IOP Conf. Series: Earth and Environmental Science **161** (2018) 012017 doi:10.1088/1755-1315/161/1/012017

ANN-based modelling and prediction of daily global solar irradiation using commonly measured meteorological parameters

M Marzouq[1], Z Bounoua[2], A Mechaqrane[2], H E Fadili[1], Z Lakhliai[1] and K Zenkouar[3]

[1]Laboratory of Computer Sciences and Interdisciplinary Physics, National School of Applied Sciences, USMBA, Fez, PO. Box 72 Fez Morocco
[2]Laboratory of Renewable Energies and Intelligent Systems Electrical Engineering department, Faculty of Sciences and Technologies, USMBA, Fez, PO. Box 2202 Fez Morocco
[3]Intelligent System and Application Laboratory, Faculty of Sciences and Technologies, USMBA, Fez, PO. Box 2202 Fez Morocco

E-mail: manal.marzouq@usmba.ac.ma

Abstract. The aim of this work is to develop an artificial neural network (ANN) based model for accurately predicting the daily global solar irradiation in the city of Fez. The potential of the developed model is verified and appraised through the local collected database for the period 2009-2015 from the radiometric station of the Faculty of Sciences and Technology of Fez. The obtained model is MLP with feed forward back-propagation algorithm containing three input parameters and a single hidden layer with nine neurons. Coefficient of determination R^2, the mean absolute percentage error MAPE and the relative root mean square error RRMSE are respectively equal to 97.16%, 21.77% and 18.79%.

1. Introduction

Reliable knowledge of solar irradiation is very important for the design and deployment of solar energy systems [1, 2]. Since it is difficult to obtain measurements of solar irradiation and its components from most meteorological stations, several models have been developed to estimate these necessary data.

ANNs has been shown to be more suitable to predict solar irradiation than other empirical models [3]. In particular, the prediction of daily solar irradiation with ANN's has been developed in the last two decades. Recently, in the work of Chiteka and al. [4], altitude, latitude, longitude, clearness index, average temperature, humidity and pressure have been used as input parameters. Other authors used sunshine duration with commonly used parameters [5, 6]. Some additional inputs were considered for estimation like the amount of suspended particulate matters [7].

In this paper, we try to develop an ANN-based model using the most commonly accessible parameters in meteorological stations: relative humidity, temperature, solar irradiation on the top of the atmosphere and so on. Our second contribution is to study the relevance of the aforementioned input parameters individually and combined. Afterwards, we will evaluate the feasibility of the incremental combination method and test the best model using different number of neurons in the hidden layer.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

IOP Conf. Series: Earth and Environmental Science **161** (2018) 012017 doi:10.1088/1755-1315/161/1/012017

Nomenclature			
G_s	Global solar irradiation (Wh/m2)	$S\Delta T$	Square temperature (°C)
ANN	ANN Artificial Neural Networks	Rh_{max}	Maximum daily relative humidity (%)
MLP	Multi-Layers Perceptron	Rh_{min}	Minimum daily relative humidity (%)
N	Number of hidden neurons	MAPE	Mean absolute percentage error (%)
Rh_{moy}	Mean relative humidity (%)	RRMSE	Relative root mean square error (%)
WS	Wind speed (m/s)	R^2	Coefficient of determination (%)
WD	Wind direction (°)	D_y	Day of the year
Rf	Rainfall (mm)	T_{max}	Maximum daily Temperature (°C)
SA	Solar altitude angle (°)	T_{min}	Minimum daily Temperature (°C)
ΔT	Difference between daily maximum and minimum temperatures (°C)	G_{toa}	Solar irradiation at the top of the top of the atmosphere (Wh/m^2)

This paper is organized as follows: the second section is dedicated to the presentation of the site and the used database followed by our ANN model description. Section 4 deals with the selection of the input parameters. Section 5 and 6 present the selection of the ANN architecture and concluding remarks.

2. Site and Database

In this study, the measurement of global horizontal solar irradiation at ground surface and other meteorological parameters was performed by a radiometric station placed on the roof of the Faculty of Sciences and Technology building at sidi Mohamed Ben Abdellah University, Fez, Morocco (latitude 33° 56' N, longitude 4° 59' W, altitude 579 m). The device collecting the solar data used in this work is a Kip & Zonen model CM-11 pyranometer. For measuring the precipitation, we use the rain gauge while air temperature and relative humidity was measured by means of a thermohygrometer. The wind speed and direction are measured with an anemometer. The measured data between the years 2009 and 2015 were used with 75% for model training and 25% for testing. Before applying our model, a quality control procedure was performed on dataset [8].

3. ANN model description

In this work, a multilayer perceptron (MLP) using Levenberg-Marquardt back-propagation training algorithm with one hidden layer is used for modelling daily values of global solar irradiation on a horizontal surface in the city of Fez.
Hyperbolic tangent and linear functions were used as activation functions in the hidden and output layers respectively. To evaluate the ANN prediction performance, many well-known prediction accuracy indices are adopted in the literature [9]. In this work, we use the coefficient of determination R^2, the mean absolute percentage error MAPE and the relative root mean square error RRMSE defined in [9, 10].

4. Selection of input parameters

In this section, we start by assessing the relevance of each input of our model. Then, we adopt an incremental method by adding progressively the most relevant inputs obtained in the first step in order to show the impact of these combinations on the accuracy of our prediction model. After that, we explore other combinations in order to evaluate the feasibility of the incremental method: the aim is to find the best combinations. In all these experimented models we used an ANN with 5 nodes in the hidden layer. Finally, the best obtained combination is tested using different number of hidden neurons. In this study, we use only one hidden layer which allows achieving several runs for each combination of inputs and hidden neurons within an acceptable time.

Table 1: Individual input parameters performances obtained for the test daily dataset on 10 runs.

Input variables	Performance indicators		
	R^2 (%)	MAPE (%)	RRMSE (%)
G_{toa}	93.01	45.72	28.26
T_{max}	92.72	39.15	28.40
SA	92.45	100.62	24.98
Rh_{min}	91.80	36.63	30.84
ΔT	89.46	39.16	34.85
Rh_{moy}	89.28	94.26	28.15
$S\Delta T$	89.15	40.06	35.40
Rh_{max}	87.57	65.51	36.34
T_{min}	87.05	39.55	32.21
Rf	85.21	54.62	40.12
WS	81.77	72.84	43.67
WD	77.62	92.41	47.82

4.1. Individual input parameters performances
The input parameters selection is the first step in developing our ANN model. The measured and calculated input data are: temperature, minimum temperature, maximum temperature, relative humidity, minimum relative humidity, maximum relative humidity, solar altitude angle, solar irradiation at the top of the atmosphere, rainfall, wind speed, wind direction and day of the year. In the variable selection process, we choose to evaluate individually the average performance of each input variable for 10 runs to avoid random effects because of the initial weight selection and to ensure the reliability of the prediction model [11]. Table 1 summarizes the obtained performance indicators for the test dataset ranked in deceasing order.

4.2. Incremental combination of input parameters
In order to take into account the effect of each input variable evaluated individually in Table 1, we suggest an incremental method by combining progressively the inputs regarding their individual performances. Six combinations of input parameters were developed and dressed in Table 2. From Table 2, it can be noted that the addition of input parameters with the best performances, taken individually, does not necessarily improve the global performance of our model.

Table 2: Average performances of the incremental combinations of input parameters obtained for the daily dataset test on 10 runs.

Combinations	Performance indicators		
	R^2 (%)	MAPE (%)	RRMSE (%)
G_{toa}, T_{max}	95.89	31.45	22.01
G_{toa}, T_{max}, SA	95.65	30.98	22.92
G_{toa}, T_{max}, SA, Rh_{min}	95.99	24.55	22.11
G_{toa}, T_{max}, SA, Rh_{min}, Rf	95.56	24.84	23.42
G_{toa}, T_{max}, SA, Rh_{min}, WS	95.70	30.75	22.89
G_{toa}, T_{max}, SA, Rh_{min}, WD	92.90	29.55	26.25

Moreover, it is clearly shown in this Table that increasing the number of input parameters does not automatically increase the performances of the ANN model: we obtain, for example,

Table 3: Average statistical indicators obtained with various combinations of inputs for the test dataset on 10 runs.

Combinations												Performance indicators		
ΔT	Rh_{moy}	SA	G_{toa}	T_{max}	T_{min}	D_y	Rh_{max}	Rh_{min}	$S\Delta T$	WS	WD	R^2 (%)	MAPE (%)	RRMSE (%)
			X	X								95.69	31.83	22.55
X	X		X									96.77	22.73	19.99
X	X		X			X						96.79	22.31	20.01
X	X	X										96.98	21.22	19.35
	X		X						X			97.02	21.18	19.31
	X	X							X			96.79	22.53	19.93
				X	X	X	X	X				96.68	22.04	20.36
				X	X		X	X				94.57	27.60	25.53
				X	X	X	X	X				96.68	22.04	20.36
		X		X	X							96.46	23.77	21.02
	X		X			X			X			96.64	20.47	22.78
		X	X	X	X							96.37	24.93	21.30
X	X	X									X	96.95	22.98	19.41
X		X	X								X	96.74	22.25	20.05
X	X	X								X	X	96.87	23.69	19.78
X	X	X									X	96.95	22.98	19.41

comparable performances with two and five input parameters presented in the first and third raw respectively.

4.3. Choice of the best combination of input parameters

The remaining of the experimental tests aims to explore other combination possibilities even with input parameters having lower performances as reported in Table 1. Table 3 shows the experimental results for different combinations. As for the other results the average performances are obtained on 10 runs. Table 3 shows that the best combination is obtained using $S\Delta T$, G_{toa} and Rh_{moy} as inputs, with R^2= 97.02%, RRMSE =19.31% and MAPE of 21.18%. Another good combination is given by the inputs (ΔT, SA, Rh_{moy}) with very similar performances, even if the corresponding input parameters are not relevant individually.

5. Selection of the number of neurons in the hidden layer

In this section, the best previously obtained combination is tested with different numbers of neurons in the hidden layer in order to find the best architecture for our model. Table 4 and Figure 1 show the corresponding results.

Table 4: Average statistical indicators obtained with various combinations of inputs for the test dataset on 10 runs.

N	R^2 (%)	MAPE (%)	RRMSE (%)	N	R^2 (%)	MAPE (%)	RRMSE (%)
1	96.41	23.30	21.06	11	96.67	22.28	20.25
2	96.73	21.50	20.17	12	96.67	22.28	20.25
3	96.84	22.53	19.75	13	96.45	24.29	21.00
4	96.81	22.22	19.75	14	96.08	25.56	22.10
5	97.02	21.18	19.31	15	96.02	23.93	22.28
6	96.98	21.32	19.30	16	96.20	23.65	21.86
7	96.93	21.55	19.51	17	96.39	24.48	21.19
8	96.63	23.37	20.47	18	96.31	25.16	21.33
9	97.16	21.77	18.79	19	96.49	22.79	20.94
10	96.51	22.66	20.78	20	96.51	23.43	20.84

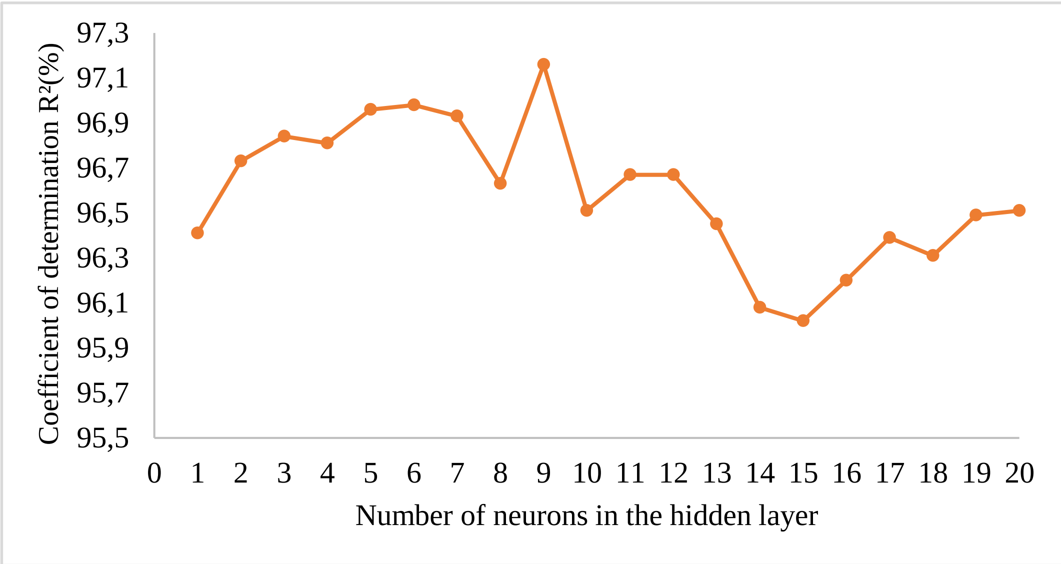

Figure 1: Coefficient of determination of the selected ANN model with different number of neurons in the hidden layer

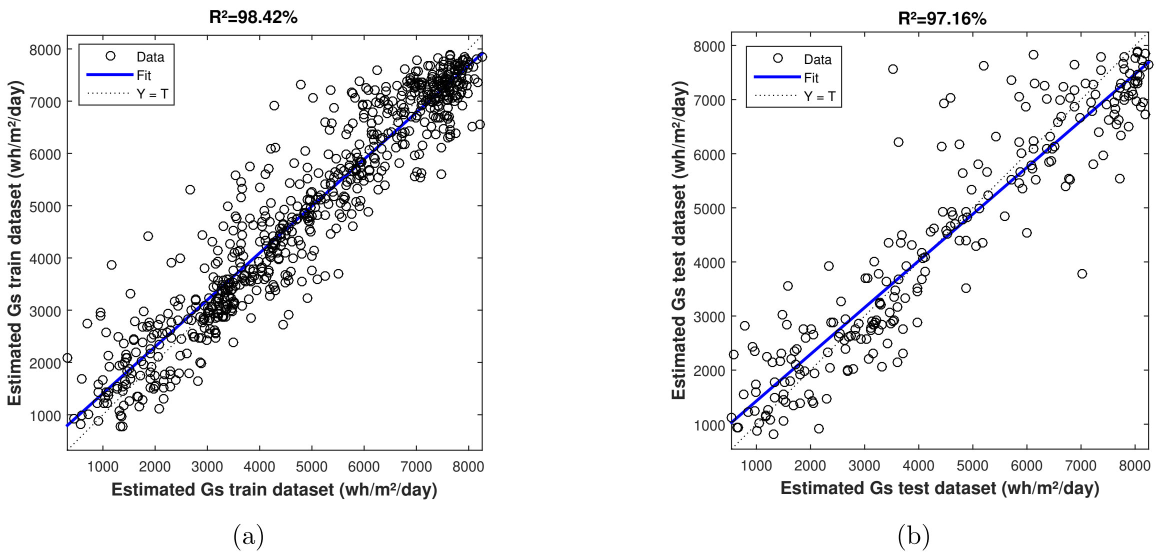

Figure 2: Predicted versus measured daily global irradiation values for our model for (a) training dataset and (b) test dataset.

Figure 2 shows the predicted values of our model with 9 neurons in the hidden layer versus the measured values. It indicates good agreement between measured and estimated values.
The results show that the suitable configuration for our model contains 9 neurons in the hidden layer. This configuration gives as performance 97.16% , 18.79% and 21.77% for R^2 , RRMSE and MAPE respectively.

6. Conclusion

This study shows the results of an effort to predict the daily global solar irradiation according to measured values such as relative humidity, temperature and solar irradiation at the top of the atmosphere which are commonly accessible parameters in the meteorological stations. During this work, we have studied the relevance of the input parameters individually and explored the feasibility of the incremental method. Finally, we have tested the effect of hidden neurons on the performances of our model.

An important outlook for this study is to propose a methodology for automatically determining the relevance of input parameters to enhance the overall performances of the ANN model.

References

[1] Iqbal M, 1979 A study of Canadian diffuse and total solar radiation data-ll Monthly average horizontal radiation *Solar Energy* vol **22** pp 87-90

[2] Sukamongkol Y, Chungpaibulpatana S and Ongsakul W 2002 A simulation model for predicting the performance of a solar photovoltaic system with alternating current loads *Renewable Energy* vol **27** pp 237-258

[3] Reddy K S and Ranjan M 2003 Solar resource estimation using artificial neural networks and comparison with other correlation models *Energy Convers. Manage.* vol **44** pp 2519-30

[4] Chiteka K and Enweremadu C C 2016 Prediction of global horizontal solar irradiance in Zimbabwe using artificial neural networks *Journal of Cleaner Production* vol **135** pp 701–711

[5] Lam J C, Wan K K W and Yang L 2008 Solar radiation modelling using ANNs for different climates in China *Energy Conversion and Management* vol **49** pp 1080-90

[6] Yadav A K, Malik H and Chandel S S 2014 Selection of most relevant input parameters using WEKA for artificial neural network based solar radiation prediction models *Renewable and Sustainable Energy Reviews* vol **31** pp 509-519

[7] Vakilia M, Sabbagh-Yazdi S-R, Kalhorb K and Khosrojerdi S 2015 Using artificial neural networks for prediction of global solar radiation in Tehran considering particulate matter air pollution *Energy Procedia* vol **74** pp 1205-12

[8] Bounoua Z and Mechaqrane A 2017 Quality control of hourly global and diffuse solar irradiation on horizontal surface at Fez city, Morocco (Fez) Proc. du 2ème colloque fronco-marocain sur les Energies Renouvelables et leur intégration aux réseaux de transport et de distribution pp 122-123

[9] Marzouq M, El Fadili H, Lakhliai K and Zenkouar K 2017 A review of solar radiation prediction using artificial neural networks Inter. Conf. on Wireless Technologies, Embedded and Intelligent Systems (Fez) IEEE pp 1-6

[10] Voyant C, Notton G, Kalogirou S, Nivet M L, Paoli C, Motte F and Fouilloy A 2016 Machine Learning methods for solar radiation forecasting: a review *Renewable Energy* vol **105** pp 569-82

[11] Bosch L, López G and Batlles F J 2008 Daily solar irradiation estimation over a mountainous area using artificial neural networks *Renewable Energy* vol **33** pp 1622-28