Available online at www.sciencedirect.com

ScienceDirect

Solar Energy 105 (2014) 91–98

SOLAR ENERGY

www.elsevier.com/locate/solener

Daily global solar radiation prediction based on a hybrid Coral Reefs Optimization – Extreme Learning Machine approach

S. Salcedo-Sanz [a,*], C. Casanova-Mateo [b], A. Pastor-Sánchez [a], M. Sánchez-Girón [a]

[a] Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, Spain
[b] Department of Applied Physics, Universidad de Valladolid, Valladolid, Spain

Received 21 December 2013; received in revised form 4 March 2014; accepted 7 April 2014

Communicated by: Associate Editor Christian A. Gueymard

Abstract

This paper discusses the performance of a novel Coral Reefs Optimization – Extreme Learning Machine (CRO–ELM) algorithm in a real problem of global solar radiation prediction. The work considers different meteorological data from the radiometric station at Murcia (southern Spain), both from measurements, radiosondes and meteorological models, and fully describes the hybrid CRO–ELM to solve the prediction of the daily global solar radiation from these data. The algorithm is designed in such a way that the ELM solves the prediction problem, whereas the CRO evolves the weights of the neural network, in order to improve the solutions obtained. The experiments carried out have shown that the CRO–ELM approach is able to obtain an accurate prediction of the daily global radiation, better than the classical ELM, and the Support Vector Regression algorithm.

Keywords: Daily global solar radiation prediction; Coral Reefs Optimization algorithm; Extreme Learning Machines

1. Introduction

Solar radiation prediction is an important problem with direct applications in renewable energy. Solar is one of the most important green sources of energy, that is currently under expansion in many countries of the world, specially in those with more solar potential, such as mid-east and southern Europe countries (Kalogirou, 2014). An accurate estimation of the energy production in solar energy systems involves the accurate prediction of solar radiation, depending on different atmospheric variables (Khatib et al., 2012; Inman et al., 2013; Sozen et al., 2004; Voyant et al., 2011).

In recent years, several works have been developed to try to predict solar radiation using machine learning techniques and environmental parameters. They used different input geographical and atmospheric parameters like latitude, longitude, temperature, wind speed and direction, daily global irradiation, sunshine duration or precipitation (Mellit and Kalogirou, 2008; Mubiru, 2008). According to Bilgili and Ozoren (2011), sunshine duration, air temperature and relative humidity are the most widely used meteorological parameters to predict daily solar radiation and its components. All these parameters are well correlated with the daily solar global radiation (Yacef et al., 2012). In López et al. (2005) a Bayesian framework for artificial neural networks, named as automatic relevance determination method, was developed to evaluate the more relevant input parameters in modelling solar irradiation. In fact, neural computation paradigm has been massively

* Corresponding author. Address: Department of Signal Processing and Communications, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain. Tel.: +34 91 885 6731; fax: +34 91 885 6699.
E-mail address: sancho.salcedo@uah.es (S. Salcedo-Sanz).

http://dx.doi.org/10.1016/j.solener.2014.04.009
0038-092X/

applied to this prediction problem, like in Benghanem and Mellit (2010), where it is shown that Radial Basis Functions (RBF) neural networks obtain excellent performance in the estimation of solar radiation. In Dorvlo et al. (2002) a comparison between Multi-Layer Perceptrons (MLPs) and RBF neural networks in a problem of solar radiation estimation is carried out. Experiments in eight stations in Oman show the good results obtained with the neural algorithms. A similar approach, also comparing MLPs and RBFs (with different predictive variables) has been recently proposed in Behrang et al. (2010). In this case, the authors test the neural network with data obtained in Iran. In Ji and Chee (2011) a hybrid approach based on ARMA and time delay neural networks has been successfully tested in data from a solar station in Singapore. Other approaches that also use neural networks as prediction methodology, include novel predictive variables, such as satellite data (Senkal and Kuleli, 2009) or temperature and relative humidity (Rehman and Mohandes, 2008). Alternative machine learning algorithms, such as Support Vector Machines (SVMs) have been also applied to solar radiation prediction problems from meteorological predictive variables (Chen et al., 2011; Zeng and Qiao, 2013). Specifically, a least-square SVM is proposed in that work, comparing the results obtained with that of auto-regressive and RBF neural networks. In Rahimikoob (2010) the potential of multi-layer perceptron neural networks with back-propagation training algorithm is shown in a problem of global solar radiation estimation in Iran. Results comparing the performance of the neural networks with that of an empirical equation for global solar radiation prediction (Hargreaves and Samani equation) show good performance of the neural approach. In Bhardwaj et al. (2013) a hybrid approach that includes hidden Markov models and generalized fuzzy models has been proposed and tested in real solar irradiation data in India.

Recently, the so called Extreme Learning Machine (ELM) has been introduced as an extremely fast training method for multi-layer perceptron type neural networks (Huang et al., 2006). The ELM is currently a state of the art approach to train neural networks, quite extended due to its excellent performance in many different problems. The ELM has also been successfully applied to solar radiation prediction problems, like in Sahin et al. (2014), where the ELM approach is applied to a solar radiation prediction problem from satellite measures. In Alharbi (2013) a case study of solar radiation prediction in Arabia Saudi is discussed comparing the performance of artificial neural networks with classical training and ELMs. In Dong et al. (2014) a hybrid wavelet-ELM approach is tested in a problem of solar irradiation prediction for application in a photovoltaic power station. Finally, in Salcedo-Sanz et al. (2013) a comparison of a Support Vector Regression algorithm and an ELM is carried out in a problem of direct solar radiation prediction, with application in solar thermal energy systems.

In the last few years, different works have tried to enhance the ELM performance by hybridizing it with evolutionary computation algorithms. Basically, two main approaches have been proposed: the first one, consists of carrying out a feature selection approach using the ELM as wrapper classifier or regressor (Chyzhyk et al., 2014; Landa-Torres et al., 2012). In this case, the evolutionary methods try to select the best set of features in terms of the ELM performance (classification accuracy or probability of error if we are dealing with regression problems). The second approach, which is the one we are interested in this paper, consists of using micro-evolution in order to obtain the best set of weights and biases in the input layer of the ELM (Zhu et al., 2005). This approach is quite sensitive to the global search algorithm used, and an excess of evolution may lead to overfitting and therefore to poor results. However this, positive results have been recently reported using micro-evolutionary algorithms (Lahoz et al., 2013), particle swarm (Han et al., 2013) or evolutionary ensembles (Wang and Alhamdoosh, 2013).

In this paper we discuss the performance of a hybrid evolutionary-ELM algorithm in a problem of daily global solar radiation prediction. Specifically, we explore the performance of a recently proposed evolutionary-type approach for global optimization, the so called Coral Reefs Optimization (CRO) algorithm, hybridized with ELM in this problem of solar radiation prediction. The CRO has excellent properties of fast convergence to optimal values, and can be used for carrying out evolution in ELM weights, in order to enhance the performance of these machines. In addition, we also explore in this work the effect of including new atmospheric predictive variables as inputs of the ELM. Meteorological variables such as the total ozone content of the atmosphere, the aerosol optical depth and precipitable water (intrinsically related to clearness index and Relative Air Mass Liou, 2002) are included in the prediction system, together with variables from atmospheric models such as the prediction of cloudiness in the zone under study. In the experimental part of the paper we show how the proposed CRO–ELM algorithm is able to successfully solve this solar radiation problem, improving the performance of the classical ELM and SVMr approaches.

2. Daily global solar radiation prediction from novel meteorological variables

In this section we describe the problem of daily global solar radiation we tackle, including a brief description of the variables involved in the prediction problem and the objective data of radiation available for the study. Note that the total set of meteorological variables included as input predictive variables in this paper has not been, to our knowledge, considered in other studies about radiation prediction, it is a novel contribution of this research.

First of all, the objective variable (prediction target) for this problem is the real global solar irradiation that reaches the ground. Data from the Meteorological State Agency of Spain (AEMET) in the radiometric observatory of Murcia (Southern Spain, 38.0°N, 1.2°W) were used. Specifically, global daily mean values from the measurements of a pyrheliometer mounted over an automatic solar tracker have been considered. These radiation data ranges from the 1st January 2010 to 31st December 2011, two years of daily measurements.

The description of the predictive input meteorological variables considered to tackle the global solar radiation prediction is as follows:

It has been recently reported that Clearness Index (horizontal global irradiation/horizontal extraterrestrial irradiation) and Relative Air Mass are the more relevant input variables to the neural network in problems of solar radiation prediction (Mellit and Kalogirou, 2008; López et al., 2005). To some degree or another, both parameters are related to the two processes involved in the solar radiation extinction: scattering and absorption. Considering this, and taking into account that the aim of this study is to predict global solar irradiation, we assume that the regressor techniques considered would work more accurately if they are trained with parameters related to atmospheric scattering and absorption processes. Scattering is a physical process by which a particle in the path of an electromagnetic wave continuously abstract energy from the incident wave and re-radiates that energy in all directions. In the atmosphere one of the most important particles responsible for scattering are aerosols. They are known to be produced by natural processes (volcanic dust, particles from sea spray, windblown dust, etc.) as well as by human activity (Liou, 2002). Its concentration varies with locality and it generally decrease rapidly with height in the troposphere. Then, in order to take into account the presence of atmospheric aerosols, we consider the daily mean aerosol optical depth product obtained from a Cimel CE318 sunphotometer as input parameter. This instrument makes direct sun measurements at wavelengths 340, 380, 440, 500, 670, 870 and 1020 nm. with a field of view of 1.2°. In this case, the instrument belongs to AEMET, it is located in the radiometric observatory of Murcia and is part of the NASA Aerosol Robotic Network (AERONET) (Holben et al., 1998).

The utilization of the aerosol optical depth is also interesting because some aerosols can absorb solar energy (Wang et al., 2009). Furthermore, the numerous gases that make up the atmosphere can scatter and absorb solar radiation to varying degrees. The permanent constituents (mainly nitrogen, oxygen and argon) account for more than 99.96% of the atmosphere by volume (Liou, 2002) and their extinction effect could be consider as constant when the period of study is short (in our case, 2 years). Nevertheless, it is interesting to have information about the variable concentration (in space and time) of two very important gases in terms of energetic absorption: ozone and water vapour.

Ozone concentration can be derived from Brewer spectrophotometer measurements. Thanks to this instrument it is possible to derive total ozone amount from the ratio of measured sunlight intensities at five wavelengths between 306 and 320 nm with a resolution of 0.6 nm, where the absorption by ozone presents large spectral structures (Anton et al., 2008). Thus, we have used the daily mean ground-based total ozone amount derived from the Murcia Brewer spectrophotometer. AEMET operates a national Brewer spectrophotometer network, having one of its instruments located at the radiometric station of Murcia.

Water vapour present in the atmosphere could be considered by selecting the total precipitable water (the amount of liquid water, in mm, if all the atmospheric water vapour in the column were condensed) product derived from an atmospheric sounding as an input parameter. Fortunately the radiometric observatory of Murcia also hosts an upper-air sounding station. Although an atmospheric sounding is launched every twelve hours (00:00 and 12:00 UTC) we have calculated the mean Total Precipitable Water (TPW) value for every two soundings in order to have the same temporal resolution as the Cimel and Brewer data. Murcia TPW data are freely available on the internet (http://weather.uwyo.edu/upperair/sounding.html).

For a predicting operative approach, it is necessary to contemplate the presence of clouds, because its existence over the area of study clearly affects the amount of solar irradiation that reaches a ground surface (Mellit and Kalogirou, 2008). For that, we have used data from the numerical weather prediction model GFS (Global Forecast System) maintained by the National Center for Environmental Prediction (USA) (Kanamitsu et al., 1991). Although its horizontal resolution is not too fine, this model has the advantage that its data are freely available on Internet. To overcome the spatial limitation, the daily mean cloud amount forecasted by the GFS model was taken at the grid point closest to the region of interest.

Finally, the theoretical extraterrestrial solar irradiation calculated with the classical equations (Iqbal, 1983) has been also considered as an input variable.

3. The hybrid CRO–ELM algorithm for solar radiation prediction

In this section we present the hybrid CRO–ELM algorithm proposed in this paper for solar radiation prediction. The actual prediction of solar radiation can be carried out with the ELM approach, and the CRO is used to evolve the input weights of the network in order to improve its performance. We briefly describe both algorithms and also the encoding and methodology carried out to solve this problem with the CRO–ELM algorithm proposed.

3.1. Basic Extreme Learning Machines

The ELM is a novel and fast training method for multilayer perceptrons type neural networks, recently proposed in Huang et al. (2006) and applied thereafter to a large number of classification and regression problems (Huang et al., 2011; Huang and Chen, 2007, 2008). The ELM's structure is similar to the network given in Fig. 1. The most significant characteristic of the basic ELM training is that it is carried out just by randomly setting the network weights, and then obtaining the inverse of the hidden-layer output matrix. The advantages of this technique are its simplicity, which makes the training algorithm extremely fast, and also its outstanding performance when compared to alternative learning methods, usually better than other established approaches such as classical multi-layer perceptrons or support vector machines. Moreover, the universal approximation capability of the ELM network, as well as its classification capability, have been already proven (Huang et al., 2006, 2012).

Mathematically, the ELM method can be described in the following way: Given a training set $\aleph \triangleq \{(\mathbf{x}_i, \mathbf{t}_i) \mid \mathbf{x}_i \in \mathbb{R}^n, \mathbf{t}_i \in \mathbb{R}, i = 1, \ldots, N_T\}$, an activation function $g(x)$ and a number of hidden nodes ($\widetilde{N}$), the ELM algorithm is summarized in a number of steps:

1. Randomly assign inputs weights $\mathbf{w}_i$ and bias b_i, with $i = 1, \ldots, \widetilde{N}$.
2. Calculate the $N_T \times \widetilde{N}$ hidden-layer output matrix $\mathbf{H}$, defined as

$$\mathbf{H} \triangleq \begin{bmatrix} g(\mathbf{w}_1\mathbf{x_1} + b_1) & \cdots & g(\mathbf{w}_{\widetilde{N}}\mathbf{x_1} + b_{\widetilde{N}}) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_1\mathbf{x_{N_T}} + b_1) & \cdots & g(\mathbf{w}_{\widetilde{N}}\mathbf{x_{N_T}} + b_{\widetilde{N}}) \end{bmatrix}. \quad (1)$$

3. Calculate the output weight vector β as

$$\beta = \mathbf{H}^{\dagger}\mathbf{T}, \quad (2)$$

where $\mathbf{H}^{\dagger}$ stands for the Moore–Pennrose inverse of matrix $\mathbf{H}$ (Huang et al., 2006), and $\mathbf{T} \triangleq [t_1, \ldots, t_{N_T}]^T$ is the training output vector.

Note that the number of hidden nodes $\widetilde{N}$ is a free parameter of the ELM training, and must be estimated to obtain good results. Usually, scanning a range of $\widetilde{N}$ values is the most practical solution for this problem.

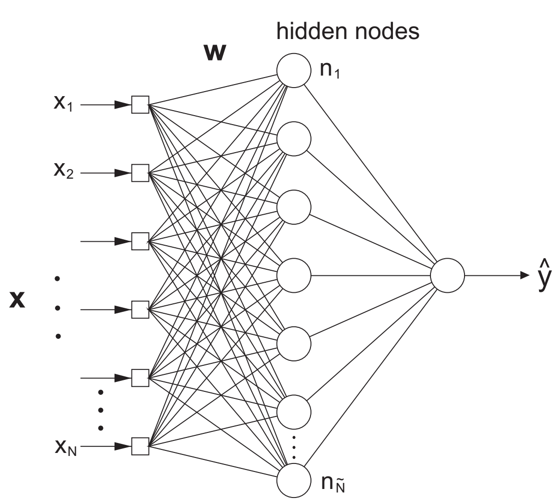

Fig. 1. Outline of the neural network structure trained with the ELM approach.

In this paper, we explore the possibility of slightly evolving the ELM input weights in order to obtain a better performance of the algorithm. We propose to use a recently developed meta-heuristic with very good properties of convergence to do this task, the Coral Reefs Optimization algorithm (CRO).

3.2. The Coral Reefs Optimization algorithm

The CRO is a class of evolutionary meta-heuristic algorithm based on corals' reproduction and reefs formation, first proposed in Salcedo-Sanz et al. (2013) and recently applied to renewable energy problems (Salcedo-Sanz et al., 2014). The CRO is based on the artificial simulation of a coral reef, Λ, consisting of a $\mathcal{N} \times \mathcal{M}$ square grid. It is then assumed that each square in the grid (i, j) is able to allocate a coral (or colony of corals) $\Xi_{i,j}$, that represents a given solution to the considered optimization problem. The CRO algorithm is usually initialized at random by assigning some squares in Λ to be occupied by corals (solutions to the optimization problem tackled) and some other squares in Λ to be empty, i.e. places in the simulated reef where new corals can freely settle down and grow up in the future. The rate between free/occupied squares in Λ at the beginning of the algorithm is denoted as ρ. Each coral is associated with a *health* function $f(\Xi_{ij}) : \mathcal{I} \to \mathbb{R}$, which represents the problem's objective function. The CRO is based on the fact that when the reef will progress, as long as healthier (stronger) corals (which represent better solutions to the problem at hand) survive, while less healthy corals perish.

The second step of the CRO is devoted to simulate a reef formation, by means of different operators that mimic all the processes occurring in a reef: corals' reproduction, larvae setting and depredation of unhealthy corals. This is carried out by means of different operators:

1. *Broadcast Spawning (external sexual reproduction)*: this type of reproduction consists of the following steps:
   1.a. Select a fraction of corals in the reef (F_b) to be broadcast spawners. The rest of the corals in the reef (i.e. $1 - F_b$) will reproduce by brooding at a later step in the algorithm.
   1.b. Broadcast spawner couples are selected to reproduce, and each of such couples form one or two coral larvae by sexual crossover. Note that, once two corals have been selected to be the parents of a larva, they are not chosen anymore in this reproduction step.

2. *Brooding (internal sexual reproduction)*: as previously mentioned, the fraction of corals that will reproduce by brooding is $1 - F_b$. The brooding modeling consists of the formation of a coral larva by means of a random mutation of the brooding-reproductive coral (self-fertil-

Table 1
Variables considered in this problem of global solar radiation prediction. The source of each variable is also reported in this table.

Source	Data	Statistics	Units
Cimel sunphotometer	Aerosol Optical Depth (at 340, 380, 440, 500, 670, 870 and 1020 nm)	Daily mean	–
Brewer spectrophotometer	Total Ozone Amount	Daily mean	Dobson
Atmospheric sounding	Total Precipitable Water	Daily mean	mm
GFS	Cloud amount	Daily mean	%
Classical equations	Theoretical extraterrestrial solar irradiation	Daily mean	kJ/m^2
Pyranometer	Measured global solar irradiance	Daily mean	kJ/m^2

Table 2
Different statistical errors (RMSE, MAE and Bias, in (W/m^2)) of the daily prediction in the test set, by the classical ELM and the Support Vector Regression approaches.

Algorithm	RMSE	Variance	MAE	Variance	Bias	Variance
Classical ELM	0.07243	0.0015	0.0199	3.8×10^{-4}	5.9×10^{-4}	8.5×10^{-6}
SVMr	0.0692	0.02	0.0186	2.3×10^{-4}	4.8×10^{-4}	5.7×10^{-6}

ization considering hermaphrodite corals). The produced larva is then released out together with larvae formed by broadcast spawning.

3. *Larvae setting*: This is the key step in the algorithm. Once all the larvae are formed at a given step k (either through broadcast spawning or by brooding), they will try to settle down and grow up in the reef. The process is the following: first, the health function of each coral larva is computed. Second, each larva will randomly try to set in a square (i, j) of the reef. If the square is empty (free space in the reef), the coral larva grows therein, no matter the value of its health function. On the other hand, if a coral is already occupying the square at hand, the new larva will set only if its health function is better than that of the existing coral. We define a number κ of attempts for a larva to settle down in the reef: after κ unsuccessful tries, it will be depredated by animals in the reef.
4. *Asexual reproduction*: in the modeling of asexual reproduction (budding or fragmentation), the overall set of existing corals in the reef are sorted as a function of their level of healthiness (given by $f(\Xi_{ij})$). Then, a small fraction F_a of the best corals duplicates and try to settle in a different part of the reef, by following the setting process described in Step 3. Note that a maximum number of identical corals (μ) may be allowed in the reef.
5. *Depredation in polyp phase*: corals may die during the reef formation phase of the CRO algorithm. At the end of each reproduction step k, a small number of corals in the reef can be depredated, thus liberating space in the reef for next coral generation. The depredation operator is applied with a very small probability P_d at each step k, and exclusively to a fraction F_d of the poorer healthy corals in Λ.

3.3. Problem encoding and methodology

Finally, the problem encoding and the methodology followed in the experiments is described in this section. The encoding of the problem in the CRO is the first issue to be solved. In this case, we use the CRO algorithm to carry out an evolution of the ELM weights, so we need to encode these weights in the corals of the CRO. To do this, it is necessary to set the number of neurons in the hidden layer of the ELM, since the number of weights depends on this number of neurons. Previous experimental work

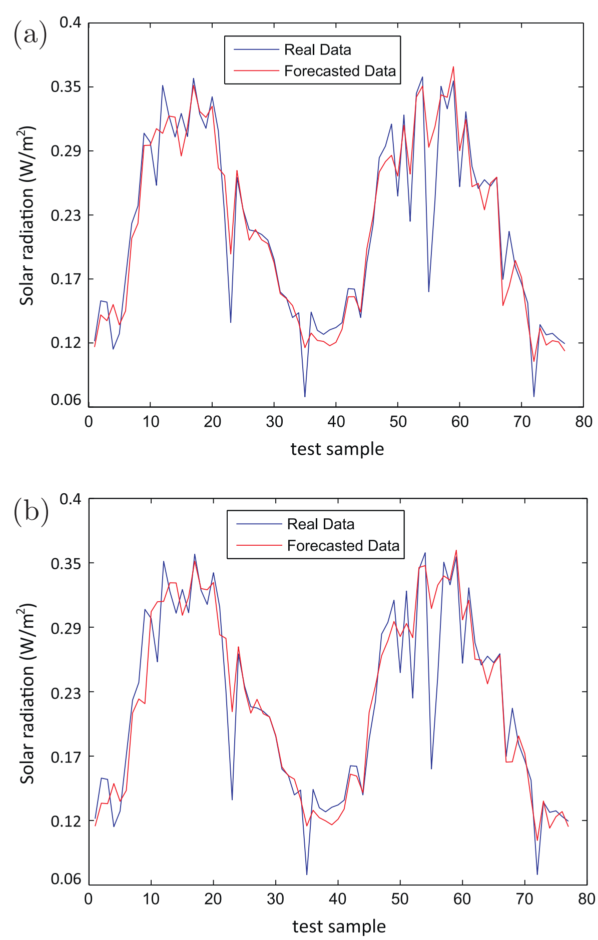

Fig. 2. Prediction obtained with the CRO–ELM and the classical ELM algorithm (random weights); (a) CRO–ELM and (b) classical ELM.

Table 3
Different statistical errors (RMSE, MAE and Bias, in (W/m^2)) of the daily prediction in the test set, by the CRO–ELM approach, for different number of generation in the evolution.

Algorithm	RMSE	Variance	MAE	Variance	Bias	Variance
CRO–ELM (3 gen)	0.0678	0.0028	0.0181	3.4×10^{-5}	4.7×10^{-4}	1.9×10^{-6}
CRO–ELM (5 gen)	0.0663	0.0037	0.0179	2.7×10^{-5}	4.5×10^{-4}	2.0×10^{-6}
CRO–ELM (10 gen)	0.0667	0.0024	0.0184	5.0×10^{-5}	4.7×10^{-4}	1.8×10^{-6}
CRO–ELM (30 gen)	0.0683	0.0019	0.0185	2.3×10^{-5}	4.8×10^{-4}	2.4×10^{-6}

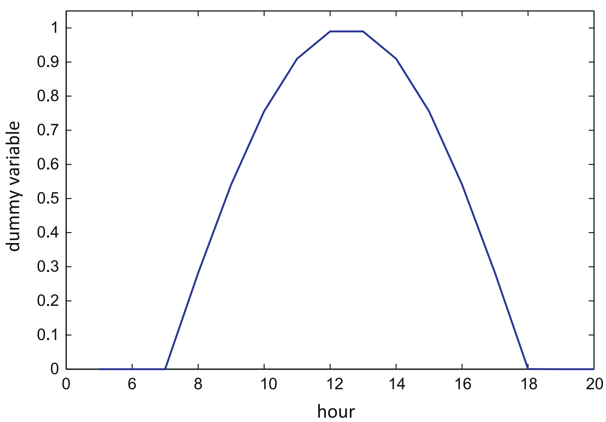

Fig. 3. Dummy variable used to tackle the hourly solar radiation prediction problem with the proposed CRO–ELM.

Table 4
RMSE (W/m^2), hourly radiation prediction in the test set, by the classical ELM and the CRO–ELM approach.

Algorithm	RMSE	Variance
Classical ELM	0.00136	1.8×10^{-6}
CRO–ELM (5 gen)	0.00125	1.6×10^{-6}

with the ELM in this problem showed that 20 neurons in the hidden layer produce good results in terms of error prediction. We therefore set the number of neurons in the hidden layer to 20 in all the experiments carried out.

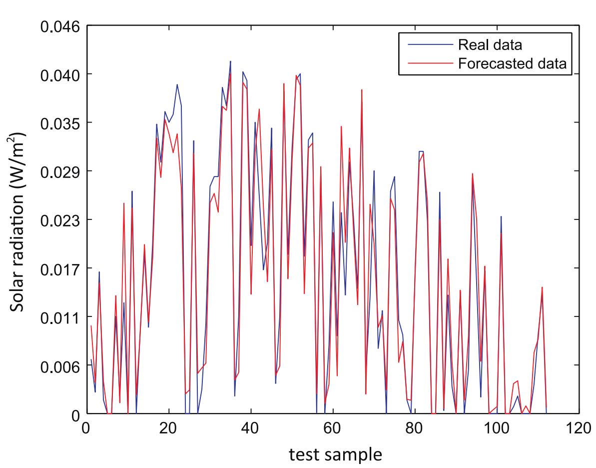

Fig. 4. Prediction obtained with the CRO–ELM in the hourly time horizon case (performance in the test samples, randomly obtained).

Taking into account the 11 input meteorological variables we consider, the encoding of the CRO is a vector of 220 real numbers, in the range $[-1, 1]$.

Each coral in the CRO (encoding a different set of weights in the input layer of the ELM), is evaluated by means of the mean root square error provided by the ELM, in a validation set. First, the available data of global solar radiation (objective) and predictive meteorological variables where divided into a training and test set (80% and 20% of the data, respectively). The train set is divided again in train and validation sets, in order to guide the CRO search. After the CRO evolution, the best coral in the reef (final solution to the problem), is evaluated by the ELM in the test set, and this result is the one reported as result of the CRO–ELM algorithm (see Table 1).

4. Numerical simulations and results

We have carried out different experiments in the data described in Section 2, using the methodology described in Section 3.3. First, we have applied the classical ELM approach without CRO evolution of input weights (i.e. random weights in the input layer). Table 2 shows these classical ELM results, in terms of the mean root square error, where the average of one hundred runs of the algorithm is displayed. This table also shows the results obtained by a Support Vector Regression algorithm (Smola and Schölkopf, 1998), in 10 different permutations of the data. A previous selection of the best hyper-parameters of the SVMr has been carried out in order to optimize the algorithm's performance in this problem. As can be seen, the SVMr approach seems to perform better than the classical ELM in terms of prediction error.

These results can be compared to that of the proposed CRO–ELM, summarized in Table 2. Evolution for 3, 5, 10 and 30 generations in the 10 permutations of the data were carried out and the results obtained have been displayed in this table. Evolution for a larger number of generations led to overfitting and poorer results than in the classical ELM were obtained. As can be seen, the evolution of the CRO–ELM approach during 10 generations obtained the best results, improving those by the SVMr and classical ELM. Micro-evolution for 3 generations improved the results of the classical ELM, but did not reach to the SVMr performance. Evolution for 30 generations of the CRO–ELM approach obtained worse result than applying a small number of generations, what indicates that overfitting is starting to appear. On the other hand, the result obtained with 30 generations was still better than the one obtained with the classical ELM approach.

The performance of the CRO–ELM in terms of prediction can be seen by depicting the real test data against the best CRO–ELM prediction. We can compare this result with the one obtained by the classical ELM approach (random weights). Fig. 2(a) and (b) show these predictions. It can be seen that both predictions are quite adjusted to the real data, and the trend is perfectly captured by both algorithms, though the CRO–ELM approach seems to be slightly more accurate than the classical ELM. Both approaches have difficulties in accurately following sudden radiation peaks, and they are usually underestimated in the predictions obtained (see Table 3).

4.1. Further discussion: application of the model to an hourly solar irradiation prediction problem

The approach presented in this paper is focused on daily solar irradiation prediction. However, in this final analysis, we explore the possibility of reducing the prediction horizon to hourly radiation prediction. The main issue with this is that the input (predictive) variables are daily averaged, so the challenge here is to adapt the system by including an extra variable that models the solar cycle. We introduce for this end a dummy variable, consisting of a sine, adapted to the solar duration of the day. Fig. 3 shows an example of the dummy variable used. In order to show the performance of the system, we consider the 2010 winter months (January, February and March), and discard the days with missing data. A total of 48 days were finally available, where we consider hourly solar radiation prediction. We train the CRO–ELM approach using 80% of the hours, whereas the test set is formed with the remainder 20% of the data, randomly chosen. Table 4 shows a comparison of the classical ELM with the proposed CRO–ELM in this problem. Note that the RMSE is quite low, so the prediction of the solar radiation from the input data is quite good. The CRO–ELM is able to obtain a better prediction than the ELM. Fig. 4 shows the hourly prediction obtained with the proposed CRO–ELM algorithm in the test set. It is possible to see how the prediction given by the CRO–ELM fits really well with the measured solar radiation.

5. Conclusions

This paper presents a novel hybrid Coral Reefs Optimization – Extreme Learning Machine approach for a problem of daily global solar radiation prediction. Different new predictive variables have been considered, both from measurements and also from meteorological models, such as the daily mean aerosol optical depth product, the ozone concentration, estimation of total precipitable water data or the presence of clouds in the zone of study from a numerical weather prediction model. These new predictive variables are completely related to the global irradiation at a given point, so it is intuitive that their inclusion in a prediction model could make it accurate. On the other hand, it has been shown that the novel hybrid CRO–ELM approach proposed is very effective in solving this solar radiation prediction problem, obtaining better results than alternative state-of-the art methods such as classical ELMs or Support Vector Regression algorithms. The testing of the proposed prediction system has been carried out in real predictive and objective data at radiometric station of Murcia (Southern Spain).

Acknowledgements

This work has been partially supported by Spanish Ministry of Economy, under Project No. ECO2010-22065-C03-02. We would like to thank AEMET for providing the observed meteorological data used in this work.

References

Alharbi, M.A., 2013. Daily Global Solar Radiation Forecasting Using ANN and Extreme Learning Machines: A Case Study in Saudi Arabia, Master of Applied Science Thesis, Dalhousie University, Halifax, Nova Scotia.

Anton, M., Loyola, D., Navascues, B., Valks, P., 2008. Comparison of GOME total ozone data with ground data from the Spanish Brewer spectroradiometers. Ann. Geophys. 26, 401–412.

Behrang, M.A., Assareh, E., Ghanbarzadeh, A., Noghrehabadi, A.R., 2010. The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. Sol. Energy 84 (8), 1468–1480.

Benghanem, M., Mellit, A., 2010. 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 35 (9), 3751–3762.

Bhardwaj, S., Sharma, V., Srivastava, S., Sastry, O.S., Bandyopadhyay, B., Chandel, S.S., Gupta, J.R., 2013. Estimation of solar radiation using a combination of Hidden Markov Model and generalized Fuzzy model. Sol. Energy 93, 43–54.

Bilgili, M., Ozoren, M., 2011. Daily total global solar radiation modeling from several meteorological data. Meteorol. Atmosph. Phys. 112 (3–4), 125–138.

Chen, J.L., Liu, H.B., Wu, W., Xie, D.T., 2011. Estimation of monthly solar radiation from measured temperatures using support vector machines – a case study. Renew. Energy 36, 413–420.

Chyzhyk, D., Savio, A., Graña, M., 2014. Evolutionary ELM wrapper feature selection for Alzheimer's disease CAD on anatomical brain MRI. Neurocomputing 128, 73–80.

Dong, H., Yang, L., Zhang, S., Li, Y., 2014. Improved prediction approach on solar irradiance of photovoltaic power station. TELKOMNIKA Indones. J. Electr. Eng. 12 (3), 1720–1726.

Dorvlo, A.S., Jervase, J.A., Al-Lawati, A., 2002. Solar radiation estimation using artificial neural networks. Appl. Energy 71 (4), 307–319.

Han, F., Yao, H.F., Ling, Q.H., 2013. An improved evolutionary extreme learning machine based on particle swarm optimization. Neurocomputing 116, 87–93.

Holben, B.N., Eck, T.F., Slutsker, I., Tanré, D., Buis, J.P., et al., 1998. AERONET-A federated instrument network and data archive for aerosol characterization. Remote Sens. Environ. 66, 1–16.

http://weather.uwyo.edu/upperair/sounding.html (last visit 20.02.14).

Huang, G.B., Chen, L., 2007. Convex incremental extreme learning machine. Neurocomputing 70, 3056–3062.

Huang, G.B., Chen, L., 2008. Enhanced random search based incremental extreme learning machine. Neurocomputing 71, 3460–3468.

Huang, G.B., Zhu, Q.Y., Siew, C.K., 2006. Extreme learning machine: theory and applications. Neurocomputing 70 (1), 489–501.
Huang, G.B., Chen, L., Siew, C.K., 2006. Universal approximation using incremental constructive feedforward networks with random hidden nodes. IEEE Trans. Neural Networks 17 (4), 879–892.
Huang, G.B., Wang, D.H., Lan, Y., 2011. Extreme learning machines: a survey. Int. J. Mach. Learn. Cybernet. 2 (2), 107–122.
Huang, G.B., Zhou, H., Ding, X., Zhang, R., 2012. Extreme learning machine for regression and multi-class classification. IEEE Trans. Syst. Man Cybernet., Part B 42 (2), 513–529.
Inman, R.H., Pedro, H.T., Coimbra, C.F., 2013. Solar forecasting methods for renewable energy integration. Prog. Energy Combust. Sci. 39 (6), 535–576.
Iqbal, M., 1983. An Introduction to Solar Radiation. Academic, San Diego.
Ji, W., Chee, K.C., 2011. Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN. Sol. Energy 85 (5), 808–817.
S.A. Kalogirou, 2014. Designing and Modeling Solar Energy Systems, Solar Energy Engineering, second ed., pp. 583–699 (Chapter 11).
Kanamitsu, M., Alpert, J.C., Campana, K.A., Caplan, P.M., Deaven, D.G., Iredell, M., et al., 1991. Recent changes implemented into the Global Forecast System at NMC. Weather Forecast. 6 (3), 425–436.
Khatib, T., Mohamed, A., Sopian, K., 2012. A review of solar energy modeling techniques. Renew. Sustain. Energy Rev. 16, 2864–2869.
Lahoz, D., Lacruz, B., Mateo, P.M., 2013. A multi-objective micro genetic ELM algorithm. Neurocomputing 111, 90–103.
Landa-Torres, I., Ortiz-García, E.G., Salcedo-Sanz, S., Segovia-Vargas, M.J., Gil-Lṕez, S., Miranda, M., Leiva-Murillo, J.M., Del Ser, J., 2012. Evaluating the internationalization success of companies through a hybrid grouping Harmony Search – Extreme Learning Machine approach. IEEE J. Select. Top. Signal Process. 6 (4), 388–397.
Liou, K.N., 2002. An introduction to atmospheric radiation. Int. Geophys. Ser..
López, G., Batlles, B., Tovar-Pescador, J., 2005. Selection of input parameters to model direct solar irradiance by using artificial neural networks. Energy 30, 1675–1684.
Mellit, A., Kalogirou, S.A., 2008. Artificial intelligence techniques for photovoltaic applications: a review. Prog. Energy Combust. Sci. 34 (5), 574–632.
Mubiru, J., 2008. Predicting total solar irradiation values using artificial neural networks. Renew. Energy 33, 2329–2332.
Rahimikoob, A., 2010. Estimating global solar radiation using artificial neural network and air temperature data in a semi-arid environment. Renew. Energy 35, 2131–2135.
Rehman, S., Mohandes, M., 2008. Artificial neural network estimation of global solar radiation using air temperature and relative humidity. Energy Policy 36 (2), 571–576.
Sahin, M., Kaya, Y., Uyar, M., Yidirim, S., 2014. Application of extreme learning machine for estimating solar radiation from satellite data. Int. J. Energy Res. 38 (2), 205–212.
Salcedo-Sanz, S., Casanova-Mateo, C., Pastor-Sánchez, A., Gallo-Marazuela, D., Labajo-Salazar, A., Portilla-Figueras, A., 2013. Direct solar radiation prediction based on soft-computing algorithms including novel predictive atmospheric variables. Intell. Data Eng. Automated Learn. – IDEAL 2013, Lect. Notes Comput. Sci. 8206, 318–325.
Salcedo-Sanz, S., Del Ser, J., Gil-López, S., Landa-Torres I., Portilla-Figueras, J.A., 2013. The Coral Reefs Optimization Algorithm: A new metaheuristic algorithm for hard optimization problems. In: Proc. of the 15th International Conference on Applied Stochastic Models and Data Analysis, (ASMDA), Mataró, Barcelona.
Salcedo-Sanz, S., Gallo-Marazuela, D., Pastor-Sánchez, A., Carro-Calvo, L., Portilla-Figueras, A., Prieto, L., 2014. Offshore wind farm design with the Coral Reefs Optimization algorithm. Renew. Energy 63, 109–115.
Senkal, O., Kuleli, T., 2009. Estimation of solar radiation over Turkey using artificial neural network and satellite data. Appl. Energy 86 (7-8), 1222–1228.
Smola, A.J., Schölkopf, B., 1998. A tutorial on support vector regression. Stat. Comput..
Sozen, A., Arcakliogblu, E., Ozalp, M., 2004. Estimation of solar potential in Turkey by artificial neural networks using meteorological and geographical data. Energy Convers. Manage. 45, 3033–3052.
Voyant, C., Musellia, M., Paolia, C., Niveta, M.L., 2011. Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation. Energy 36 (1), 348–359.
Wang, D., Alhamdoosh, M., 2013. Evolutionary extreme learning machine ensembles with size control. Neurocomputing 102, 98–110.
Wang, C., Jeong, G.R., Mahowald, N., 2009. Particulate absorption of solar radiation: anthropogenic aerosols vs. dust. Atmosph. Chem. Phys. 9, 3935–3945.
Yacef, R., Benghanem, M., Mellit, A., 2012. Prediction of daily global solar irradiation data using Bayesian neural network: a comparative study. Renew. Energy 48, 146–154.
Zeng, J., Qiao, W., 2013. Short-term solar power prediction using a support vector machine. Renew. Energy 52, 118–127.
Zhu, Q.Y., Qin, A.K., Suganthan, P.N., Huang, G.B., 2005. Evolutionary extreme learning machine. Pattern Recogn. 38 (10), 1759–1763.