Renewable Energy 111 (2017) 63–76

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

A CRO-species optimization scheme for robust global solar radiation statistical downscaling

S. Salcedo-Sanz [a], S. Jiménez-Fernández [a, *], A. Aybar-Ruiz [a], C. Casanova-Mateo [b, c], J. Sanz-Justo [b], R. García-Herrera [d, e]

[a] Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, Spain
[b] LATUV: Remote Sensing Laboratory, Universidad de Valladolid, Valladolid, Spain
[c] Department of Civil Engineering: Construction, Infrastructure and Transport, Universidad Politécnica de Madrid, Madrid, Spain
[d] Department of Astrophysics and Atmospheric Sciences, Universidad Complutense de Madrid, Spain
[e] Department of Sedimentary Geology and Environmental Change, Instituto de Geociencias IGEO, UCM-CSIC, Madrid, Spain

A R T I C L E I N F O

Article history:
Received 22 December 2016
Received in revised form
21 February 2017
Accepted 26 March 2017
Available online 28 March 2017

Keywords:
Coral reefs optimization algorithm
CRO with species
Global solar radiation
Solar energy
Extreme learning machines

A B S T R A C T

This paper tackles the prediction of the global solar radiation (GSR) at a given point, using as predictive variables the outputs of a numerical weather model (the WRF meso-scale model) obtained at a different grid points. Prediction is obtained in this work using a Multilayer Perceptron (MLP) trained with Extreme Learning Machines (ELMs). Provided that the number of WRF outputs is vast, we propose the use of a Coral Reefs Optimization algorithm with species (CRO-SP) to obtain a reduced number of significant predictive variables, therefore improving the global solar radiation prediction attained without feature selection. The proposed system has been tested on real data from a radiometric station located at Toledo (Spain) and average best results of RMSE of 69.19 W/m^2 have been achieved, resulting in a 21.62% improvement over the average prediction without considering the CRO-SP for the feature selection.

1. Introduction

Solar energy is a clean and sustainable renewable source, with a high potential for significant growth in future years. Solar energy development is specially important in the Middle-East region, southern Europe and in the USA, places where the solar resource can be exploited all year around [1]. An important problem faced by this renewable resource is its integration in the grid system, because the energy produced by solar facilities is intrinsically stochastic due to the presence of clouds, atmospheric particles, dust, etc. In order to predict the solar production, an accurate global solar radiation (GSR) at the solar plant is needed, and this radiation depends completely on different atmospheric variables [2–4].

A vast set of different Machine Learning and Artificial Intelligence techniques has been applied for the prediction of global solar radiation [5]. Most of them use different machine learning techniques, with inputs based on meteorological and geographical parameters such as sunshine duration, air temperature, relative humidity, wind speed, wind direction, cloud cover, precipitation, etc [6,7]. According to [8] the design, control and operation of solar energy systems requires long-term series of meteorological data such as solar radiation, temperature, or wind data. Artificial Neural Networks (ANNs) are one of the most applied methods in solar radiation prediction problems. In Ref. [9] an exhaustive review on solar radiation prediction using ANNs is presented, describing forty two different researches, each one using as input variables several of the above-mentioned parameters to predict the solar radiation. In Ref. [10], the authors present a vast set of Artificial Intelligence techniques (ANN, Fuzzy logic, Genetic algorithms, Expert systems, etc.) applied to different photovoltaic applications: sizing of photovoltaic (PV) systems, modeling, simulation and control of PV systems or prediction of PV production using atmospheric or meteorological data. In some of the works cited, also meteorological and geographical parameters are used to increase the accuracy of the systems implemented, such as in Ref. [11], where ANNs are used with geographical parameters to estimate solar energy potential in Turkey. In Ref. [12], different combinations of input variables are considered and tested with Multi-Layer Perceptrons

* Corresponding author. Department of Signal Processing and Communications, Universidad de Alcalá, 28871, Alcalá de Henares, Madrid, Spain.
E-mail address: silvia.jimenez@uah.es (S. Jiménez-Fernández).

http://dx.doi.org/10.1016/j.renene.2017.03.079
0960-1481/

(MLP) and Radial Basis Function (RBF) neural networks. Moreover, results are then compared to conventional GSR prediction models, concluding that the ANN approaches perform better. In a similar research approach [13], carries out a comparison between MLP and RBF neural networks in a problem of solar radiation estimation. Experiments in eight stations in Oman show the good results obtained with the neural algorithms. The work in Ref. [14] presents the performance of ANN in solar energy prediction in of Kuwait. Two different training approaches, gradient descent and Levenberg-Marquart algorithm are tested in five different Kuwaiti locations. Recently, the use of Extreme Learning Machines (ELM) as a fast training method for ANNs has been applied due to the good GSR prediction results obtained. Sahin et al. [15] apply ELMs using satellite measures, concluding that the ELM model performs better than ANN with back-propagation in terms of GSR estimation and computational time. The ELM's performance in alternative radiation prediction systems is also described in Refs. [16–18].

Alternative strong regression algorithms have been applied in solar radiation prediction problems. Support Vector Regression (SVR) is one of these approaches, which has been successfully exploited in solar radiation prediction. In Ref. [19] the SVR algorithm has been mixed with a wavelet transform, in order to improve the performance of the former. Results in an Iranian coastal city have shown the performance of this hybrid proposal. Also with a hybrid algorithm involving SVRs [20], presents a firefly meta-heuristic with a SVR for global solar radiation prediction in different locations of Nigeria. In Ref. [21] solar irradiation mapping is tackled with SVR with exogenous data. Variable selection using a genetic algorithm is also considered in order to improve the performance of the SVR approach. In Ref. [22] a comparison between the performance of SVR and ANNs is carried out, in a problem of photovoltaic power generation. The work in Ref. [26] tackles a problem of global solar energy prediction with SVRs considering different prediction time horizons. Finally, in Refs. [23–25], a least-square SVR has been implemented using meteorological and atmospheric data as predictive variables. Specifically, the results obtained with the least-square SVR are compared to that of an auto-regressive neural network and a RBF neural network.

Bayesian methodology has also been applied in solar energy prediction problems. In Ref. [27] the description of a Bayesian methodology to determine the most relevant meteorological input parameters to an Artificial Neural Network (ANN) is introduced, concluding that there are relevant input parameters (clearness index and relative air mass) to estimate the direct irradiance. [28] also discusses the performance of Bayesian networks in global solar radiation prediction. Specifically, a comparison of GSR prediction performance between Bayesian networks, multi-layer perceptrons and empirical models is carried out. In Ref. [29], several clusters are formed and a prediction model is trained for each cluster to represent a different pattern in the stochastic component of the solar radiation, obtaining better results than ARMA or Time Delay Neural Networks. In Ref. [30], an approach for daily global solar irradiation prediction based on temporal Gaussian processes is discussed, improving the performance of a number of alternative regressors such as neural networks, SVR or regression trees.

Other approaches use all-sky or satellite images in order to obtain solar radiation prediction. In Ref. [31], the prediction of the solar radiation is based on different features extracted from all-sky images, such as the number of cloud pixels, frame difference, gradient magnitude, intensity level, accumulated intensity along the vertical line of the sun or the number of corners in the image. On the other hand [32], and [33] tackle the problem of solar radiation prediction from satellite images, in different locations in Turkey and Australia, respectively.

Hybrid approaches, i.e. algorithms which mix some kind of regression techniques with predictors from different sources have been quite used in solar energy prediction problems. In Ref. [34] solar irradiation in India is analyzed using a hybrid approach that combines Hidden Markov Models and generalized fuzzy models. According to Diagne et al. [35], forecasting of global horizontal irradiance (GHI) can be categorized according to the input variables used, that also determine the forecast horizon where they perform best. For instance, for time horizons from 4 to 6 h, the use of numerical weather prediction (NWP) models typically outperforms satellite-based predictions. Moreover, the use of the Weather Research and Forecasting (WRF) meso-scale model (a regional NWP model) hybridized with a Kalman filter reduces de GHI hour-ahead forecast relative root mean square error (rRMSE) from 35.20% to 22.33% [36]. In Ref. [37] a hybrid approach formed by Time Delay Neural Networks and Auto-regressive Moving Average (ARMA) models is proposed for short-term (hourly) solar radiation prediction in Singapore. A similar approach which also mixes ANN and ARMA models is proposed in Ref. [38]. In this case, the performance of the methodology is evaluated in different location of the Southern French coast and Corsica. In Ref. [39] ANNs are mixed with 2-D linear filters to obtain a hybrid approach for hourly solar radiation prediction. In Ref. [40] ANNs are hybridized with numerical weather prediction models for solving a problem of surface solar irradiance prediction in Brazil. In Ref. [41] a mixed approach formed by an ANN with wavelet transform is proposed for solving a problem of solar irradiance forecasting in 25 different locations of Singapore.

Finally, in Ref. [42] a Grouping Genetic Algorithm (GGA) is mixed with an ELM in a problem of solar radiation prediction at different time horizons. In that work, the GGA determines which WRF output variables best refine the GSR forecast (performed using the above-mentioned ELM approach). In Ref. [43] a hybrid ELM–Coral Reefs Optimization is proposed for solar radiation prediction in Southern Spain, in which the CRO is used to modify the ELM weights in order to improve its performance.

In this paper we propose the use of a novel hybrid approach to optimally predict the GSR at a given location. For this purpose, we use a vast set of meteorological and atmospheric variables provided by a numeric meso-scale model, the WRF, at different points close to the target location under study. Then, a Coral Reefs Optimization with species (CRO-SP) algorithm is used to determine the best subset of WRF variables that lead to a best forecast. This problem is known in the literature as *statistically downscaling* the GSR prediction of a meso-scale model to a given point [44]. Therefore, the ultimate goal of this approach is to evaluate what features (predictive variables) from the numerical model are useful for this forecast. The proposed CRO-SP algorithm is a novel co-evolution algorithm recently described in Ref. [45], especially well-suited for optimization problems with variable-length encodings. In our work, each species in the CRO-SP algorithm represents the use of a different number of WRF variables. Furthermore, the proposed complete solar prediction system (CRO-SP + ELM) will be tested with real data from a radiometric station in Toledo, Spain.

The structure of the rest of the paper is the following: next section presents the problem formulation. Section 3 describes the objective GSR dataset used to train and test the forecast system, and the outputs of the WRF meso-scale model used as input (predictive) variables. Section 4 describes the CRO-SP algorithm that will perform the feature selection and the ELM method for solar radiation downscaling. Section 5 shows the main results obtained with the proposed hybrid approach in a real solar radiation prediction problem at Toledo, Spain. Finally, Section 6 closes the paper giving some final conclusions and remarks.