An evolutionary method for creating ensembles with adaptive size neural networks for predicting hourly solar irradiance

Raka Jovanovic*, Luis M. Pomares†, Yasir E. Mohieldeen‡, Daniel Perez-Astudillo§ and Dunia Bachour¶
Qatar Environment and Energy Research Institute,
Hamad bin Khalifa University,
PO Box 5825, Doha, Qatar
Email: *rjovanovic@hbku.edu.qa, †lpomares@hbku.edu.qa, ‡ymohieldeen@hbku.edu.qa,
§dastudillo@hbku.edu.qa, ¶dbachour@hbku.edu.qa

***Abstract*—In this paper we propose a hybridized approach for finding high quality artificial neural network (ANN) for calculating hourly estimates of solar irradiance. These properties are essential for performance analysis of solar based energy generation. To be more precise the hourly global horizontal irradiance (GHI), direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI) are estimated based on ANNs which are trained using satellite and ground measurement data. In the proposed method we explore the effect of combining the measured data with properties derived from the standard physical models. The performance of the method is improved by using a genetic algorithm in two ways. First by selecting the parameters that are used for training the ANN. Secondly by adapting the size of the hidden layer of the ANN based on the number of selected input parameters. The adaptive size based approach proves to be especially suitable for ANN ensembles. In our computational experiments we evaluate the effectiveness of the proposed method on feedforward neural network. The results show that the adaptability of the ANN manages to notably improve the performance when compared to the standard approach using a fixed size of the hidden layer.**

***Index Terms*—Global solar radiation, Artificial neural network, Ensemble model, Evolutionary artificial neural network**

I. Introduction

In the last decade there has been an astonishing increase in solar generated electrical energy. Information regarding global horizontal irradiance (GHI), direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI) are essential for the performance analysis and financial evaluation of solar power systems. The problem comes from the fact that these values are only measured at a small number of radiometric stations but their values are required over vast areas, and consequently they need to be estimated. The two most important applications of such predictions are: first, the development of solar maps which are used for finding optimal locations for solar facilities [1]; and second, the forecasting of GHI, DNI and DHI which are of utmost importance for the integration of solar generated energy into the electrical grid [2].

These values can be predicted using different mathematical models based on the physics and the dynamic of the atmosphere [3] or by using statistical approaches. Early mathematical models considered the relation between solar radiation, and sunshine duration as linear [4], [5]; later more effective non-linear ones have been proposed based on the maximum possible duration of sunshine [6], [7] and temperature [7].

One of the most successful statistical tools for this task are artificial neural networks (ANNs), for which a systematic review can be found in [8], [9]. In such applications there is a general preference for feedforward neural network (FNN). Some examples are the use ANNs for estimating hourly [10], daily [11], [12] and monthly [13] solar irradiation properties. Another important application is in short-term forecasting [14] for power predication of grid-connected photovoltaic (PV) plants. However, networks based on radial basis function (RBF) have also been used in this field. One example is the use of RBFs for estimating daily global solar radiation [15]. In most cases multilayer perceptron networks (MLP) manage to out perform RBFs; one such example can be seen in article [16] in the case of estimation of GHI. Research has also been conducted on applying recurrent neural networks for forecasting the solar radiation using the past solar radiation and solar energy [17]. The performance of ANNs is commonly improved by exploiting information about cloud conditions [1] and using different types of preprocessing steps. For example, the performance of ANNs was enhanced using wavelet based denoising [18] or by using linear based filters [19], [20].

ANNs for predicting solar irradiance, in general, use visible and/or infrared channels from satellite data in combination with time/space based parameters. The latter group of parameters are latitude, longitude, date and time, but due to better performance, in a vast majority of implementations, the derived physical properties (calculated using additional measurements for the region of interest), are used instead. It has been shown that the predictions of solar irradiance by ANNs can be significantly improved by selecting a good subset of all the potential input parameters. This selection can be done effectively by using genetic algorithms (GA) [12], [21].

In this paper we extend the work on hybridizing ANNs using GA for parameter selection. In the currently published research only the physical properties that directly affect the solar irradiance are used as potential input parameters of the ANN. In the proposed research we explore the effect of extending the set of potential parameters with the basic time/space parameters from which these values are derived. One of the drawbacks in the existing approaches for input parameter selection for an ANN using GA is that this is done for networks with a fixed size of the hidden layers. It is a well known fact that the best performance of an ANN can be achieved if some relations between the number of input parameters and the number of neurons in the hidden layer are satisfied. In this approach we explore the effectiveness of having an adaptive size of the ANN that is used in the GA. In our computational experiments we explore the effectiveness of such an approach on FNN and on a basic ANN ensemble. The training of the networks is done based on hourly ground measurements and satellite data for a period of six months.

The paper is organized as follows. In the next section we explain the motivation for the proposed method. In the third section the details of the hybridized method are presented. In the last section the results of the conducted computational experiments are shown.

II. Motivation for Method

In this section we give the motivation for the hybridized GA/ANNs method for estimating GHI, DNI and DHI. For estimating these values we use 13 input parameters and a single output value based on ground measurements. More details regarding the input and output parameters will be given in the results section. As stated in article [12], the optimal estimations of solar irradiance, achieved by an ANN, are dependent on the selected input parameters. This is due to the fact that input parameters, corresponding to physical properties of the system, potentially have redundant or nonessential information for the values being predicted. Because of this the performance of the ANN can be improved by choosing an optimal selection of input parameters. The problem is that such a selection cannot, in the case of solar irradiance estimation, be done based only on known physical properties. It should be understood that the use of ANNs comes from the fact that existing physical models are not able to produce sufficiently precise estimates.

In case of the problem of interest it is not possible, on desktop computers, to evaluate all potential combinations of input parameters due to the large number of potetntial combinations and the high computational cost of training an ANN. Because of this the selection of optimal parameters is done using some optimization method. Published research has proven that GAs are very suitable for this task. The problem with existing methods is that they use the GA for selecting the set of input parameters for training an ANN with a fixed size of the hidden layer. There are two problems with this approach. First, the size of the hidden layer is selected using a rule of thumb through trial and error. Secondly, it is a well known fact that the optimal number of neurons in the hidden layer is dependent on the number of input/output parameters. Some of the most common rules are

- The number of hidden neurons should be between the size of the input layer and the size of the output layer.
- The number of hidden neurons should be 2/3 the size of the input layer, plus the size of the output layer.
- The number of hidden neurons should be less than twice the size of the input layer.

It is evident that a different number of input parameters will give us a different range for the optimal number of neurons based on the basic rules. In the proposed method we address this issue by adapting the size of the hidden layer of the ANN based on the number of parameters that are selected inside the fitness function of the GA. In the following subsections we give details on how this concept is applied.

III. Hybridized Method

For the sake of clarity the method will be presented in the following way. First we will give specifics of the ANN and the used notation, next the definition of GA and finally the interaction between the two algorithms. In the following text we will use the following notation. The set of input parameters I will be a matrix $n \times m$, where each row represents one set of input parameters. In relation we will have the target values O as a matrix $1 \times m$. Although the ANN will be used to a large extent as a black box we will give its specifics. It will be an FNN, using the hyperbolic tangent sigmoid transfer function. The training method will be Levenberg-Marquardt backpropagation. The ANN will be trained in the standard way using a set of T training, V validation, E test rows from I and corresponding ones in O.

A. GA Chromosome

The goal of the proposed method is to generate the best ANN for finding estimates of solar irradiance properties. The idea is to achieve this through using a GA for finding the best set of parameters and have an adequate size of the hidden layer of the ANN. We will define the chromosomes for the GA in a natural way. A chromosome will consist of n binary values, were each one states if a parameter is being used in designing the ANN. This can be better understood by observing Figure 1. The adaptive size of the ANN has been calculated using the following simple formula

$$N = \left\lceil \alpha \sum_{1}^{n} c_i \right\rceil \quad (1)$$

In Eq. 1 N is the number of neurons in the hidden layer, c_i is a binary variable which states whether input parameter i is used in the ANN. The number of the neurons will be the ceiling function of α times the number of used parameters.

B. GA Fitness Function

The objective of the fitness function is to give us a method of evaluating the quality of neural networks that can be generated based on a specific chromosome. In the common application of GA there is a unique correspondence between

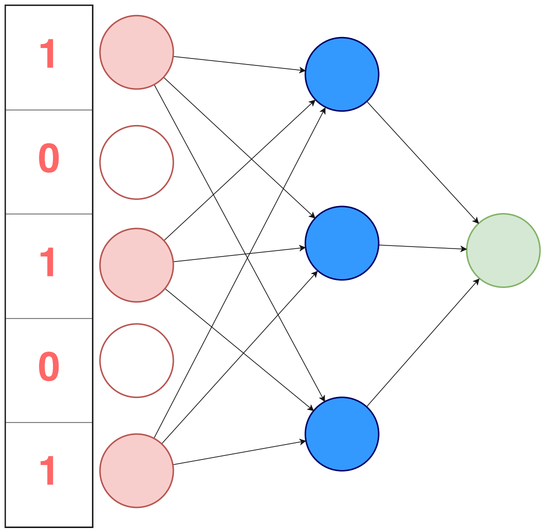

Fig. 1. Illustration of the relation between the chromosome in the GA and the corresponding ANN. The red nodes represent potential input parameters, the filled ones are used as input in the ANN. Blue/green nodes are used for the hidden/output layer of the ANN. The size of the hidden layer is selected based on $\alpha = 0.75$

a chromosome and the value of the fitness function, which is also the value that we wish to minimize. In case of the problem of interest this relation is more complex. To be more precise the goal of the proposed method is to find the best ANN for estimating the output parameter based on input parameters. The problem comes from the fact that a chromosome only gives us information about the structure of the ANN. It is well known that the quality of estimates that can be achieved using an ANN depends on the training set. On the other hand although a trained good structure will not always give us a better estimate than a lower quality one, it is expected that it will be better in a majority of cases. Because of this the hybridized algorithm should aim to train promising structures multiple times.

Let us assume that we have an available a set of input parameters I and corresponding output values O. The set I/O will be further divided into subsets I_g/O_g for comparison of generated ANNs and I_l/O_l for training ANNs inside the fitness function. As previously stated, for one chromosome value an ANN can be trained in different ways depending on the training set. Because of this each time a chromosome is evaluated inside the fitness function the sets I_l/O_l will be partitioned into different random validation, training and test sets.

The standard way of evaluating the quality of an estimate for solar irradiance is by using the relative root-mean-square error ($rRMSE$). For the sake of completeness we include its definition

$$rRMSE = \frac{1}{\bar{o}}\sqrt{\sum_{i=1}^{n}(o_i - \hat{o}_i)^2} \tag{2}$$

In Eq. 2 o_i is the output value corresponding to input parameter set p_i. $\hat{o}_i$ is the estimate acquired by applying the ANN to input parameter set p_i. Further, $\bar{o}$ is the mean value of the output value set. In the proposed method the fitness value will be defined using the following formula.

$$f(c) = \frac{1}{2}(rRMSE_l + rRMSE_g) \tag{3}$$

Eq. 3 gives the value of the fitness function from a specific value of the chromosome c. The fitness value will be equal to the average of two $rRMSE$. The first one is $rRMSE_l$, based on the selected training set which we can consider local. The second one is used as a global test which is calculated based on the sets I_g and O_g which will be used for all ANNs that are tested. The purpose of the second pair is to have an additional measure that can be considered ”more uniform” for all the ANNs.

C. Implementation Details

In the majority of applications of GA to specific problems, the metaheuristic is used through existing packages or tool boxes. In the general case only the fitness function and chromosomes are defined and implemented by the user. On the other hand, mutation, crossover, elitism and selection operations of the GA are, except for specifying performance parameters, used as a black box. Due to the fact that the proposed method was implemented in this manner, we only give the details of the fitness function for which the pseudocode can be seen in Algorithm 1.

Algorithm 1 Fitness function implementation

function FITNESS(C)
  global I_l, O_l, I_G, I_G
  global $Best$

  Normalize I,O, and store $NormParam$
  Randomly divide I_l into I_t, I_v, I_e
  Randomly divide O_l into O_t, O_v, O_e

  $Ann = train(I_t, O_t, I_v, O_t, I_e, O_e)$

  $rRMSE_l = rRMSE(Ann, O, I, NormParam)$
  $rRMSE_g = rRMSE(Ann, O_G, I_G, NormParam)$

  $Result = \frac{1}{2}(rRMSE_l + rRMSE_g)$

  if $Result < Best.Value$ **then**
    $Best.Value = Result$
    $Best.NN = ANN$
    $Best.Param = C$
    $Best.NormParam = NormParam$
  end if
  return $Result$

end function

In the fitness function global variables are used for the previously described sets I_l, O_l, I_g, I_g. The first step is normalizing the data and storing the normalization parameters $NormParam$. Next, the available data will randomly divided for training the artificial neural network Ann. The trained network Ann will be used to calculated the value of the fitness function based on Eq. 3. As previously stated the results of training of ANN will not always be the same as it is dependent on the selected training set, because of this it is necessary to store the best found one. This is done by using a global structure $Best$ which contains information about the best value of the fitness function, the trained ANN, the set of used parameters and the normalization parameters.

D. Improvement using ANN Ensemble

One of the standard and most effective ways of improving the performance of ANNs is the use of ensembles. Ensembles exploit the simple concept of using multiple neural networks jointly to solve a problem. Research has shown that the generalization capabilities of such systems can outperform those of single networks [22]. The use of ensembles has proven to be very effective for enhancing the performance of ANNs for solar irradiance prediction [10], [12], [1], [23].

It is a well known fact that for such aggregate systems to be effective the individual networks must be as accurate and diverse as possible [24], [22]. The diversity of the used ANNs is generally achieved in two ways: by using different training sets and by having ANNs with different structures. The proposed hybridized method has been designed in a way that multiple executions of the probabilistic algorithm generate ANNs having these properties. To be more precise, the fitness function uses a variety of different training sets and the adaptive size of the hidden layer provides diversity in the structure.

Although there is a wide range of possible methods for creating ANN ensembles even the use of simple averaging produces a high level of improvement [25]. We have used this simple approach to enhance the performance of the method, based on the following equation

$$ANN_e(k) = \frac{1}{n}\sum_{i=1}^{n} ANN_i(k) \qquad (4)$$

In Eq. 4 the notation ANN_i is used for an ANN having a specific input parameters set and training data. The estimate acquired using the ensemble ANN_e for some input parameters k is equal to the average value of $ANN_i(k)$.

IV. Results and Discussion

In this section we present results of the conducted computational experiments. The goal was to evaluate the effectiveness of using adaptive size ANNs for input parameter selection. To be more precise, we explore the effect of having an adaptive size of the hidden layer of an ANN, instead of a fixed one as in currently published research, when the set of input parameters is selected using GA. Further, we conduct a comparison of the quality of estimates that can be achieved with ensembles consisting of ANNs generated in these two ways. In the further text we first give details of the data used for training the ANNs and later the experimental setup and discuss the results.

A. Training Data

The objective of the proposed method was to estimate the hourly GHI, DNI and DHI values. The ground measurements used in this study were collected with a radiometric station operated by the Qatar Environment and Energy Research Institute (QEERI) in Doha, latitude $25.33°$ N, longitude $51.43°$ E, and altitude $10m$. The station [26] is equipped with BSRN-quality (Baseline Surface Radiation Network, http://bsrn.awi.de) instrumentation mounted on a Solys2 sun tracker with sun sensor, measuring DNI with a CHP1 pyrheliometer, while GHI and DHI are measured with separate CMP11 pyranometers, one of them (for DHI) shaded, and both of them with ventilator units; the station is maintained regularly by cleaning all sensors and checking the alignment and level of the sensors and tracker. Data are collected as one-minute averages in W/m^2, and were averaged within each hour to obtain the hourly values of GHI, DNI and DHI used here.

The input parameters that have been used are the following:

- Year
- Month
- Day
- Hour
- Julian day
- Instantaneous albedo
- Cosine of sun zenith angle
- Scattering angle. Angle between sun and satellite
- Extraterrestrial irradiance. Top of the atmosphere irradiance
- Clear sky Global Horizontal Irradiance
- Clear sky Direct Normal Irradiance
- Air Mass. Number of atmospheres the sun rays have to traverse compared with an air mass of one, which is the normal depth of the atmosphere for sun zenithal angle of 0.
- GEO satellites images from Meteosat First Generation (MFG-IODC) having a $3km$ resolution for which the pixel containing ground station location was extracted. Only the visible channel has been used.

It is important to note that the available fifteen-minute satellite raw data were averaged and converted into hourly effective radiances in the same way as in article [23]. The data used for training the ANNs was collected in the period from 1.1.2015 to 31.5.2015.

B. Computational Experiments

In this subsection we give details of the performed computational experiments. Both hybridized algorithms, using fixed ($fANN$) and adaptive ($aANN$) size of the hidden layer of the ANN, for estimating the solar irradiance properties have been implemented in MatLab R2013a. The training of the ANNs has been done using the built-in MatLab functions. The details of the chosen ANNs are as described in Section III. The GA has been applied using the MatLab built-in function using the default parameters, with the previously described fitness function. The parameter for determining the size of the hidden layer was $\alpha = 0.75$. In case of the method having a fixed size hidden layer the number of neurons was 10. This value was chosen empirically from a large number of tests that were conducted for having the best performance. In the case of both algorithms, 15% of the available data was used for evaluating the performance based on $rRMSE$ as test data I_{MT}, O_{MT}. The rest of the data has been used for finding optimal ANNs in the following way. Of the remaining data, 15% was used for sets I_g, O_g. All remaining data has been used for training the ANNs in the fitness function of the GA, out of which 15% was used for validation and another 15% for testing. All the calculations have been done on a machine with Intel(R) Core(TM) i7-2630 QM CPU 2.00 GHz, 4GB of DDR3-1333 RAM, running on Microsoft Windows 7 Home Premium 64-bit.

Due to the probabilistic nature of the algorithm multiple runs of each of the algorithms have been conducted. To be more precise, for each method the algorithm was executed 50 times using different random seeds. With the intention of having a fair comparison the tests have been paired in the following way. The same data sets I_{MT}, O_{MT}, I_g and O_g have been used for both methods. Each of the algorithms was executed for a fixed time limit of 120 seconds for each ANN. In our experiments we have explored the performance of ensembles having from 2 to 5 ANNs. Again to have a fair comparison we have used up to 5 pairs of ANNs, for fixed and adaptive hidden layer size, as described in the previous text. Such test have been done separately for DNI, GHI and DHI.

TABLE I
COMPARISON OF HYBRIDIZED METHODS FOR ESTIMATING DNI USING PARAMETER HAVING FIXED (F) OR ADAPTIVE (A) SIZE OF HIDDEN LAYER OF THE ANN. N IS USED TO INDICATE THE NUMBER OF ANNS IN THE ENSEMBLE.

N	Type	Avg	Stdev	Min	Max
1	F	31.6	2.7	27.5	39.0
1	A	31.4	3.3	26.0	37.4
2	F	30.1	3.0	25.5	37.3
2	A	30.0	3.2	25.0	35.9
3	F	29.6	3.1	24.7	35.7
3	A	29.2	3.0	24.7	35.2
4	F	29.4	3.1	24.5	35.8
4	A	29.0	3.1	24.6	35.2
5	F	29.1	3.2	24.0	35.8
5	A	28.8	3.1	24.5	35.1

The evaluation of the performance of the two methods has been done by observing the average $rRMSE$ for the 50 performed runs. To have a better comprehension of the performance we have also included the standard deviation, maximal and minimal values of the $rRMSE$. The results of the performed computational experiments can be seen in Tables I, II, III. The first thing that can be observed is that for both $fANN$ and $aANN$ the improvement achieved using the simple ensemble is notable. For both methods the best estimates have been achieved for DNI, GHI and DHI when ensembles with 5 ANNs are used and it gives an improvement of around 10% when compared to a single ANN of the same type. When we observe the average value of the $rRMSE$, for each of the predicted values, $aANN$ shows an improvement compared to $fANN$ for all estimated values and all sizes of ensembles. In case of minimal and maximal values of $rRMSE$ $aANN$ shows an overall better performance but it is not consistent. The robustness of both methods, in the sense of the range of quality of estimates, is not very strong. The difference between the best and worst $rRMSE$ is around $30-40\%$. It is important to note that similar behavior can be observed for other existing methods used to predict hourly solar irradiance [10].

TABLE II
COMPARISON OF HYBRIDIZED METHODS FOR ESTIMATING GHI USING PARAMETER HAVING FIXED (F) OR ADAPTIVE (A) SIZE OF HIDDEN LAYER OF THE ANN. N IS USED TO INDICATE THE NUMBER OF ANNS IN THE ENSEMBLE.

N	Type	Avg	Stdev	Min	Max
1	F	12.9	1.9	10.0	16.5
1	A	12.5	1.5	9.7	15.0
2	F	12.7	2.2	9.3	17.0
2	A	12.3	1.6	9.4	14.8
3	F	12.6	2.0	9.1	15.6
3	A	12.2	1.6	9.3	14.8
4	F	12.3	1.9	9.0	15.2
4	A	12.1	1.7	9.4	15.1
5	F	12.3	1.9	9.1	14.9
5	A	12.0	1.7	9.2	15.0

TABLE III
COMPARISON OF HYBRIDIZED METHODS FOR ESTIMATING DHI USING PARAMETER HAVING FIXED (F) OR ADAPTIVE (A) SIZE OF HIDDEN LAYER OF THE ANN. N IS USED TO INDICATE THE NUMBER OF ANNS IN THE ENSEMBLE.

N	Type	Avg	Stdev	Min	Max
1	F	29.3	3.0	24.2	35.9
1	A	29.2	3.1	24.5	37.8
2	F	28.3	2.7	24.8	34.7
2	A	27.6	2.6	24.1	34.5
3	F	27.6	2.9	23.9	34.9
3	A	27.1	2.6	22.6	34.4
4	F	27.1	2.7	23.5	33.6
4	FA	26.8	2.6	22.9	34.3
5	F	27.1	2.6	23.6	32.8
5	A	26.5	2.5	23.2	34.2

With the goal of having a more exhaustive comparison of the $fANN$ and $aANN$ we have also performed tests to confirm statistical significance. To be more precise, a paired single tailed Student t-Test has been done with a p value of 0.05. In the case of a single $ANNs$ statistical significance has only been shown in the case of GHI. In the case of ensembles consisting of 2-5 ANNs the better performance of $aANN$ has been confirmed in all the cases, except for predicting DNI with an ensemble with two ANNs.

V. Conclusion

In this paper we have presented an hybridized GA/ANN approach for predicting solar irradiance properties. The proposed method extends existing research for selecting optimal input parameters for training the ANN using a GA. The novelty of the approach is in using ANNs with an adaptive size of the hidden layer. The method is designed in a way that it is especially suitable for ANN ensembles. The performance of the method was tested using ground measured and satellite data for predicting DNI, GHI and DHI. The conducted computational experiments show that the use of an adaptive number of neurons in the hidden layer outperforms the standard approach, in the case of single ANN or ensembles. In almost all the cases these results have proven to be statistically significant.

In the future we plan to extend the work in several directions. First by using a heuristic approach for selecting the best ANNs to use in the ensemble since this type of approach has shown to provide significant improvement [27]. Further, we plan to include cloud conditions in the model and adapt it to specifics of the Gulf Cooperation Council (GCC) region. Finally, extend the number of potential input parameters to include all the available channels from the satellite data and use a larger number of ground measuring stations from the GCC region.

References

[1] M. H. Alobaidi, P. R. Marpu, T. B. M. J. Ouarda, and H. Ghedira, "Mapping of the solar irradiance in the uae using advanced artificial neural network ensemble," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 8, pp. 3668–3680, Aug 2014.

[2] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renewable and Sustainable Energy Reviews*, vol. 27, pp. 65 – 76, 2013.

[3] L. Martín, L. F. Zarzalejo, J. Polo, A. Navarro, R. Marchante, and M. Cony, "Prediction of global solar irradiance based on time series analysis: application to solar thermal power plants energy production planning," *Solar Energy*, vol. 84, no. 10, pp. 1772–1781, 2010.

[4] M. Rietveld, "A new method for estimating the regression coefficients in the formula relating solar radiation to sunshine," *Agricultural Meteorology*, vol. 19, no. 2, pp. 243–252, 1978.

[5] R. Benson, M. Paris, J. Sherry, and C. Justus, "Estimation of daily and monthly direct, diffuse and global solar radiation from sunshine duration measurements," *Solar energy*, vol. 32, no. 4, pp. 523–535, 1984.

[6] R. Said, M. Mansor, and T. Abuain, "Estimation of global and diffuse radiation at tripoli," *Renewable energy*, vol. 14, no. 1, pp. 221–227, 1998.

[7] A. Katiyar, V. Katiyar, A. Kumar, and C. Pandey, "Sixth order empirical relationship for estimating global solar radiation from sunshine hours," *Jour. PAS*, vol. 15, pp. 156–165, 2008.

[8] A. Qazi, H. Fayaz, A. Wadi, R. G. Raj, N. Rahim, and W. A. Khan, "The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review," *Journal of Cleaner Production*, vol. 104, pp. 1 – 12, 2015.

[9] F. S. Tymvios, S. C. Michaelides, and C. S. Skouteli, "Estimation of surface solar radiation with artificial neural networks," in *Modeling Solar Radiation at the Earths Surface*. Springer, 2008, pp. 221–256.

[10] A. Linares-Rodriguez, S. Quesada-Ruiz, D. Pozo-Vazquez, and J. Tovar-Pescador, "An evolutionary artificial neural network ensemble model for estimating hourly direct normal irradiances from meteosat imagery," *Energy*, vol. 91, pp. 264 – 273, 2015.

[11] O. Şenkal, "Solar radiation and precipitable water modeling for turkey using artificial neural networks," *Meteorology and Atmospheric Physics*, vol. 127, no. 4, pp. 481–488, 2015.

[12] A. Linares-Rodriguez, J. A. Ruiz-Arias, D. Pozo-Vazquez, and J. Tovar-Pescador, "An artificial neural network ensemble model for estimating global solar radiation from meteosat satellite images," *Energy*, vol. 61, pp. 636 – 645, 2013.

[13] O. Şenkal, "Modeling of solar radiation using remote sensing and artificial neural network in turkey," *Energy*, vol. 35, no. 12, pp. 4795–4801, 2010.

[14] F. Wang, Z. Mi, S. Su, and H. Zhao, "Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters," *Energies*, vol. 5, no. 5, p. 1355, 2012.

[15] A. Mellit, M. Menghanem, and M. Bendekhis, "Artificial neural network model for prediction solar radiation data: application for sizing stand-alone photovoltaic power system," in *IEEE Power Engineering Society General Meeting, 2005*, June 2005, pp. 40–44 Vol. 1.

[16] H. A. Hejase, M. H. Al-Shamisi, and A. H. Assi, "Modeling of global horizontal irradiance in the united arab emirates with artificial neural networks," *Energy*, vol. 77, pp. 542 – 552, 2014.

[17] N. Zhang and P. K. Behera, "Solar radiation prediction based on recurrent neural networks trained by levenberg-marquardt backpropagation learning algorithm," in *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*, Jan 2012, pp. 1–7.

[18] L. Lyu, M. Kantardzic, and E. Arabmakki, "Solar irradiance forecasting by using wavelet based denoising," in *Computational Intelligence for Engineering Solutions (CIES), 2014 IEEE Symposium on*, Dec 2014, pp. 110–116.

[19] A. A. Moghaddam and A. R. Seifi, "Study of forecasting renewable energies in smart grids using linear predictive filters and neural networks," *IET Renewable Power Generation*, vol. 5, no. 6, pp. 470–480, November 2011.

[20] F. O. Hocaolu, mer N. Gerek, and M. Kurban, "Hourly solar radiation forecasting using optimal coefficient 2-d linear filters and feed-forward neural networks," *Solar Energy*, vol. 82, no. 8, pp. 714 – 726, 2008.

[21] R. Marquez and C. F. Coimbra, "Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the {NWS} database," *Solar Energy*, vol. 85, no. 5, pp. 746 – 756, 2011.

[22] P. Granitto, P. Verdes, and H. Ceccatto, "Neural network ensembles: evaluation of aggregation algorithms," *Artificial Intelligence*, vol. 163, no. 2, pp. 139 – 162, 2005.

[23] S. Quesada-Ruiz, A. Linares-Rodríguez, J. Ruiz-Arias, D. Pozo-Vázquez, and J. Tovar-Pescador, "An advanced ann-based method to estimate hourly solar radiation from multi-spectral msg imagery," *Solar Energy*, vol. 115, pp. 494–504, 2015.

[24] A. Krogh, J. Vedelsby *et al.*, "Neural network ensembles, cross validation, and active learning," *Advances in neural information processing systems*, vol. 7, pp. 231–238, 1995.

[25] A. J. C. Sharkey, "On combining artificial neural nets," *Connection Science*, vol. 8, no. 3-4, pp. 299–314, 1996.

[26] D. Perez-Astudillo and D. Bachour, "DNI, GHI and DHI Ground Measurements in Doha, Qatar," *Energy Procedia*, vol. 49, pp. 2398–2404, 2014, Proceedings of the SolarPACES 2013 International Conference.

[27] Z.-H. Zhou, J. Wu, and W. Tang, "Ensembling neural networks: Many could be better than all," *Artificial Intelligence*, vol. 137, no. 1, pp. 239 – 263, 2002.