

Solar Radiation Prediction for Dimensioning Photovoltaic Systems Using Artificial Neural Networks

E. Noriega-Angarita^{#1}, V. Sousa-Santos^{#2}, M. Quintero-Duran^{#3}, C. Gil-Arrieta^{#4}

[#] Faculty of Engineering, Universidad de la Costa
Calle 57 No. 55-66, Barranquilla, Colombia

¹ enoriega2@cuc.edu.co

² vsousa1@cuc.edu.co

³ mquinter5@cuc.edu.co

⁴ cesargil2012@gmail.com

Abstract— this paper presents a prediction model of solar radiation for dimensioning photovoltaic generation systems in the Atlantic Coast of Colombia, using artificial neural networks. As a case of study is presented the municipality “El Carmen de Bolivar” located in this region. To obtain the model, the average data of daily temperature, relative humidity and solar radiation from the last ten years, reported by weather stations in this city were used. Six neural networks were designed with six variants of input variables (temperature, humidity and month) and the output variable (solar radiation). The best result was obtained using all input variables. In the training process, the correlation index (R) between solar radiation estimated by the model and the recorded data was 0.8. In validating the correlation index was 0.77.

Keyword- Modelling, prediction, solar radiation, artificial neural networks, photovoltaic generation systems

I. INTRODUCTION

The world energy scenario has been changing in recent years. Countries that for decades were energy resources importers, are now exporters, and exporters are emerging as the fastest growing in demand. On the other hand, there are very few long-term solutions to global energy challenges. There is a new focus on energy efficiency, but CO₂ emissions keep rising. In addition, there are 1.3 billion people with no electricity and 2.6 billion lack clean cooking facilities [1].

Searching solutions, Colombia approved the law 1715 from 2014, which encourages the use of non-conventional renewable energy. Within non-conventional renewable sources, solar energy has a high energy potential in the Atlantic Coast of Colombia [2], so it is necessary to get the required information, solar radiation mainly, for developing projects with photovoltaic installations.

Solar radiation is one of the most important parameters for dimensioning photovoltaic farms. This parameter can be directly measured or estimated, using statistical models or artificial intelligence techniques such as Artificial Neural Networks (ANN), considering meteorological variables that have been previously measured in the specific place where installation is estimated.

Several authors have addressed the issue related to the study and statistical estimation of time series of solar radiation, where the most recognized approach is the Angstrom model [3].

In [4] a statistical model that allows selecting the best value of solar constant was developed in order to estimate the daily monthly global solar radiation average, with better results than the Angstrom model. In [5], linear, quadratic, logarithmic and exponential models were applied to estimate global solar radiation, obtaining good results. An empirical model for estimating daily global solar radiation per year on a horizontal surface is proposed in [6]. The model was validated in nine representative stations in China and it was demonstrated that adequately respond to highly variable weather conditions. In [7], the authors investigate several models for estimating solar radiation, showing that the quality of models is very dependent on the location where the station, in which measurements are made, is installed.

A model of dynamic harmonic regression is proposed in [8] that allows obtaining good results for a 24 hours forecast. In [9], three methods based on daily measurements of maximum and minimum temperature at meteorological stations, cloud cover and extra-terrestrial global radiation were developed. In [10], 54 models based on the bandwidths procedure for calculating the global radiation and diffuse over horizontal surface are evaluated. In [11] a model for estimating radiation is developed, which considers the difference between the

daily maximum and minimum air temperature and saturation vapor pressures at the maximum and minimum temperature, providing good results.

In [2] a work developed in Barranquilla, Colombia is reported. Several statistical models are analyzed to determine the meteorological variables that are correlated with solar radiation. The analyzed variables are temperature, relative humidity and sunshine hours. As a result it was found that the best correlation is obtained with the relative humidity, the rest of the variables have a virtually constant behavior throughout the year.

Solar radiation can also be estimated by using artificial neural networks (ANN). This artificial intelligence technique has been used in many engineering applications [12], [13]. Its application to the prediction of solar radiation is based on historical data collection, and with network training processes, allowing predictions for the design of photovoltaic generation systems. Scientific literature reports the development of several models of ANN, and in most cases the superiority of the use of this artificial intelligence technique is shown in relation to statistical methods.

In [14] several models of solar radiation estimation reported in literature that uses ANN were applied. The results show that the accuracy of outcomes depends on the combination of input parameters, training algorithms and architecture configuration. In [15] a prediction model of global solar radiation distribution on horizontal surfaces is presented. The designed model processes the correlation between weather conditions, the duration of daylight and peak value of radiation. As in [7], it is shown that the effectiveness of the results is highly dependent on the location of the weather station.

In [16] is shown that using temperature and humidity data as input, it can be obtained good results in the prediction of global solar radiation (GSR). In [17] is described the development of a model for estimating daily monthly diffuse solar radiation data of 9 stations having different climatic conditions distributed in different geographic locations in China using ANN. The results of this study demonstrated the model generalization ability in the rest of China. In [18] a model of neural network is presented using the features of interrelation of direct, diffuse and global solar radiation in India, developing an algorithm that includes the estimation of these radiations through clear sky conditions, and deviations due to random weather events.

In [19] six ANN models were developed for the estimation and modeling of daily global solar radiation. The input data were global radiation, diffuse radiation, air temperature and relative humidity. It was shown that results can be obtained with good accuracy using the sunshine duration and air temperature. In [20] an ANN is developed to estimate the hourly solar radiation in the province of Cordoba, Argentina. This model uses as input data the temperature, relative humidity, wind speed and rain.

There are several works in Colombia that use ANN for prediction of meteorological variables. However, in [21], [22] solar radiation is not studied, while in [23] a model for predicting solar radiation is proposed, but applied in a location outside Colombia.

This work proposes to implement an ANN model for obtaining daily solar radiation prediction in El Carmen de Bolivar, belonging to the Colombian Atlantic coast. In the proposed model, it is considered as input variables the daily average data of temperature, relative humidity and solar radiation from the last ten years, reported by weather stations in the city. As variable input is also used the month of the year. Before the neural networks design, an analysis of variability of studied parameters and a correlation between them are presented.

The contribution of this paper is to present a forecast analysis in a region where it has never done such a study, which constitutes a starting point for the projection of photovoltaic power plants in Colombia. The remainder of the paper is as follows. Material and methods are specified in section II. Section II shows the results got from the ANN process and their analysis. Finally, some conclusions are mentioned.

II. MATERIALS AND METHODS

The study is divided into two main stages. In the first stage a statistical analysis about the behavior of the used variables is achieved, and the correlation between these is done. The second step is the development of six neural networks, considering equal number of variants of combinations from input variables (temperature, humidity and month) and the output variable (solar radiation). To select the most appropriate network, behavioral analysis is performed during the training phase and validation.

A. Stage 1

In stage1, the following steps are executed:

- Behavior analysis of the average for daily temperature, humidity and solar radiation per month.
- Statistical parameters determination of the average for daily temperature, humidity and radiation per month.
- Behavior analysis of the average for daily temperature, humidity and solar radiation.
- Statistical parameters determination of the average for daily temperature, humidity and radiation.
- Correlation index (R) analysis between radiation vs temperature, radiation vs humidity, and humidity vs temperature.

B. Stage 2

Six neural networks were developed in this step with input and output combinations shown in Table 1. For the development of networks, the "nntool" graphical interface from Matlab program was used.

TABLE I
Input combinations and output of the proposed neural networks

Network	Inputs	Output
1	Temperature (°C)	Radiation
2	Month Temperature (°C)	Radiation
3	Humidity (%)	Radiation
4	Month Humidity (%)	Radiation
5	Temperature (°C) Humidity (%)	Radiation
6	Month Temperature (°C) Humidity (%)	Radiation

For all networks the same properties were used, in order to have a reference for comparing the performance of all networks and select the most appropriate. The applied properties were obtained from a sensitivity analysis and addressed criteria reported in specialized literature [15], [18], [24]. The properties of networks are presented in table 2.

TABLE III
ANN properties used

Network type	"FEED-FOWARD BACKPROP"
Training function	TRAINLM
Adaptive learning function	LEARNGDM
Performance function	MSE
Number of layers	2
Number of neurons in layer 1	1000
Transfer function in Layer 1	LOGSIG
Number of neurons in layer 2	1
Transfer function in Layer 2	TANSIG

The developed neural network architecture neural network is shown in Fig. (1). the difference is the number of input variables. Networks 1 and 3 have an input variable. Networks 2, 4 and 5 have two input variables, while the network 6 has three input variables.

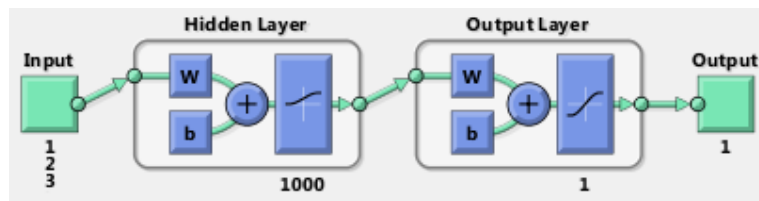


Fig. 1. Architecture of developed neural networks

Neural networks design was conducted in two phases; training and validation. In training phase, 70% of data were used, meanwhile the validation phase was carried out with 30% of the remaining data. This data selection method was applied in [24].

For the selection of the most appropriate network, were taken as criteria the minimum square error (mse) determined during the training process of each network and the correlation index (R) between the radiation values obtained by the model of networks and the real value. These criteria were applied in [16] and [19].

III. RESULTS AND DISCUSSION

"El Carmen de Bolivar" is a city located inside the Department of Bolivar in the Colombian Atlantic Coast. It is at 114 km southeast of Cartagena de Indias. In this study, the daily average data of temperature, relative humidity and solar radiation from the last ten years were used. These data were collected from a weather station close to where the installation of photovoltaic power plant is projected. The weather station equipment is certified by the requirements of the Institute of Hydrology, Meteorology and Environmental Studies of

Colombia (IDEAM). Thus it meets the conditions suggested in [7] and [14], which recommends prediction studies near the measuring points of the meteorological variables.

The results of the two stages above are presented.

A. Stage 1

In Fig. (2), the average variation of daily temperature, humidity and solar radiation per month is shown. This average was determined per month as the sum of each daily parameter, divided by the number of days in month. Table 3 shows the statistical parameters of minimum, maximum, their average and standard deviation corresponding to graphics.

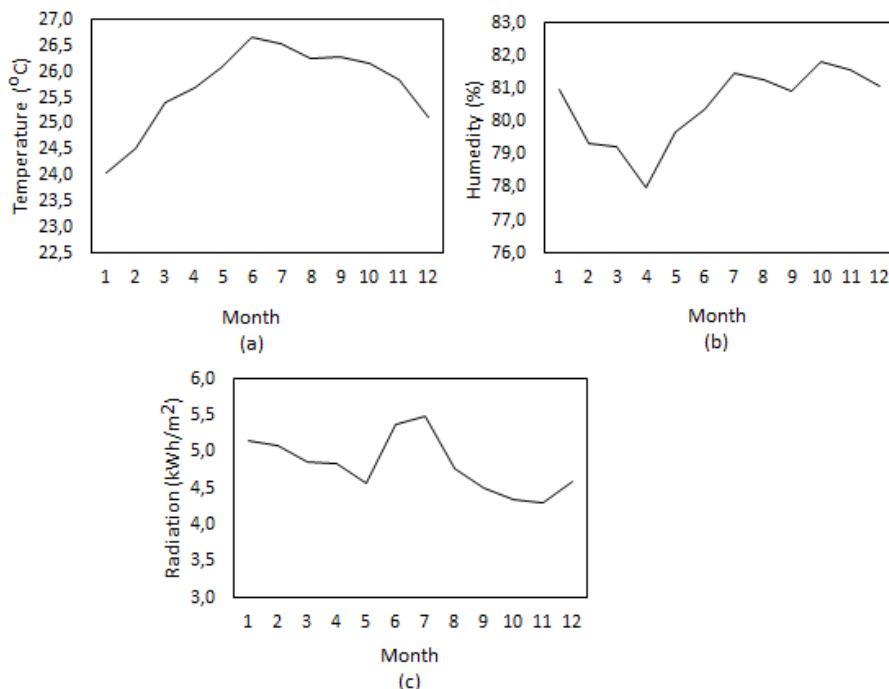


Fig. 2. Average behavior for daily temperature (a), humidity (b) and radiation (c) per month

TABLE IIIII
Statistical parameters of the average for daily temperature, humidity and radiation per month

Variable	Minimum	Maximum	Average	Standard deviation
Temperature (°C)	24.04	26.66	25.71	0.81
Humidity (%)	77.97	81.80	80.46	1.17
Radiation (kWh/m ²)	4.31	5.48	4.83	0.38

It can be seen from table 3 according to the standard deviation values, that the parameters are not very variable, having the humidity the greater variation, this is also evident in the study reported in [2] for the city of Barranquilla, belonging to the same region. Nevertheless, although generally temperatures are high in Fig. (2-a), it is clear that the greatest records are reported from June to August, while December to February temperatures are lower. Regarding humidity, Fig. (2-b) shows that from February to May the lowest values are recorded, while in the remaining months values remain almost constant. According to Fig. (2-c), the highest solar radiation is registered in both June and July, while in November is lower.

Due to the observed influence of month in the behavior of meteorological variables, it is decided to incorporate it as one of the input variables in the development of neural networks.

In Fig. (3) the daily average record of the variables under study from the last ten years are shown. Table 4 shows the corresponding statistical parameters. The solar radiation annual average is 1371.5 kWh/m², which represents a high potential of solar energy [11].

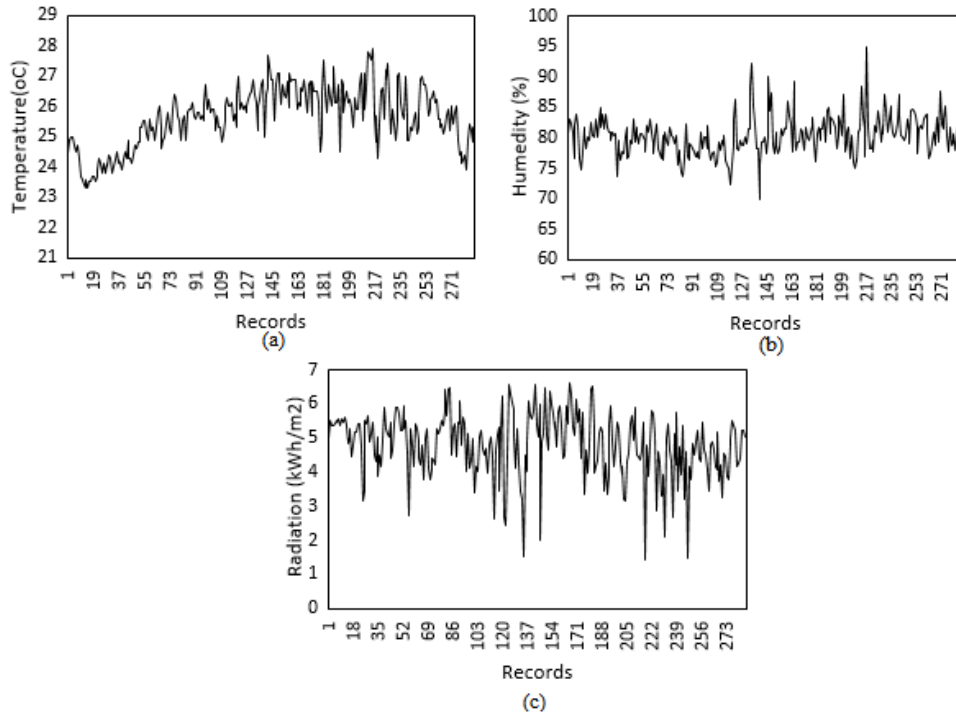


Fig. 3. Average behavior for daily temperature (a), humidity (b) and radiation (c)

Table 4 shows similarly to the previous analysis of the behavior per month, that humidity is the most variable parameter. In Fig. (3) an instability in the values are observed at three cases. In the case of temperature, a defined pattern by the months of higher and lower temperature indicated above is observed. In the case of both humidity and radiation, it is not possible to clearly define a pattern in the variation.

TABLE IVV
Statistical parameters of the average for daily temperature, humidity and radiation per month

Variable	Minimum	Maximum	Average	Standard deviation
Temperature (°C)	23.30	27.90	25.62	0.98
Humidity (%)	70.00	95.00	80.36	3.10
Radiation (kWh/m ²)	1.45	6.64	4.81	0.91

In order to identify whether there is relationship between the behavior of meteorological variables under study, a correlation analysis is performed between these. In Fig. (4) the correlation between radiation vs temperature, radiation vs. humidity, and humidity vs temperature are shown.

In Fig. 4 the low values observed from correlation coefficient (R) between the radiation vs temperature (0.004), radiation vs humidity (-0.4), and humidity vs temperature (0.1) shows that there is no simple relationship between these variables. This performance is also evident in [19]. For this reason, it is useful the artificial neural networks application as a tool to obtain a model for estimating solar radiation from other meteorological variables, such as humidity and temperature, also including the month of the year.

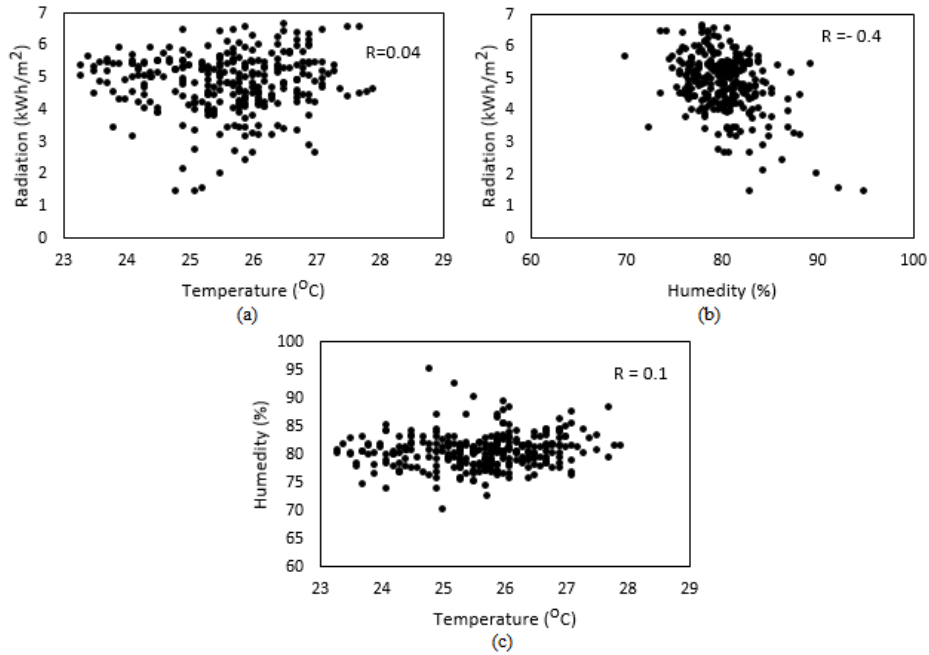


Fig. 4. Correlation between radiation vs. temperature (a), radiation vs. humidity (b), and humidity vs. temperature (c)

B. Stage 2

Results concerning the training and validation phases of neural networks are shown.

Training phase results:

In Fig. (5) the graphics of radiation behavior obtained from neural networks models and the radiation recorded corresponding to the training stage are shown. The dispersion plots between these variables and the corresponding correlation (R) index are presented in Fig. (6).

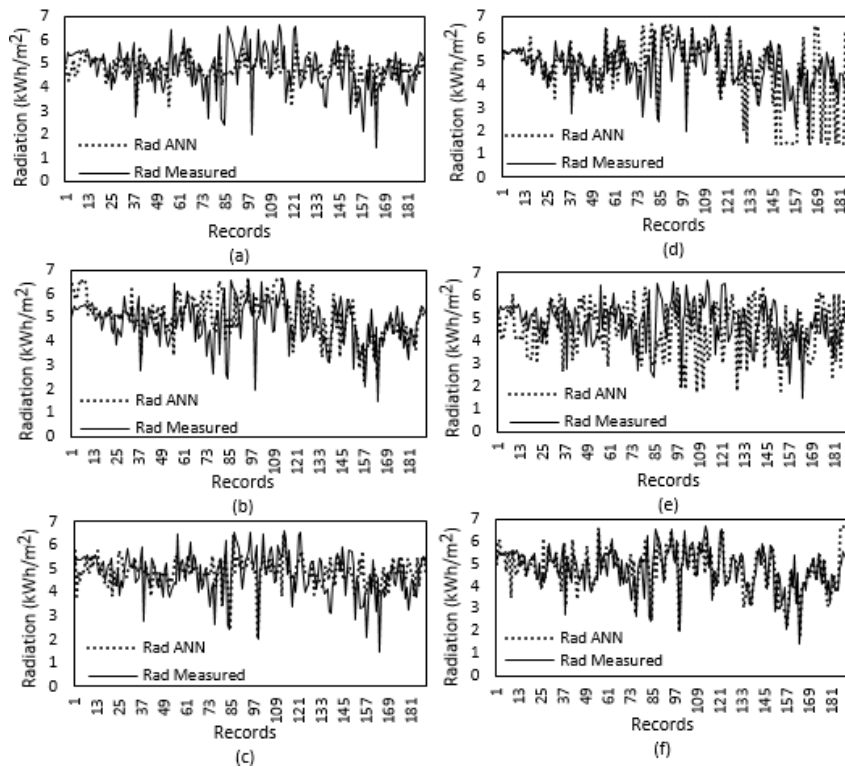


Fig. 5. Radiation obtained from the neural network (Rad ANN) model the and associated records (Rad Measured), (a) network 1, (b) network 2, (c) network 3 (d) network 4 (e) network 5 (f) network 6

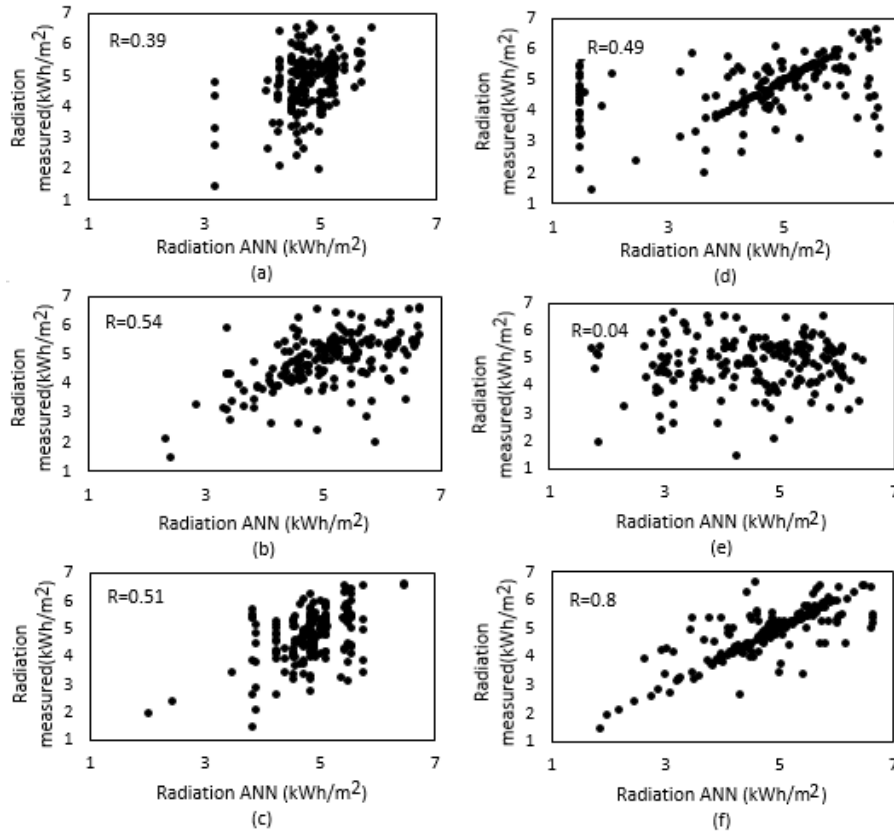


Fig. 6. Dispersion plot and correlation index (R) of radiation obtained from the neural network (Rad ANN) model and corresponding records (Rad Rec), (a) network 1, (b) network 2, (c) network 3, (d) network 4, (e) network 5, and (f) network 6

Table 5 shows the minimum square error (mse) obtained during the training stage per network and the correlation index (R) between the radiation obtained from the neural network model and the recorded radiation sample.

TABLE V
Minimum square error (mse) and index correlation (R) per network

Network	Inputs	Output	mse	Correlation index (R)
1	Temperature (°C)	Radiation	0.0117	0.39
2	Month Temperature (°C)		0.00471	0.54
3	Humidity (%)		0.0101	0.51
4	Month Humidity (%)		0.0196	0.49
5	Temperature (°C) Humidity (%)		0.000272	0.04
6	Month Temperature (°C) Humidity (%)		0.000203	0.8

From Fig. (5), Fig. (6) and table 5, it can be concluded that the best trained network, and therefore, which provides the best results is the network 6, with a correlation index of 0.8 and a minimum square error of 0.000203. These results are considered good, considering that no data was filtered. If network 5 with network 6 are compared, it is evident that considering the month in the model provided better results.

Results of the validation phase for network 6:

The validation phase of the network was carried out with the rest of records. In Fig. (7) the graphics about the radiation obtained from the neural networks model and recorded radiation are shown. Correlation index (R) is also observed.

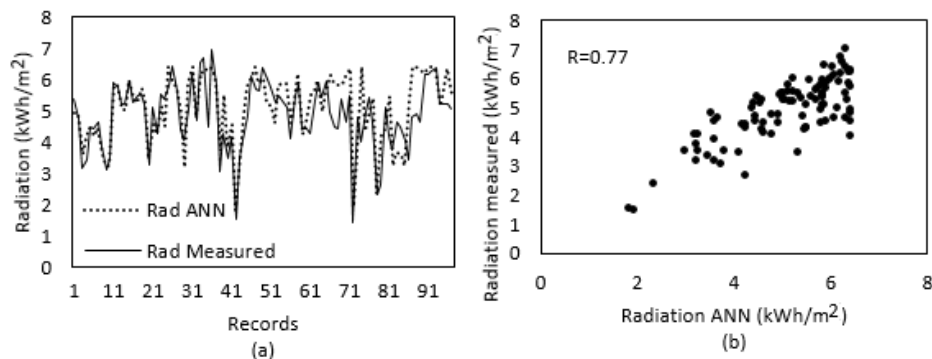


Fig. 7. (a) radiation obtained from the neural network (Rad ANN) 6 model and corresponding records (Rad Rec), (b) correlation index (R) between the two variables

As shown in Fig. (7), network 6 provides good results for the rest of the data that were not used in the training process, which validate applicability. This way the model can be used for dimensioning photovoltaic systems and analyze possible generation scenarios.

IV. CONCLUSIONS

Photovoltaic power generation is an alternative with great potential in Colombia, as a means of supplementing energy deficits due to natural phenomena such as "El Niño", which significantly affects heavily hydropower dependent systems.

This work shows the great potential offered by the Atlantic coast, where one of its areas, the city of "El Carmen de Bolívar", has a capacity of 1371.5 kWh/m² per year and 4.81 kWh/m² daily average.

In the evaluated models, was evident the importance of considering the month as an element that complements the use of meteorological variables, and thus improves the quality of the obtained model.

The obtained model with a correlation index of 0.8 during training and 0.7 in the validation, allows dimensioning photovoltaic systems and predict with good accuracy, the ability to generate daily, monthly and annually power to different scenarios of temperature and humidity.

The results show that it is possible to generalize the tool used in this work to tropical climates areas, such as the Colombian Atlantic coast, in which the installation of photovoltaic systems is projected. This is further facilitated by the requirement of few weather variables and the similarity in climate characteristics.

ACKNOWLEDGMENT

The authors thank the Colombian Administrative Department of Science and Technology – Colciencias as financing entity of the project "Development and implementation of a comprehensive program evaluation and utilization of wind and solar resource in the Colombian Caribbean region with emphasis on La Guajira", which this work is part of it. We thank the teamwork of the Atlantic University, Bolívar Technology University, the University of La Guajira and IDEAM on its Barranquilla headquarter for providing the information used in this work.

REFERENCES

- [1] M. Van der Hoeven, *World Energy Outlook 2014*. Paris, 2014.
- [2] A. Vélez-Pereira, E. Vásquez, W. Coronell, and Et.al, "Determinación de un modelo paramétrico para estimar la radiación solar," *Ingenium*, vol. 7, no. 18, pp. 11–17, 2013.
- [3] A. Angstrom, "Solar and terrestrial radiation," *J. R. Meteorol. Soc.*, vol. 50, no. 210, pp. 121–126, 1924.
- [4] H. Li, Y. Lian, X. Wang, W. Ma, and L. Zhao, "Solar constant values for estimating solar radiation," *Energy*, vol. 36, no. 3, pp. 1785–1789, Mar. 2011.
- [5] M. Yorukoglu and A. N. Celik, "A critical review on the estimation of daily global solar radiation from sunshine duration," *Energy Convers. Manag.*, vol. 47, no. 15–16, pp. 2441–2450, Sep. 2006.
- [6] H. Li, W. Ma, Y. Lian, and X. Wang, "Estimating daily global solar radiation by day of year in China," *Appl. Energy*, vol. 87, no. 10, pp. 3011–3017, Oct. 2010.
- [7] Z. Jin, W. Yezheng, and Y. Gang, "General formula for estimation of monthly average daily global solar radiation in China," *Energy Convers. Manag.*, vol. 46, no. 2, pp. 257–268, Jan. 2005.
- [8] J. R. Trapero, N. Kourentzes, and A. Martín, "Short-term solar irradiation forecasting based on Dynamic Harmonic Regression," *Energy*, vol. 84, pp. 289–295, May 2015.
- [9] M. El-Metwally, "Simple new methods to estimate global solar radiation based on meteorological data in Egypt," *Atmos. Res.*, vol. 69, no. 3–4, pp. 217–239, Jan. 2004.
- [10] V. Badescu, C. A. Gueymard, S. Cheval, C. Oprea, M. Baciú, A. Dumitrescu, F. Iacobescu, I. Milos, and C. Rada, "Computing global and diffuse solar hourly irradiation on clear sky. Review and testing of 54 models," *Renew. Sustain. Energy Rev.*, vol. 16, no. 3, pp. 1636–1656, Apr. 2012.
- [11] J. Almorox, C. Hontoria, and M. Benito, "Models for obtaining daily global solar radiation with measured air temperature data in

- Madrid (Spain),” *Appl. Energy*, vol. 88, no. 5, pp. 1703–1709, May 2011.
- [12] G. Winter Althaus, B. Gonzalez Landin, A. Pulido Alonso, B. Galvan Gonzalez, M. Maarouf, J. Gonzalez Guerra, and M. Cruz Perez, “Predicción de la demanda de la energía eléctrica a largo plazo: un reto en ingeniería computacional,” *Dyna Ing. e Ind.*, vol. 90, no. 3, pp. 582–584, 2015.
- [13] J. Garcia Martin, “Clasificación del tratamiento térmico de aceros con ensayos no destructivos por corrientes inducidas mediante redes neuronales,” *Dyna Ing. e Ind.*, vol. 89, no. 3, pp. 526–532, 2014.
- [14] A. K. Yadav, H. Malik, and S. S. Chandel, “Selection of most relevant input parameters using WEKA for artificial neural network based solar radiation prediction models,” *Renew. Sustain. Energy Rev.*, vol. 31, pp. 509–519, Mar. 2014.
- [15] P. L. Zervas, H. Sarimveis, J. A. Palyvos, and N. C. G. Markatos, “Prediction of daily global solar irradiance on horizontal surfaces based on neural-network techniques,” *Renew. Energy*, vol. 33, no. 8, pp. 1796–1803, Aug. 2008.
- [16] S. Rehman and M. Mohandes, “Artificial neural network estimation of global solar radiation using air temperature and relative humidity,” *Energy Policy*, vol. 36, no. 2, pp. 571–576, Feb. 2008.
- [17] Y. Jiang, “Prediction of monthly mean daily diffuse solar radiation using artificial neural networks and comparison with other empirical models,” *Energy Policy*, vol. 36, no. 10, pp. 3833–3837, Oct. 2008.
- [18] N. D. Kaushika, R. K. Tomar, and S. C. Kaushik, “Artificial neural network model based on interrelationship of direct, diffuse and global solar radiations,” *Sol. Energy*, vol. 103, pp. 327–342, May 2014.
- [19] M. Benghanem, A. Mellit, and S. N. Alamri, “ANN-based modelling and estimation of daily global solar radiation data: A case study,” *Energy Convers. Manag.*, vol. 50, no. 7, pp. 1644–1655, Jul. 2009.
- [20] S. Sayago, M. Bocco, G. Ovando, and Et.al, “Radiación solar horaria: modelos de estimación a partir de variables meteorológicas básicas,” *Av. en Energías Renov. y Medio Ambient.*, vol. 15, no. 1, pp. 51–57, 2011.
- [21] J. Bonilla, J. Ramirez, and O. Ramirez, “Metodología para el diseño de un modelo univariado de red neuronal para el pronóstico de la temperatura mínima en la zona de Mosquera (Cundinamarca, Colombia),” *Meteorol. Colomb.*, vol. 1, no. 10, pp. 111–120, 2006.
- [22] N. Obregón, F. Monsalve, and S. Garavito, “Predicción de series de tiempo hídricas mediante modelos de redes neuronales artificiales y modelos de árboles de decisión,” in *Memorias del Seminario Internacional La Hidroinformática en la Gestión Integrada de los Recursos Hídricos*, 2003.
- [23] A. González-Rodríguez, “Modelo para la predicción de la radiación solar a partir de redes neuronales artificiales,” *Escuela de Ingeniería de Antioquia*, 2014.
- [24] J. R. G. Sarduy, K. G. Di Santo, and M. A. Saidel, “Linear and non-linear methods for prediction of peak load at University of São Paulo,” *Measurement*, vol. 78, pp. 187–201, Jan. 2016.

AUTHOR PROFILE

Eliana Noriega Angarita is working as professor at Universidad de la Costa. His interest areas are renewable energies and power policies.

Vladimir Sousa Santos is working as professor at Universidad de la Costa. His interest areas are power quality and renewable energies.

Michell Josep Quintero Duran is working as professor at Universidad de la Costa. His interest areas are power systems and renewable energies.

Cesar Gil Arrieta is pursuing his Master degree at Universidad de la Costa. His interest areas are renewable energies and statistics.