

# Contribution with global solar radiation forecasting by learning methods in arid climate

Belaid Boualit Sabrina<sup>1</sup>, Mellit Adel<sup>2\*</sup>, Chellali Farouk<sup>1</sup>

<sup>1</sup> *Unité de Recherche Appliquée en Energies Renouvelables, URAER, Centre de Développement des Energies renouvelables, CDER, 47133, Ghardaïa, Algeria*  
sabribelaid@yahoo.fr

<sup>2</sup> *Faculty of Sciences and Technology, Electronics Department, Renewable Energy Laboratory, Jijel University, Jijel 18000, Algeria*

*\*The International Center of Theoretical Physics, ICTP, Trieste, Italy*

**Abstract**— A prior knowledge of Daily Global Solar Radiation (DGSR) is very important for better management and control of various solar installations, mainly photovoltaic and thermal systems. But the complexity of daily behaviour of solar radiation in terms of variability and non-stationary saw its random characteristic leads to consider more robust modelling and forecasting means of this meteorological phenomenon that remains essentially characteristic to the observed region. In this paper, we contribute to develop forecasting models of DGSR in southern region of Algeria (Ghardaia) by Support Vector Machine (SVM) method, for this purpose 42 SVM models are constructed with different possible combinations of measured temperatures (maximum temperature ( $T_{max}$ ), minimum temperature ( $T_{min}$ ) and average temperature ( $T_{mean}$ )) with calculated extraterrestrial radiation ( $H_0$ ) and maximum sunshine duration ( $S_0$ ) as input. Four models are selected for their good forecasting with an NRMSE arranged between 13.163% and 13.305% and a correlation coefficient ( $r$ ) exceeding 89.4%. To prove the effectiveness of the proposed SVM models a comparative study is conducted with neural network models working with the same inputs.

**Keywords**— Forecasting, Daily Global Solar Radiation, Support Vector Machine, Neural Network.

## I. INTRODUCTION

The solar radiation is a very random phenomenon, its modeling and its forecasting has always been an actuality subject and a challenge for the researchers, A literature review shows that many researchers have focused on accurate GSR prediction using artificial neural network ANN [1-4]. Although, the forecast of GSR by SVM technique (Support Vector Machine), which was developed by Vapnik [5] and proved his achievement in computer science, bioinformatics, and environmental science, was applied recently and gives good results.

In 2013, Chen et al [6] presents an application of SVM to estimate DGSR using different combination in input attributes

based on sunshine duration and five empirical sunshine-based models are evaluated using meteorological data at three stations in Liaoning province in China. The SVM models outperform the empirical models. And results good performances with RMSE < 2.4 MJ/m<sup>2</sup> and RRMSE < 18%. Zeng et al [7] propose a least-square support vector machine (LS-SVM) based model for short-term solar power prediction (SPP, One-hour-ahead) in Denver-USA. The input of the model includes atmospheric transmissivity in a novel two-dimensional (2D) form, sky cover, relative humidity (Rh), and wind speed (WS). The output of the model is the predicted atmospheric transmissivity, which then is converted to solar power according to the latitude of the site and the time of the day. The coefficient of correlation  $r$  is 0.9740. In 2014, Ekici [8] developed a Least Squares Support Vector Machine (LS-SVM) based intelligent model to predict the next day's solar insolation in Turkey location with 99.294% accuracy. The prediction model has five inputs; The number of the day from 1st January, daily mean temperature, daily maximum temperature, sunshine duration, and the insolation of the day before. And recently (2015), Mohammadi et al [9] developed an hybrid approach by combining the Support Vector Machine (SVM) with Wavelet Transform (WT) algorithm to predict daily and monthly horizontal global solar radiation in Iran. The different inputs are; relative sunshine duration ( $S/S_0$ ) which is the ratio of sunshine duration ( $S$ ) to the maximum possible sunshine duration ( $S_0$ ), difference between maximum and minimum ambient temperatures ( $T_{max}-T_{min}$ ), relative humidity (Rh), water vapor pressure (VP), average ambient temperature ( $T_{avg}$ ) and extraterrestrial global solar radiation on a horizontal surface ( $H_0$ ). Performance of model gives an MAPE, MABE, RMSE, RRMSE and  $r$  for daily estimation are 6.9996%, 0.8405 MJ/m<sup>2</sup>, 1.4245 MJ/m<sup>2</sup>, 7.9467% and 0.9086, respectively.

Concerning our contribution, we have to develop SVMs models to forecast daily global solar radiation DGSR for one

step ahead by using simple inputs. The obtained results are compared with those given by NN models to prove the effectiveness of SVM.

## II. SUPPORT VECTOR MACHINE (SVM) THEORY

The formulation of SVM employs the Structural Risk Minimization (SRM) principle, which has been shown to be superior to the traditional Empirical Risk Minimization (ERM) principle employed in conventional learning algorithms (e.g. neural networks). This difference makes SVM more attractive in statistical learning applications [10].

Given a set of data points:

$$D = \{(x_i, y_i)\} \in R^d * R, \quad i = 1:n$$

The principle is to find a function  $f$  establishing a relationship between the variables  $x$  and grandeur to modelit  $y$ ;  $y = f(x)$  from the set of measurements  $D$ .

$$f(x) = w\phi(x) + b, \quad et b \in R \quad (1)$$

$\phi(x)$  : is the high dimensional feature space which is nonlinearly mapped from the input space. For more detailed information could be found in (Vapnik 1995 and 1998). The principle is to solve quadratic problem with constraints, so find the Lagrange multipliers  $\alpha_i$  and  $\alpha_i^*$ ,  $i = 1:n$  by minimizing:

$$\begin{aligned} \min L(\alpha_i, \alpha_i^*) \\ = - \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) + \varepsilon \sum_{i=1}^n y_i (\alpha_i + \alpha_i^*) \\ + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i, x_j) \end{aligned} \quad (2)$$

With constraints:

$$0 \leq \alpha_i, \alpha_i^* \leq C, \quad i = 1, \dots, n \quad (3)$$

$$\sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \quad (4)$$

Where  $C$  is a user specified constant and determines the trade-off between the empirical risk and the regularization term.

The regression function is given by:

$$f(x, \hat{\alpha}_i, \hat{\alpha}_i^*) = \sum_{i=1}^n (\hat{\alpha}_i - \hat{\alpha}_i^*) K(x, x_i) + b^* \quad (5)$$

$K(x, x_i) = \phi(x_i) * \phi(x_j)$ : is defined as the kernel function. Hervalue is equal to the scalar product of two vectors  $x_i$  and  $x_j$  in the feature space  $\phi(x_i)$  and  $\phi(x_j)$ .

The elegance of using the kernel function is that one can deal with feature spaces of arbitrary dimensionality without having to compute the map  $\phi(x)$  explicitly. Any function satisfying Mercer's condition can be used as kernel function [11]. Use of RBF function is recommended for several reasons, among other because RBF handles the case where the relationship between the labels and attributes is nonlinear.

Fig 1 Relationship between DGSR and

To develop these models, we have exploited the measures taken between February 2012 and February 2015 at Applied Research Unit for Renewable Energies (ARURE)Ghardaia. Situated in southern region of Algeria, of which latitude:  $+32.37^\circ$ , longitude:  $+3.77^\circ$ , and altitude: 450 m above the mean sea level. This site is characterized by semi-arid to arid climate. Two years are chosen for training the SVM model and one year is reserved to the test. According to Zhao et al [12], SVM are highly effective models in solving non-linear problems even with small quantities of training data.

We will proceed to forecast daily global solar radiation DGSR of the following day ( $D + 1$ ) using as inputs to the model SVM, different daily temperature of D-Day such as; maximum temperature ( $T_{max}$ ), minimum temperature ( $T_{min}$ ),  $T_{max}-T_{min}$ , average temperature ( $T_{mean}$ ), extraterrestrial radiation ( $H_0$ ) and maximum sunshine duration ( $S_0$ ) in situ.

To justify our choice of inputs, we illustrate in Fig. 1 the relationship between DGSR and each of inputs used by calculating their correlation coefficients  $r$ .

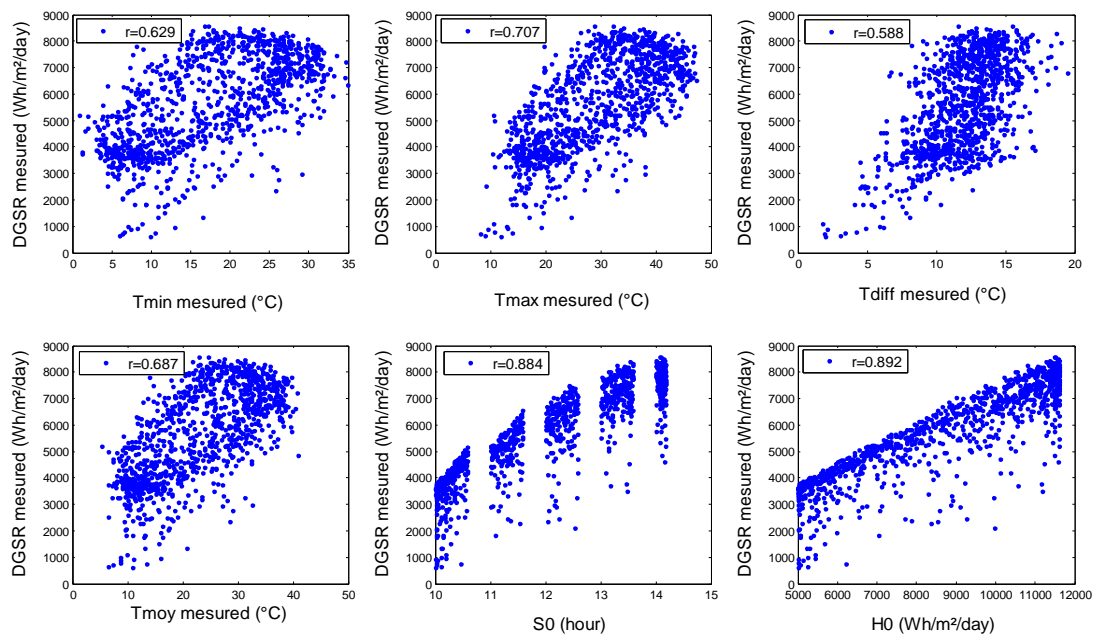


TABLE 1.  
 THE VARIOUS INPUTS USED TO DEVELOP SVM MODELS

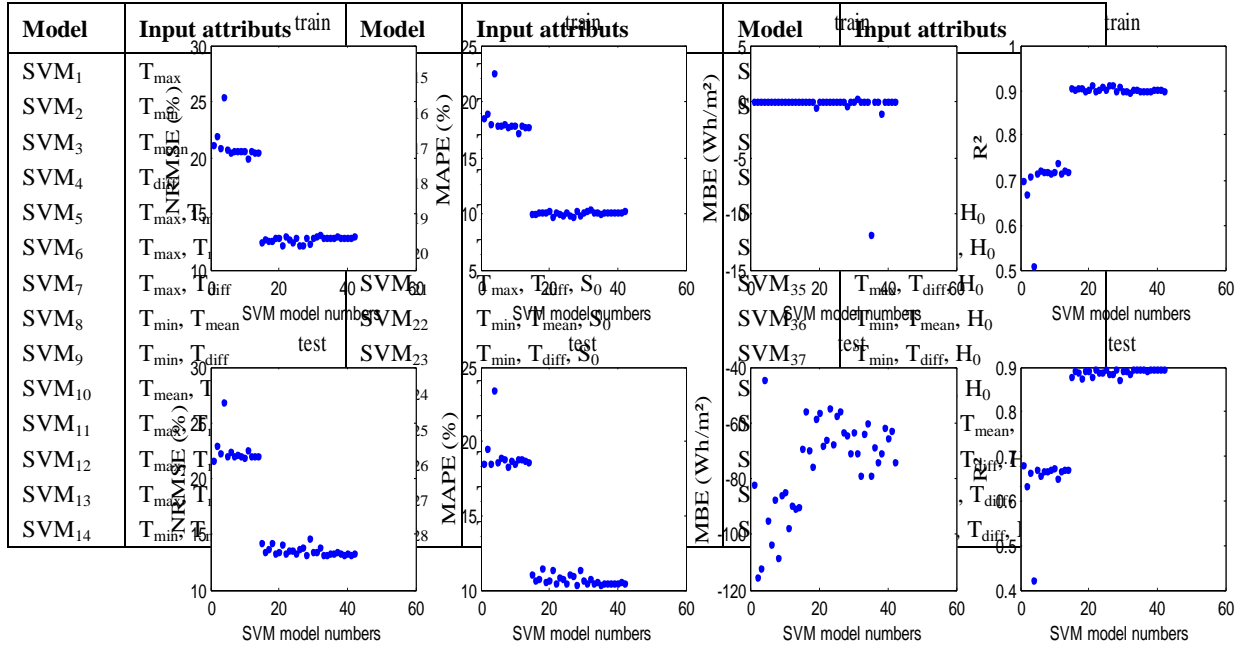


TABLE 1. shows the different possible combinations. To simplify the writing, we put:  $T_{max}-T_{min}=T_{diff}$

To interpret the results of training and forecasting by different models (SVM<sub>1</sub> to SVM<sub>42</sub>), we will calculate some performance tests; NRMSE, RMSE, MAPE, MBE and the correlation coefficient  $r$ . And we chose the plot of representative curves of these criteria that will allow us better reading (see Fig. 2.) and a better selection of models.

#### IV. DISCUSSION

SVM<sub>1</sub>toSVM<sub>14</sub>modelswhoseinputsareonlytemperatures give not goodresults, NRMSEvariesbetween20.095 % 26.844% and  $r$  between 0.421 to 0.740. by against, SVM<sub>15</sub>to SVM<sub>42</sub>modelswhom weintroduced eitherS<sub>0</sub> orH<sub>0</sub>, the results are betterand approachingthem, butwe manage todistinguish fourmodelswhose performances arebettercompared to theother with respect tothe prediction,which we summarizein table 2.

Fig 2Performance criteriaof the differentSVMmodels developedforDGSR

TABLE 2  
 PERFORMANCE RESULTS OF THE FOUR SELECTED MODELS

Inputs			RMSE (Wh/m <sup>2</sup> )	RMSE (Cal/cm <sup>2</sup> )	RMSE (Mj/m <sup>2</sup> )	NRMSE (%)	MAPE (%)	MPE (Wh/m <sup>2</sup> )	r
SVM <sub>25</sub>	T <sub>max</sub> , T <sub>min</sub> , T <sub>mean</sub> , S <sub>0</sub>	Test	757.376	65.123	2.727	12.740	10.181	29.087	0.900
		Train	777.346	66.840	2.798	13.266	10.503	-57.599	0.894
SVM <sub>28</sub>	T <sub>min</sub> , T <sub>mean</sub> , T <sub>diff</sub> , S <sub>0</sub>	Test	762.664	65.577	2.746	12.829	10.293	31.988	0.900
		Train	771.301	66.320	2.777	13.163	10.403	-64.432	0.896
SVM <sub>39</sub>	T <sub>max</sub> , T <sub>min</sub> , T <sub>mean</sub> , H <sub>0</sub>	Test	757.479	65.131	2.727	12.742	10.137	28.027	0.901
		Train	771.815	66.364	2.779	13.172	10.458	-61.500	0.896
SVM <sub>42</sub>	T <sub>min</sub> , T <sub>mean</sub> , T <sub>diff</sub> , H <sub>0</sub>	Test	765.401	65.813	2.755	12.875	10.058	19.156	0.898
		Train	779.625	67.036	2.807	13.305	10.440	-74.242	0.894

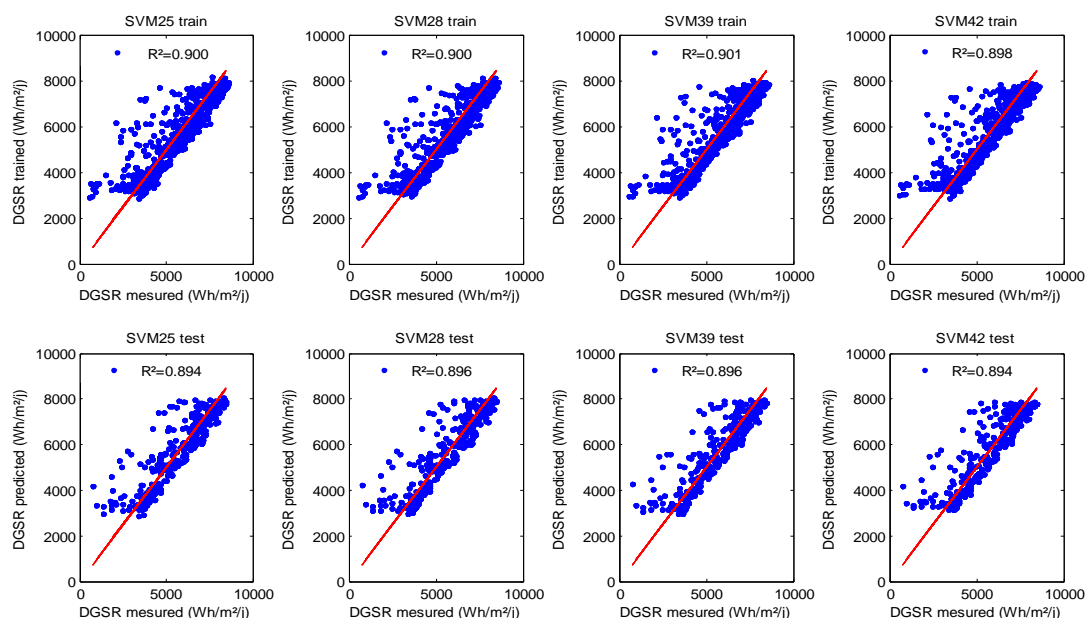


Fig 3DGSr predicted and estimated based on the measured DGSr

Fig.3 Represents the estimation and forecasting DGSr through selected models

#### V. COMPARISON WITH NN MODEL

Saw that neural networks (NN) have proved their achievement in the field of machine learning and prediction of solar radiation [15-21], so we have developed MLP (Multi Layer Perceptron) models with the same inputs as those of SVM models selected (SVM<sub>25</sub>, SVM<sub>28</sub>, SVM<sub>39</sub>, SVM<sub>42</sub>), then we compare them to prove the effectiveness of SVM.

Fig. 4. and 5. show the comparison between the correlation coefficients r of the four SVM models selected and those developed by MLP, respectively for training and forecasting. We present on the x-axis the selected inputs (ex: inputs 25 refer to the inputs used in the model SVM<sub>25</sub>). From these figures, we see that we get almost the same results for training, while for prediction, SVM models perform better.

#### VI. CONCLUSION

This work has as objective the forecasting of DGSr for one step ahead by SVM method from many combinations of simple inputs.

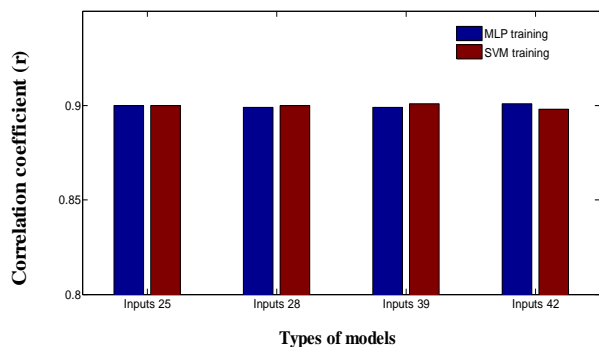


Fig 4 Comparison of SVM models to MLP models over training

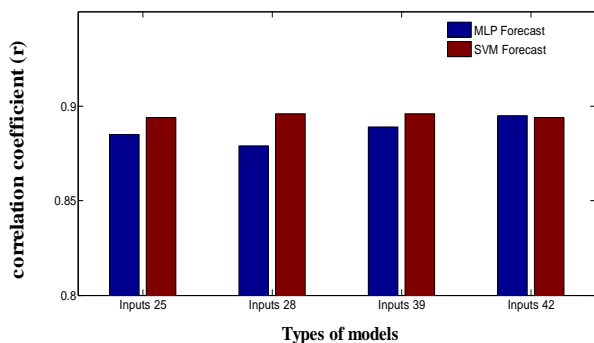


Fig 5 Comparison of SVM models to MLP models over forecasting.

Four models are selected for forecasting of DGSR with a correlation coefficient exceeding 89.4%, this models consider the three types of temperatures ( $T_{max}$ ,  $T_{min}$ , and  $T_{mean}$ ) with either  $H_0$  or  $S_0$ . Their performances gave NRMSE between 13.163% and 13.305%, MAPE of 10.403% to 10.503% and  $r$  varies of 0.894 to 0.896. To prove the effectiveness of the proposed SVM models, a comparative study is conducted with neural network models working with the same inputs. We have noted that SVM models give better results compared to NN models. Finally, it can be concluded that the SVM is a reliable technique that can be easily operated whatsoever for estimation or forecasting processes with behavior as random as that of the solar radiation.

#### REFERENCES

[1] M. Benghanem, A. Mellit, and S. N. Alamri, "ANN-based modelling and estimation of daily global solar radiation data: A case study," *Energy Conversion and Management*, vol. 50, pp. 1644-1655, 2009.

[2] A. Mellit, M. Benghanem, A. H. Arab, and A. Guessoum, "A simplified model for generating sequences of global solar radiation data for isolated sites: Using artificial neural network and a library of Markov transition matrices approach," *Solar Energy*, vol. 79, pp. 469-482, 2005.

[3] A. Mellit and A. M. Pavan, "A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy," *Solar Energy*, vol. 84, pp. 807-821, 2010.

[4] R. Yacef, M. Benghanem, and A. Mellit, "Prediction of daily global solar irradiation data using Bayesian neural network: A comparative study," *Renewable Energy*, vol. 48, pp. 146-154, 2012.

[5] V. Vapnik, *The nature of statistical learning theory* Verlag .new york: Springer, 1995.

[6] J.-L. Chen, G.-S. Li, and S.-J. Wu, "Assessing the potential of support vector machine for estimating daily solar radiation using sunshine duration," *Energy Conversion and Management*, vol. 75, pp. 311-318, 2013.

[7] J. Zeng and W. Qiao, "Short-term solar power prediction using a support vector machine," *Renewable Energy*, vol. 52, pp. 118-127, 2013.

[8] B. B. Ekici, "A least squares support vector machine model for prediction of the next day solar insolation for effective use of PV systems," *Measurement*, vol. 50, pp. 255-262, 2014.

[9] K. Mohammadi, S. Shamshirband, C. W. Tong, M. Arif, D. Petković, and S. Ch, "A new hybrid support vector machine-wavelet transform approach for estimation of horizontal global solar radiation," *Energy Conversion and Management*, vol. 92, pp. 162-171, 2015.

[10] A. A. F. Refaat M Mohamed, "Classification of Multispectral Data Using Support Vector Machines Approach for Density Estimation," presented at the IEEE Seventh International Conference on Intelligent Engineering Systems, Assiut, Egypt, 2003.

[11] J.-L. Chen, G.-S. Li, B.-B. Xiao, Z.-F. Wen, M.-Q. Lv, C.-D. Chen, Y. Jiang, X.-X. Wang, and S.-J. Wu, "Assessing the transferability of support vector machine model for estimation of global solar radiation from air temperature," *Energy Conversion and Management*, vol. 89, pp. 318-329, 2015.

[12] H.-x. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 16, pp. 3586-3592, 2012.

[13] T. Muneer, *Solar Radiation and Daylight models*, 2004 ed. Elsevier Ltd, 2004.

[14] K. Mohammadi, S. Shamshirband, M. H. Anisi, K. A. Alam, and D. Petković, "Support vector regression based prediction of global solar radiation on a horizontal surface," *Energy Conversion and Management*, vol. 91, pp. 433-441, 2015.

[15] S. M. Al-Alawi and H. A. Al-Hinai, "An ANN-based approach for predicting global radiation in locations with no direct measurement instrumentation," *Renewable Energy*, vol. 14, pp. 199-204, 1998.

[16] M. A. Behrang, E. Assareh, A. Ghanbarzadeh, and A. R. Noghrehabadi, "The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data," *Solar Energy*, vol. 84, pp. 1468-1480, 2010.

[17] A. Rahimikhoob, "Estimating global solar radiation using artificial neural network and air temperature data in a semi-arid environment," *Renewable Energy*, vol. 35, pp. 2131-2135, 2010.

[18] A. Linares-Rodríguez, J. A. Ruiz-Arias, D. Pozo-Vázquez, and J. Tovar-Pescador, "Generation of synthetic daily global solar radiation data based on ERA-Interim reanalysis and artificial neural networks," *Energy*, vol. 36, pp. 5356-5365, 2011.

[19] A. Moghaddamnia, R. Remesan, M. H. Kashani, M. Mohammadi, D. Han, and J. Piri, "Comparison of LLR, MLP, Elman, NNARX and ANFIS Models—with a case study in solar radiation estimation," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 71, pp. 975-982, 2009.

[20] S. Rehman and M. Mohandes, "Artificial neural network estimation of global solar radiation using air temperature and relative humidity," *Energy Policy*, vol. 36, pp. 571-576, 2008.

M. Benghanem and A. Mellit, "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*, vol. 35, pp. 3751-3762, 2010.