Energy Conversion and Management 118 (2016) 105–118

Contents lists available at ScienceDirect

Energy Conversion and Management

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

# Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate

S. Belaid [a], A. Mellit [b,c,*]

[a] Unité de Recherche Appliquée en Energies Renouvelables, URAER, Centre de Développement des Energies Renouvelables, CDER, 47133 Ghardaïa, Algeria
[b] Renewable Energy Laboratory, Jijel University, Jijel 18000, Algeria
[c] The International Center of Theoretical Physics, ICTP, Trieste, Italy

ARTICLE INFO

*Article history:*
Received 1 October 2015
Received in revised form 6 March 2016
Accepted 26 March 2016

*Keywords:*
Global solar radiation
Prediction
Arid climate
Simple inputs
Support vector machine
Artificial neural networks

ABSTRACT

Prior knowledge of solar radiation in situ is very important, for better management, sizing and control of solar energy installations. In this paper, an application of a support vector machine (SVM) for the prediction of daily and monthly global solar radiation on horizontal surface in Ghardaïa (Algeria) is presented. Different combinations of measured ambient temperatures, calculated maximum sunshine duration and calculated extraterrestrial solar radiation have been considered for one-step ahead prediction (one day or one month). The obtained results showed a good agreement between measured and predicted global solar radiation data. A comparative study is conducted with the developed neural networks based model and some models published in the literature. The main advantage is that the proposed SVM based models require few simple parameters to get good accuracy.



## 1. Introduction

As reported in [1], about 40 GW of solar PV capacity were installed in 2014. Due to the fast growth of photovoltaic (PV) installation, there is an increasing requirement for more exact and applicable modeling, prediction and forecasting of solar radiation. However, the complexity of the daily behavior of solar radiation in terms of variability and non-stationary as well as its random characteristic, leads to consider more sophistical modeling and prediction techniques, of this meteorological phenomenon, that remains essentially typical in a given region.

Today, the challenge is to develop models that achieve good performances and reduce significantly the errors rate of prediction in different time scales (hourly, daily and monthly), which is an important issue for solar energy applications (Photovoltaic, thermal, etc).

As example, in Algeria the potential of renewable energy is strongly dominated by solar energy source with applications (photovoltaic or thermal) that are very favorable in southern Algerian regions, and therefore, the availability of the measured solar radiation is required to design, size and analyze the performances of such systems [2].

Modeling and forecasting of solar radiation was investigated in different ways, using empirical models, stochastic process (AR, ARMA, ARIMA and SARIMA) and artificial intelligence techniques (neural networks, fuzzy logic, genetic algorithm, etc.) [3]. The most investigated technique is the application of artificial neural networks (ANNs) [4–7]. However, recent studies have demonstrated that Support Vector Machine (SVM) performs better than traditional ANNs in dealing with classification and regression problems due to its good generalization ability [8,9].

Among these works, in [10] the authors presented an application of SVM to estimate the daily global solar radiation (DGSR) in China, different combination in input attributes based on sunshine duration and five empirical sunshine-based models have been used. As results, the SVM models outperform the empirical models. In [11] the authors developed a SVM model to estimate the monthly global solar radiation (MGSR) using seven combinations of air temperatures ($T_{max}$ and $T_{min}$), including linear, polynomial and radial basic function as kernel functions. They find that SVM model with polynomial kernel and $T_{max}$, $T_{min}$ in inputs gives good results. A least-square support vector machine (LS-SVM) based model for short-term solar power prediction (One-hour-ahead) in Denver-USA is proposed in [12], the input of the model includes atmospheric transmissivity in a novel two-dimensional form, sky cover, relative humidity (RH), and wind speed (WS). The output of the model is the predicted atmospheric transmissivity, which is then converted to solar power according to the latitude of the

* Corresponding author at: Renewable Energy Laboratory, Jijel University, Jijel 18000, Algeria.
*E-mail addresses:* adelmellit2013@gmail.com, amellit@ictp.it (A. Mellit).

http://dx.doi.org/10.1016/j.enconman.2016.03.082
0196-8904/

**Nomenclature**

| | |
|---|---|
| Alt | altitude (m) |
| ANN | artificial neural network |
| ANFIS | adaptive Neuro fuzzy inference system |
| BNN | Bayesian neural network |
| $D$ | daily or the number of the day of the year, starting from first January (1–365) |
| $H_0$ | extraterrestrial global solar radiation (Wh/m$^2$/day) |
| $K_t$ | clearness index |
| $L$ | location |
| Lat | latitude (°) |
| Lon | longitude (°) |
| LLR | local linear regression |
| M | monthly mean daily |
| MLP | multi-layer perceptron |
| P | precipitation (mm) |
| $r$ | correlation coefficient |
| $R^2$ | determination coefficient |
| RBF | radial basis function network |
| RH | relative humidity (%) |
| RNN | recurrent neural network |
| $S$ | sunshine duration (h) |
| $S_0$ | potential sunshine duration (h) |
| SKT | skin temperature (surface temperature) |
| $SR_0$ | clear-sky solar radiation (Wh/m$^2$/day) |
| $T_{mean}$ | mean air temperature (°C) |
| $T_{min}$ | minimum air temperature (°C) |
| $T_{max}$ | maximum air temperature (°C) |
| $T_{diff}$ | maximum–minimum air temperature (°C) |
| TCC | total cloud cover |
| TCO | total column of ozone |
| TCW | total column of water vapor |
| VP | vapor pressure (bar) |
| WS | wind speed (m/s) |

site and the time of the day. A LS-SVM based intelligent model to predict the next day's solar insolation in Turkey location with 99.294% accuracy was developed in [13]. The model has five inputs: the number of the day from 1st January (D), daily mean temperature ($DT_{mean}$), daily maximum temperature ($DT_{max}$), daily sunshine duration (S), and DGSR of the day before.

Recently, in [14], the authors developed a hybrid approach by combining the SVM with Wavelet Transform (WT) algorithm for the prediction of daily and monthly horizontal global solar radiation in Iran. The different inputs are; relative sunshine duration ($\frac{S}{S_0}$), difference between maximum and minimum ambient temperatures ($T_{max} - T_{min}$), relative humidity (RH), water vapor pressure (VP), average ambient temperature ($T_{mean}$) and extraterrestrial global solar radiation on a horizontal surface ($H_0$). In [15] the authors developed SVM models to estimate horizontal global solar radiation in Iran based on two input elements of sunshine hours ($S$) and maximum possible sunshine hours ($S_0$) with polynomial and radial basic function as kernel and compared them with the conventional sunshine duration-based empirical models. In [16] authors present a support vector regression (SVR) model, with polynomial and radial basis function (RBF) as the kernel function, for prediction global solar radiation (GSR). Different meteorological data obtained from the only station in the studied region (Iran) were selected as the inputs of the model like; maximum and minimum temperature, the actual duration of sunshine ($S$), maximum possible duration of sunshine ($S_0$), number of days, clear-sky solar radiation ($R_{so}$) and extraterrestrial solar radiation ($H_0$). In [17] the authors evaluated the transferability of SVM to estimate global solar radiation from air temperature in subtropical zone in China and suggest that solar radiation at one site could be well estimated by SVM model developed at another site. The accuracy of estimation is affected by the distance and temperature difference between two sites, and altitude of source site. In [18] the authors described a methodology based on SVM combined with feature selection and genetic algorithms (GAs) that was applied at 14 meteorological stations in Spain, under different climates and on diverse terrain, in order to generate models of DGSR with a high capacity for generalization and selection of only those non co-correlated variables.

The aim of this work is to apply a SVM approach for one step-ahead prediction of horizontal DGSR and MGSR in an arid climate (Ghardaïa), by using simple and available parameters in situ, such as measured ambient temperatures, calculated sunshine duration and extraterrestrial solar radiation. Ghardaïa is considered as an experimental platform for solar energy applications, which is a favorable region for the implementation of photovoltaic and thermal installations due to its important solar energy potential. Thus, prior knowledge of solar radiation in situ is very important, for better management, sizing and control of solar energy installations, particularly in this site where solar radiation models are not well investigated.

This paper is organized as follows: Site and database description are reported in Section 2. Brief theory of support vector machines is presented in Section 3. In Section 4, we explain the methodology and showing the developed SVM models for the prediction of horizontal DGSR and MGSR. The obtained results are discussed and analyzed in Section 5. A comparative study with developed ANN based models for this region and other published models, is presented in Section 6. Finally, some concluding remarks and future perspectives are given in Section 7.

## 2. Site and dataset

### 2.1. Site

Fig. 1 shows the location of the site under consideration (situated in the south of Algeria, about 25 km from Ghardaïa city with latitude = +32.37°, longitude = +3.77° and altitude = 450 m). This region is characterized by an arid climate, during the period from May to October, daily average temperatures are ranging from 25 °C up to 40.8 °C. While, during the period from November to March, the average daily temperatures are lower than 19 °C, with a minimum registered in December, January and February (5 °C during the day).

The measured data used in this study are recorded at the Research Unit for Renewable Energy Applications (RUREA), located at the same site. The data are recorded every 1 h with a high precision by a radiometric station installed at the rooftop of the RUREA, as shown in Fig. 2.

### 2.2. Dataset

The used meteorological dataset consists of daily measured maximum, minimum, mean, maximum-minimum temperatures (°C) (see Fig. 3a) and horizontal daily global solar radiation (Wh/m$^2$/day) (see Fig. 3b).

Fig. 3b presents the variation of horizontal daily solar radiation along three years (2012, 2013 and 2014). It can be observed that