky day			Cloudy day		
	RMSE	MSE	R^2	RMSE	MSE	R2
Subtractive clustering	11.44	130.988	0.969	74.184	5503	0.89
Grid partitioning	11.72	137.381	0.954	84.416	7126	0.86
FCM	08.97	80.564	0.982	63.807	4071	0.92

These figures and tables show that the three methods give good estimation results on clear sky day with an R^2 above the 95 % for all algorithms. However, in the case of cloudy days, the FCM algorithm gives good results compared to others with RMSE equals to 63.80 W/m^2 and MSE equal to 4071 W/m^2 with an R^2 equal to 0.9227. In addition, we should consider the computational time, which plays a significant role because selecting a suitable method that can take time for calculations is not favorable.

5. **CONCLUSION**

The hourly global solar irradiation on a horizontal surface on the Ghardaia site was estimated using clustered nonlinear ANFIS model. FCM, subtractive clustering, and grid partitioning have been selected for clustering the inputs dataset, consisting of meteorological measurements such as temperature, wind, humidity, pressure, clear sky, and top of the atmosphere solar irradiation time series. Statistical analysis was made using root means squared error, means squared error, and R squared error. A good agreement between the measured values and those estimated by the different models that can be used for similar applications.

We can see the goodness of the proposed nonlinear ANFIS model for estimation purposes. The number of clusters plays a significant role in the quality of the results; a low number of clusters affects the non-well partitioning of the inputs, whereas a high number of clusters increases the computation time.

Received on 2 November 2022

REFERENCES

1. M.J. Pickl, *The renewable energy strategies of oil majors – From oil to energy?* Energy Strategy Reviews, **26**, pp. 100370-100378 (2019).
2. R. Dutta, K. Chanda, R. Maity, *Future of solar energy potential in a changing climate across the world: A CMIP6 multi-model ensemble analysis*, Renew Energy, **188**, pp. 819–829 (2022).
3. T. Güney, *Solar energy and sustainable development: evidence from 35 countries*, International Journal of Sustainable Development & World Ecology, **29**, *2*, pp. 187–194 (2021).
4. F. Antonanzas-Torres, R. Urraca, J. Polo, O. Perpiñán-Lamigueiro, R. Escobar, *Clear sky solar irradiance models: A review of seventy models*, Renewable and Sustainable Energy Reviews, **107**, pp. 374–387 (2019).
5. A. Teke, H.B. Yildirim, Ö. Çelik, *Evaluation and performance comparison of different models for the estimation of solar radiation*, Renewable and Sustainable Energy Reviews, **50**, pp. 1097–1107(2015).
6. B.M. Lopes, A.M.G. Costa, W. Uturbey, *Comparative analysis of estimation models for monthly average hourly solar radiation*, Journal of Control, Automation and Electrical Systems, pp. 1–18 (2022).
7. A.H. Ajil, H.A. AL-Rikabi Rashad, *An empirical model for calculating monthly average daily global solar radiation based on temperature for five cities of different climates in Iraq*, Journal of Applied Science and Engineering, **21**, *4*, pp. 497–506 (2018).
8. A.S. Tomlin, R. Crook, C.J. Smith, J. Gooding, *Solar resource estimation using a radiative transfer with shading (RTS) model*, 31st European Photovoltaic Solar Energy Conference and Exhibition, **1**, pp. 2808–2813 (2015).
9. R. Tapimo, R. Tapimo, M. Lazard, G.L. Ymeli, G.L. Ymeli, D. Yemele, *Radiative transfer model for ground surface irradiance estimation: clear sky*, JOSA A, **38**, *11*, pp. 1640–1646 (2021).
10. J. Zhang, L. Zhao, S. Deng, W. Xu, Y. Zhang, *A critical review of the models used to estimate solar radiation*, Renewable and Sustainable Energy Reviews, **70**, pp. 314–329 (2017).
11. H.B. Yıldırım, A. Teke, F. Antonanzas-Torres, *Evaluation of classical parametric models for estimating solar radiation in the Eastern Mediterranean region of Turkey*, Renewable and Sustainable Energy Reviews, **82**, pp. 2053–2065 (2018).
12. B.S. Masabi, Z. Aghashariatmadari, S. Hejabi, *Evaluation the efficiency of a parametric model based on MODIS data for solar radiation estimation in comparison with some empirical models*, Arabian Journal of Geosciences, **14**, *15*, pp. 1–13 (2021).
13. B. Jahani, Y. Dinpashoh, A. Raisi Nafchi, *Evaluation and development of empirical models for estimating daily solar radiation*, Renewable and Sustainable Energy Reviews, **73**, pp. 878–891(2017).
14. M.A. Hassan, A. Khalil, S. Kaseb, M.A. Kassem, *Independent models for estimation of daily global solar radiation: A review and a case study*, Renewable and Sustainable Energy Reviews, **82**, pp. 1565–1575 (2018).
15. A. Angstrom, *Solar, and terrestrial radiation. Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation*, Quarterly Journal of the Royal Meteorological Society, **50**, *210*, pp. 121–126 (1924).
16. A. Teke, H.B. Yildirim, Ö. Çelik, *Evaluation and performance comparison of different models for the estimation of solar radiation*, Renewable and Sustainable Energy Reviews, **50**, pp. 1097–1107 (2015).
17. D. Patel, S. Patel, P. Patel, M. Shah, *Solar radiation and solar energy estimation using ANN and Fuzzy logic concept: A comprehensive and systematic study*, Environmental Science and Pollution Research, **29**, *22*, pp. 32428–32442 (2022).
18. H. Jiang, N. Lu, J. Qin, W. Tang, L. Yao, *A deep learning algorithm to estimate hourly global solar radiation from geostationary satellite data*, Renewable and Sustainable Energy Reviews, **114**, p. 109327 (2019).
19. K. Benmouiza, A. Cheknane, *Quantification of solar radiation in Algeria, application to the sizing of photovoltaic systems*, Abou Bekr Belkaid University, Tlemcen, 2015. Accessed: Jun. 22, 2022.
20. O.D. Samuel, M.O. Okwu, O.J. Oyejide, E. Taghinezhad, A. Afzal, M. Kaveh, *Optimizing biodiesel production from abundant waste oils through empirical method and grey wolf optimizer*, Fuel, **281**, p. 118701 (2020).
21. B.B. Uzun, M. Kiliç, N. Özbay, A.E. Pütün, E. Pütün, *Biodiesel production from waste frying oils: Optimization of reaction parameters and determination of fuel properties*, Energy, **44**, *1*, pp. 347–351(2012).
22. K. Benmouiza, A. Cheknane, *Clustered ANFIS network using fuzzy c-means, subtractive clustering, and grid partitioning for hourly solar radiation forecasting*, Theor. Appl. Climatol., **137**, *1–2*, pp. 31–43 (2019).
23. D. Yu, *Estimation of pile settlement socketed to rock applying hybrid ALO-ANFIS and GOA-ANFIS approaches*, Journal of Applied Science and Engineering, **25**, *6*, pp. 979–992(2022).
24. J.C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*, Pattern Recognition with Fuzzy Objective Function Algorithms (1981).
25. J.C. Dunn, *A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters*, Journal of Cybernetics, **3**, *3,* pp. 32–57 (2008).
26. R.R. Yager, D.P. Filev, *Generation of fuzzy rules by mountain clustering*, Journal of Intelligent & Fuzzy Systems, **2,** *3*, pp. 209–219 (1994).