*Desalination and Water Treatment*
www.deswater.com

doi: 10.1080/19443994.2014.939861

55 (2015) 3210–3217
September

Taylor & Francis
Taylor & Francis Group

# Prediction of global solar radiation by artificial neural network based on a meteorological environmental data

J. Díaz-Gómez[a], A. Parrales[a], A. Álvarez[b], S. Silva-Martínez[b], D. Colorado[c], J.A. Hernández[b,*]

[a]*Posgrado en Ingeniería y Ciencias Aplicadas, Universidad Autónoma del Estado de Morelos. Av. Universidad 1001 Colonia Chamilpa Cuernavaca, México C.P. 62209, Morelos, Mexico*
[b]*Centro de Investigaciones en Ingeniería y Ciencias Aplicadas (CIICAp), Universidad Autónoma del Estado de Morelos (UAEM). Av. Universidad 1001 Colonia Chamilpa Cuernavaca, México C.P. 62209, Morelos, Mexico, Tel. +52 777 329 7084; Fax: +52 777 329 7984; email: alfredo@uaem.mx (J.A. Hernández)*
[c]*Centro de Investigación en Recursos Energéticos y Sustentables (CIRES), Universidad Veracruzana, Av. Universidad km 7.5, Col. Santa Isabel, Coatzacoalcos C.P. 96535, Veracruz, Mexico*

Received 4 March 2014; Accepted 16 June 2014

ABSTRACT

This paper describes a model based on artificial neural network (ANN) for the estimation of global solar radiation (GSR) on a horizontal surface. The GSR is the primary renewable energy in the nature. The estimation of GSR is decisive; there are many researches that need such estimation since they do not have the required instruments. A simple algorithm with ANN modeling is proposed. The configuration of the back propagation neural network giving the mean square error was a three-layer ANN (5-30-1 neurons) with tangent sigmoid transfer function at the hidden layer, sigmoid transfer function at output layer, and Levenberg–Marquardt algorithm. The sensitivity analysis was developed and showed that all studied variables (time, temperature, relativity humidity, wind speed, and atmospheric pressure) have effect in the prediction of GSR. The results showed that the ANN modeling could simulate the behavior of GSR.

*Keywords:* Global solar radiation; Artificial neural network; Modeling; Sensitivity analysis

## 1. Introduction

The global solar radiation (GSR) is necessary to be known in many processes. Most of the environmental data are usually obtained from local weather stations. For instance, knowing radiation measurements of Cuernavaca city, in Morelos State of Mexico (19˚02´ of latitude, 99˚20´ of longitude, with a height of 1,480 m above sea level), is important for the people who live in such city to use the plenty of solar radiation in their green processes. The climate of Cuernavaca is classified as moderate with long-term temperature and relative humidity varying between 5 and 100% and 6.1–36.1˚C, respectively, throughout the year. The GSR

*Corresponding author.

*Presented at the Conference on Desalination for the Environment: Clean Water and Energy 11–15 May 2014, Limassol, Cyprus*

was found to vary between a minimum of 0 and a maximum of 1,490 $W/m^2$ [1].

The development of a model using an artificial neural network (ANN) allows us to learn and simulate the GSR on a surface by knowing local weather variables specific to this city. This undoubtedly facilitates measurements of solar radiation without using a device that would provide a direct measurement of GSR, such as pyrometer, heliometer, or other device designed for this purpose.

The importance of knowing the GSR for this locality is of vital importance, since in this part of Mexico, there is a daily average of solar radiation that reaches around 241 $W/m^2$. The values of GSR are the most important parameter for the solar energy applications [2]. Thus, the development and implementation of new solar collection technology, such as solar collectors, photovoltaic panels, solar energy traps for pool heating, solar ovens, are compulsory. Precise information on the spatial distributional GSR is also crucial for selecting sites for installing power plants [3]. What entails GSR measurements is the requirement of measurements in real time to carry out the investigations conducted in different research centers, schools, and weather stations. Most of the time the GSR measurement instruments are inaccessible, display faults, lack of providing real-time data, etc. Therefore, the lack of real-time data in local cities makes compulsory usage of data from researchers that miss important and precise information on the GSR. Thus, such missing information can be provided by the proper use of models (as shown here) using available basic weather information of locality

Even in automatic weather stations, where solar radiation is measured, stored data are often lost, due to equipment failure or erroneous data due to problems of sensor calibration, or are outside the expected range [4–6].

The realization of models using an ANN has increased over the years and we can make use of this tool in many fields and applications such as constraint satisfaction, control, data compression, diagnostics, forecasting, comprehensive mapping, multisensory data fusion, optimization, pattern recognition, and risk assessment [7,8]. An ANN is known for its ability to learn data simulation and prediction. The inspiration to use a neural network comes from the biology of the human brain [9–11].

The ANN is useful in this type of prediction of radiation phenomena, thanks to the dynamism and the feasibility of using numerous inputs or outputs. At present, there are numerous models of radiation, which differ in sophistication from simple empirical formulations commonly based on weather data or more complex models which usually involve high computational costs and require also numerous input parameters [12].

There are other works on GSR, for many cities in Europe, including works such as Rehman and Mohandes [13] for the town of Saudi Arabia, Behrang et al. [14] in Iran, extensive works for the town of Turkey as Şenkal [15] where geographical parameters and satellite data are also used. Lu et al. [16] also reported a model to estimate solar radiation, specifically in the Chinese area from a geostationary satellite data, which not only include meteorological parameters but also data from different image channels satellite (IR, visible spectrum, day, hour, daily clearness index, air mass, latitude, altitude, etc.). Tymvios et al. [17] used a recurrent neural network method to estimate the maximum sun radiation using measured values of air temperature and relative humidity.

The ANN allows developing models with different input variables to the determination of the GSR; this paper focuses on using input variables which were obtained with simple instrumentation ranging from thermometers to barometers. These variables are available in any laboratory or weather station and are of easy access, such as temperature, pressure, wind speed, and relative humidity. This reduces costs and offers better and quicker access of data from GSR in real time; the model is fast and simple to use, and this reliable model is known to detect a specific error. Another advantage of the ANN model is that the results being so similar to models that use a large number of variables or expensive computer equipment, other models need to use methods for improving the prediction of ANN, where they perform a pretreatment data as shown in [18–21], or just the fact of making conventional models such as Angstrom, [22] Lewis model [23], and Swartman model [24], or described in the literature by Duffie and Beckman [25] that showed how to predict the GSR on a surface requiring time and money.

In this study, an ANN was used to estimate the daily GSR on a horizontal surface based on local environmental meteorological variables from Cuernavaca, City. Thus, this work provides an alternative for creating models based on empirical data. This research used local meteorological variables such as temperature, relative humidity, atmospheric pressure, wind speed, and GSR, with which the proposed model was developed.

## 2. Methodology

ANN was applied to obtain a mathematical model in which the inputs and targets data play an important role in setting the network. It is particularly used in the multilayer perception network which has been the most used among ANN's architectures both in the renewable energy domain and in the time series forecasting (see Fig. 1) [18]. A simplified sketch of the neuron structure and behavior is presented in Figs. 1 and 2.

The input data considered in this study were five variables: temperature, relative humidity, wind speed, atmospheric pressure, and time as shown in Figs. 3–5, and as output data, GSR incident was used on a surface see Fig. 6.

The ANN adopted corresponds to a multilayer feed-forward network with back propagation and supervised training.

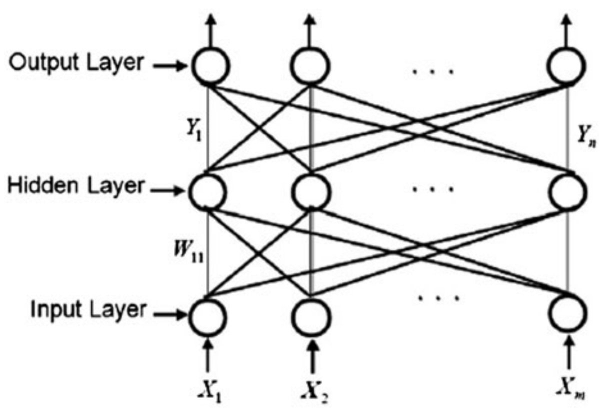


Fig. 1. ANN multilayer structure.

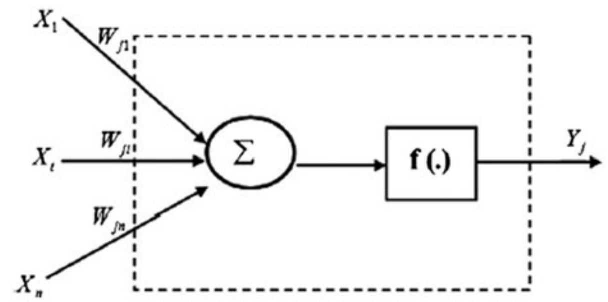


Fig. 2. Details of the neuron.

Fig. 1 shows the general structure of the neural network, while the second layer presents the details of the neuron (see Fig. 2) in which *(f)* could be a sigmoidal or hyperbolic tangent function.

### 2.1. ANN performance indicators

The feed-forward set in this study consists of three layers. In the input layer, a transfer function was used (*hyperbolic tangent sigmoid transfer function*), while in the hidden layer and output layer, a transfer function was adopted (*logarithmic sigmoid transfer function*). In the training function was chosen by the Levenberg–Marquardt algorithm, due to its adjustment, it makes that minimizes the error, since it uses an approximation for the Hessian matrix. The Hessian matrix is the square matrix of second-order partial derivatives of a function; it describes the local curvature of a function of many variables.

Data between 2010 and 2012 with an overall of 139,354 observations with time period of 10 min were measured; 60% of the data were used for training the neural network, while 20% were used for comparison and 20% for testing.

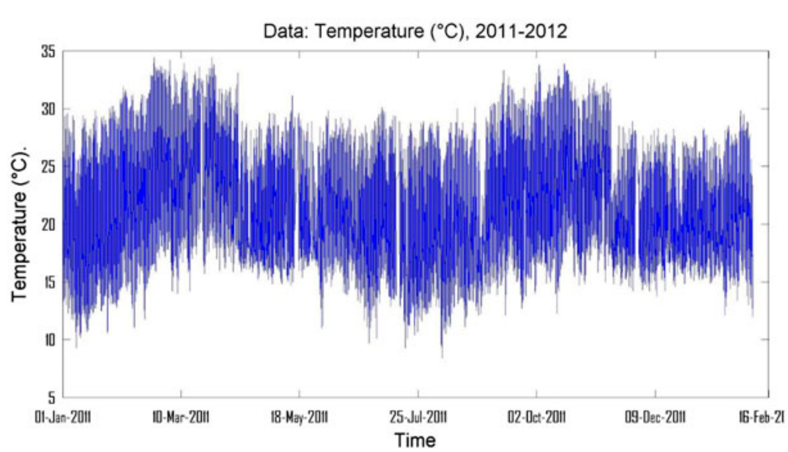


Fig. 3. Temperature (°C) for the 2011–2012 period.

The ANN was developed by a software toolbox Matlab®. The input variables were normalized between 0 and 1, which was obtained by dividing the maximum value of each variable. Table 1 shows the parameters used for the process of training of the ANN proposal.

### 2.2. Neurons number learning

Optimizing the number of neurons was determined based on the minimum mean square error (MSE) value placed data for training and prediction [26–31]. The learning was conducted by checking the MSE over a range of 1–35, the number of neurons (see Fig. 7). The MSE was 0.0137 when using a neuron and decreased to 0.00419 when used 30 neurons. When pass from 15 to 30 neurons, there was a small change in the MSE, reaching a best regression with 30

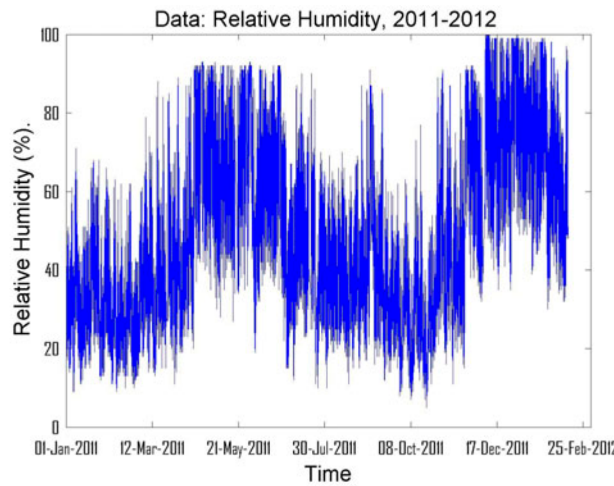


Fig. 4. Relative humidity (%) for the 2011–2012 period.

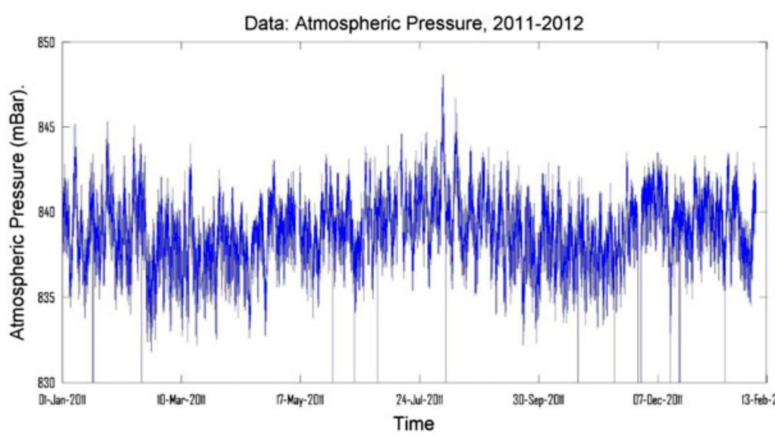


Fig. 5. Atmospheric pressure (mbar) for the 2011–2012 period.

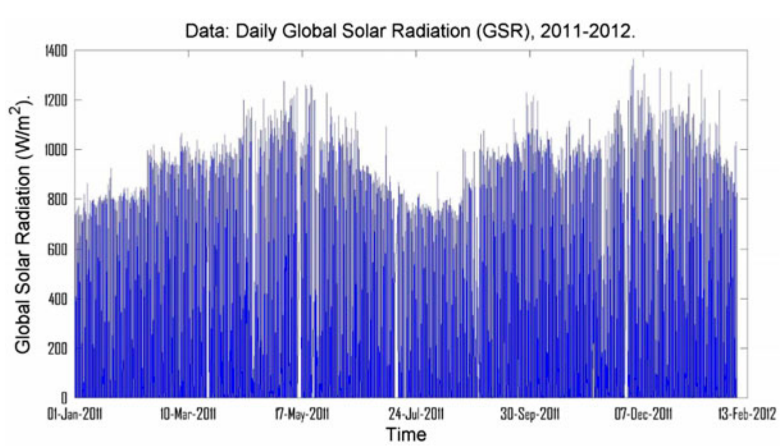


Fig. 6. GSR (W/m$^2$) for the 2011–2012 period.

Table 1
Parameters used in training

| Parameters used in the training process | |
|---|---|
| Performance function | MSE |
| Maximum number of epochs to train | 100 |
| Performance goal | 1E−3 |
| Maximum validation failures | 5 |
| Maximum performance gradient | 1E−10 |
| Initial, $\mu$ | 0.001 |
| $\mu$ Decrease factor | 0.1 |
| $\mu$ Increase factor | 10 |
| Maximum, $\mu$ | 1E+10 |
| Maximum time of train | Infinite |

neurons, increasing the number of neurons in 35 no significant change.

## 3. Results and discussion

An ANN was used to develop a model for prediction of GSR on a surface, whereby an $R^2$ of 0.88841 was obtained, with a configuration of 30 neurons in the hidden layer. Fig. 8 shows the regression obtained.

Table 2 shows the weights obtained by ANN with 30 neurons that produced the best architecture. The best performance with 15 neurons, obtained when $R^2$ = 0.8751, and with 30 neurons with a maximum in the regression of 0.88841 (see Fig. 7) due to the nature of RSG model compared to other models described, is acceptable.

### 3.1. Sensitivity analysis

Besides obtaining the model for the ANN, also a study of the relative importance of the input variables, called sensitivity analysis, which is based on the neural network weight matrix and application by the Garson equation [27,28] in their work, proposed an equation based on the partitioning of the connection weights, using the following equation [1].

$$Ij = \frac{\sum_{k=1}^{m=Mh}\left(\left(\left|W_{jm}^{ih}\right| / \sum_{k=1}^{Ni} W_{jm}^{ih}\right) x \left|W_{nm}^{ho}\right|\right)}{\sum_{k=1}^{Ni}\left\{\sum_{m=1}^{m=Nh}\left(\left|W_{km}^{ih}\right| / \sum_{k=1}^{Ni} W_{km}^{ih}\right) x \left|W_{nm}^{ho}\right|\right\}} \quad (1)$$

where $Ij$ is the relative importance of $j_{th}$ input variable in the output variable, $N_h$ and $N_i$ are the number of entries in the hidden layer and the connection weight $W$, the suffixes $i$, $o$ refer to the input layer, hidden, and output, respectively, the suffixes $k$, $m$, and $n$ refer to the hidden input, neurons, and output, respectively.

Fig. 9 shows the relative importance of the input variables when applied a sensitivity analysis. In this case, the more important input variable is the time, followed by the atmospheric pressure, the temperature,

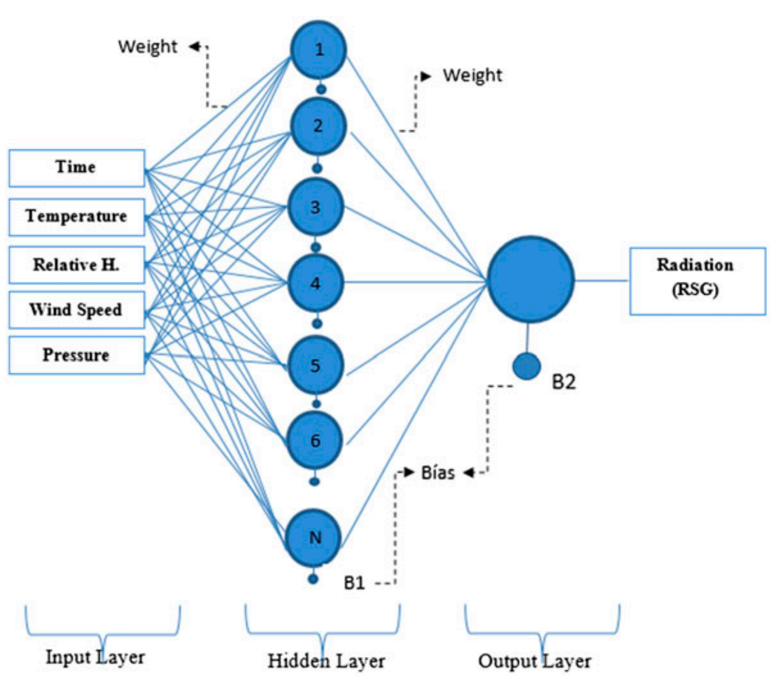


Fig. 7. Detail of ANN, input and output.

Table 2
Weights matrix

| Neuron | Time | Temperature | Relative humidity | Wind speed | Pressure | Radiation (RSG) |
|---|---|---|---|---|---|---|
| | W1 | | | | | W2 |
| 1 | 2.50E+03 | −1.91E+00 | 9.85E+00 | 9.92E+00 | −4.71E+01 | 0.48911939 |
| 2 | 1.76E+03 | 3.18E+00 | 1.17E+00 | 2.70E+00 | 1.11E+02 | −1.1198815 |
| 3 | −1.07E+02 | 3.87E+00 | 1.40E+00 | 1.98E+01 | 1.15E+02 | 0.99049345 |
| 4 | 1.64E+03 | −1.64E+00 | 2.47E+00 | 3.98E+00 | −3.28E+01 | −1.1520632 |
| 5 | −6.22E+02 | 3.59E−01 | −3.45E+00 | −3.33E+00 | 1.05E+02 | 1.0837525 |
| 6 | 2.21E+03 | 2.12E+00 | 1.64E+00 | 9.63E−02 | 1.38E+02 | 59.309229 |
| 7 | 2.20E+03 | 1.78E+00 | −8.08E−01 | −5.95E−01 | 8.00E+01 | 10.313996 |
| 8 | 2.59E+03 | 1.25E+00 | −4.54E−02 | −6.17E−01 | 7.92E+01 | 36.606521 |
| 9 | −9.82E+02 | −6.20E+00 | −2.14E+00 | 1.70E+00 | −2.24E+02 | 7.9266037 |
| 10 | −5.68E+02 | −1.04E+01 | −8.78E+00 | 2.12E+00 | −1.05E+02 | −1.0560186 |
| 11 | −1.54E+03 | −7.57E+00 | −9.70E+00 | 5.48E+00 | 4.36E+01 | 0.69781506 |
| 12 | 3.08E+03 | −1.24E+00 | 1.51E−01 | 1.42E−01 | −1.18E+00 | 54.471562 |
| 13 | 1.73E+03 | −3.78E+00 | −1.40E+00 | 8.12E−02 | −1.25E+02 | 32.336055 |
| 14 | 2.50E+03 | 2.87E+00 | −5.31E+00 | 1.80E+00 | −1.16E+02 | 0.53597143 |
| 15 | 3.74E+02 | 3.89E+00 | 1.36E+00 | −1.11E+00 | 1.53E+02 | −23.651539 |
| 16 | 4.96E+02 | −7.21E+00 | 6.41E+00 | 9.05E−01 | −1.73E+02 | 0.5234023 |
| 17 | −1.96E+03 | 4.81E+00 | 1.59E+00 | −1.50E+00 | 1.48E+02 | 32.828777 |
| 18 | 2.53E+03 | −1.72E+01 | −2.75E+00 | 7.30E+00 | −1.90E+02 | −0.8241482 |
| 19 | −2.77E+02 | −6.91E−01 | 2.07E+00 | 1.47E+00 | 5.52E+01 | −4.2419543 |
| 20 | −2.52E+03 | 1.61E+00 | 8.35E−01 | 7.05E−02 | 2.92E+01 | 50.667646 |
| 21 | 2.99E+03 | −3.52E+00 | −1.35E+00 | 5.31E−02 | −9.49E+01 | −45.789322 |
| 22 | 2.51E+03 | 3.75E−03 | −2.08E−01 | −4.07E−01 | 3.34E+01 | 48.979677 |
| 23 | 2.27E+03 | −4.80E+00 | −1.29E+00 | 6.86E−01 | −1.56E+02 | 32.336402 |
| 24 | −2.74E+03 | 4.63E+00 | 1.44E+00 | −6.13E−01 | 1.38E+02 | 42.000404 |
| 25 | 1.32E+03 | −3.70E+00 | −6.89E+00 | −2.88E−01 | 6.88E+01 | −1.6915169 |
| 26 | 1.05E+03 | −5.82E+00 | −1.62E+00 | 1.43E+00 | −2.09E+02 | 21.735065 |
| 27 | −1.84E+03 | −7.63E+00 | −4.40E+00 | 2.27E+00 | −1.43E+02 | 2.718599 |
| 28 | −8.21E+02 | −7.59E+00 | −2.57E+00 | 3.69E+00 | −2.10E+02 | −4.5912158 |
| 29 | −1.38E+03 | −4.00E+00 | −2.19E+00 | 1.74E−01 | −2.11E+02 | 42.648142 |
| 30 | 2.84E+03 | −2.23E+00 | −2.69E−01 | 2.57E−01 | −5.09E+01 | −62.883813 |

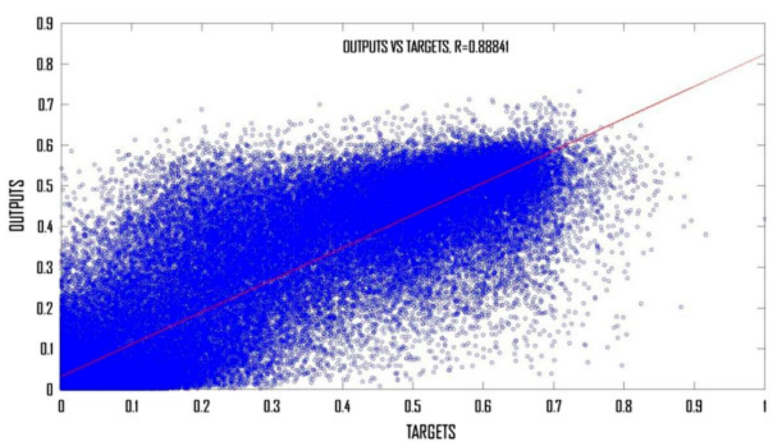


Fig. 8. Regression of 0.88841 was obtained, target vs. simulated data for ANN.

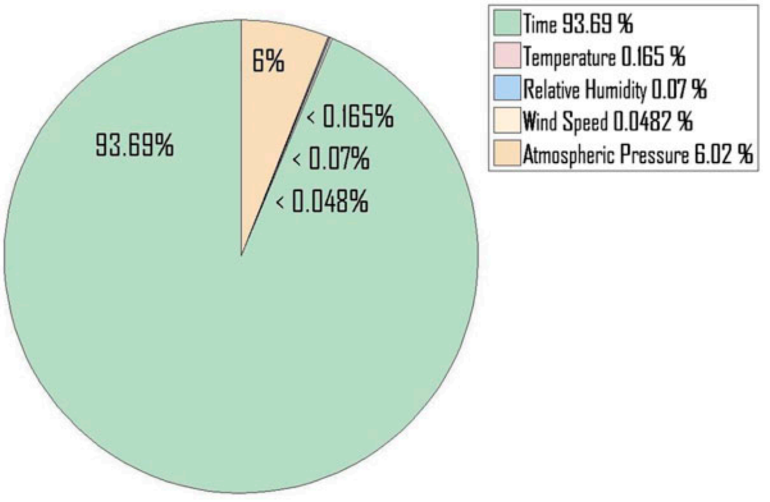


Fig. 9. Relative importance of input variables.

the relative humidity, and finally, the wind speed. Atmospheric pressure parameter was not expected to have a greater importance than the other variables, as there is no variation in magnitude, ranging from 832 to 846 (mbar) as shown in Fig. 5, but the behavior of the small variation with respect to solar radiation was a significant relationship over the temperature or relative humidity.

## 4. Conclusions

This study proposed an algorithm that uses an ANN to estimate GSR on a horizontal surface of Cuernavaca city, Morelos State of Mexico. In this algorithm, a stable and efficient model was constructed to determine the relationship between input variables and the GSR.

The algorithm was tested with data from the period 2010, 2011, and 2012. The neural network was trained with solar radiation data obtained from national meteorological synoptic station of Cuernavaca in the same period, as output data. ANN was optimized to yield the best architecture with 30 neurons, resulting in a MSE of 0.0095. With this architecture adopted, the configured ANN obtained an $R^2$ of 0.88841.

The Garson's equation was used for a sensitivity analysis. This allows us to know the variable with the most importance relative and effect that exists in the modeling of the GSR. Time was the variable that had a greater relative importance, followed by the atmospheric pressure, temperature, relative humidity, and wind speed, in that order of importance. It seems interesting that the atmospheric pressure has a tendency to solar radiation.

The ANN is a very useful tool when trying to simulate and find predictive models for a specific time.

Finally, the parameters such as temperature and relative humidity did not have a tendency to behave in the same way as the radiation. The development of this model is definitely a contribution to local cities, and specifically to Cuernavaca, since no GSR models are available for this part of the country. It is also compulsory to have GSR data in real time where the computational costs of modeling by ANN are competitive in comparison with other models and predicting weather variables. Future work will focus on a variable in order to know the relationship between atmospheric pressure and GSR.

## Acknowledgment

The authors acknowledge the support of CONAGUA, specifically the National Weather System, to the database provided for conducting this study.

## References

[1] CONAGUA, Comisión Nacional del Agua, Internet consulted in February 2014. Available from: www.conagua.gob.mx.
[2] K. Bakirci, Correlations for estimation of daily global solar radiation with hours of bright sunshine in Turkey, Energy 34(4) (2009) 485–501.
[3] J.D. Mondol, Y.G. Yohanis, B. Norton, Solar radiation modelling for the simulation of photovoltaic systems, Renew. Energy 33(5) (2008) 1109–1120.

[4] L. Hontoria, J. Aguilera, P. Zufiria, Generation of hourly irradiation synthetic series using the neural network multilayer perceptron, Sol. Energy 72(5) (2002) 441–446.
[5] L.A. Hunt, L. Kuchar, C.J. Swanton, Estimation of solar radiation for use in crop modelling, Agric. For. Meteorol. 91 (1998) 293–300.
[6] M.G. Abraha, M.J. Savage, Comparison of estimates of daily solar radiation from air temperature range for application in crop simulations, Agric. For. Meteorol. 148 (2008) 401–416.
[7] R. Aguiar, M. Collares-Pereira, TAG: A time-dependent, autoregressive, Gaussian model for generating synthetic hourly radiation, Sol. Energy 49 (1992) 167–174.
[8] D.W. Patterson, Artificial Neural Networks. Theory and Applications, Prentice Hall, Singapore, 1996.
[9] N. Prakash, S.A. Manikandan, L. Govindarajan, V. Vijayagopal, Prediction of biosorption efficiency for the removal of copper(II) using artificial neural networks, J. Hazard. Mater. 152(3) (2008) 1268–1275.
[10] N. Daneshvar, A.R. Khataee, N. Djafarzadeh, The use of artificial neural networks (ANN) for modeling of decolorization of textile dye solution containing C.I. Basic Yellow 28 by electrocoagulation process, J. Hazard. Mater. 137 (2006) 1788–1795.
[11] A. Aleboyeh, M.B. Kasiri, M.E. Olya, H. Aleboyeh, Prediction of azo dye decolorization by UV/H2O2 using artificial neural networks, Dyes Pigm. 77 (2008) 288–294.
[12] P. Martí, M. Gasque, Improvement of temperature-based ANN models for solar radiation estimation through exogenous data assistance, Energy Convers. Manage. 52(2) (2011) 990–1003.
[13] S. Rehman, M. Mohandes, Artificial neural network estimation of global solar radiation using air temperature and relative humidity, Energy Policy 36(2) (2008) 571–576.
[14] M.A. Behrang, E. Assareh, A. Ghanbarzadeh, A.R. Noghrehabadi, The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data, Sol. Energy 84(8) (2010) 1468–1480.
[15] O. Şenkal, Modeling of solar radiation using remote sensing and artificial neural network in Turkey, Energy 35(12) (2010) 4795–4801.
[16] N. Lu, J. Qin, K. Yang, J. Sun, A simple and efficient algorithm to estimate daily global solar radiation from geostationary satellite data, Energy 36(5) (2011) 3179–3188.
[17] F.S. Tymvios, C.P. Jacovides, S.C. Michaelides, C. Scouteli, Comparative study of Ångström's and artificial neural networks' methodologies in estimating global solar radiation, Sol. Energy 78(6) (2005) 752–762.
[18] C. Voyant, M. Muselli, Ch. Paoli, M.L. Nivet, Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation, Energy 36(1) (2011) 348–359.
[19] Ch. Paoli, C. Voyant, M. Muselli, M.L. Nivet, Forecasting of preprocessed daily solar radiation time series using neural networks, Sol. Energy 84(12) (2010) 2146–2160.
[20] J. Faraway, C. Chatfield, Times series forecasting with neural networks: A case study, research report 95-06 of the statistics group, University of Bath, Bath, 1995.
[21] J.D. Hamilton, Times Series Analysis, Princeton University Press, Princeton, 1994, ISBN 0-691-04289-6.
[22] A. Angstrom, Solar and terrestrial radiation, J. Roy. Meteor. Soc. 50 (1924) 121–126.
[23] G. Lewis, Estimates of irradiance over Zimbabwe, Sol. Energy 31 (1983) 609–612.
[24] R.K. Swartman, O. Ogunlade, Solar radiation estimates from common parameters, Sol. Energy 11 (1967) 170–172.
[25] J. Duffie, W. Beckman, Solar Engineering of Thermal Processes, Wiley, New York, NY, 2013.
[26] K. Yetilmezsoy, S. Demirel, Artificial neural network (ANN) approach for modeling of Pb(II) adsorption from aqueous solution by Antep pistachio (*Pistacia vera* L.) shells, J. Hazard. Mater. 153 (2008) 1288–1300.
[27] APHA, AWWA, WPCF, Standard Methods for the Examination of Water and, Wastewater, eighteenth ed., American Public Health Association/American Water Works Association/Water Pollution Control Federation, Washington, DC, 1992.
[28] G.D. Garson, Interpreting neural-network connection weights, AI Expert 6 (1991) 47–51.
[29] J.A. Hernández, D. Colorado, Uncertainty analysis of COP prediction in a water purification system integrated into a heat transformer using several artificial neural networks, Desalin. Water Treat. 51(7–9) (2013) 1443–1456.
[30] E. Cardoso, S. Silva-Martinez, A. Alvarez, J.A. Hernández, Design and experimental analysis of low-cost heat water solar collectors, Desalin. Water Treat. 51(4–6) (2013) 1302–1309.
[31] D. Juárez-Romero, N. Shah, F. Pliego-Solórzano, J.A. Hernández, J. Siqueiros, A. Huicochea, Heat and mass transfer in a horizontal pipe absorber for a heat transformer, Desalin. Water Treat. 10(1–3) (2009) 238–244.