Renewable and Sustainable Energy Reviews 81 (2018) 330–341

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

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

Comparison of four heuristic regression techniques in solar radiation modeling: Kriging method vs RSM, MARS and M5 model tree

Behrooz Keshtegar[a,*], Cihan Mert[b], Ozgur Kisi[c]

[a] Department of Civil Engineering, University of Zabol, P.B. 9861335856, Zabol, Iran
[b] Faculty of Computer Technologies and Eng., International Black Sea University, Tbilisi, Georgia
[c] School of Natural Sciences and Engineering, Ilia State University, Tbilisi, Georgia

ARTICLE INFO

Keywords:
Solar radiation
Kriging
RSM
MARS
M5 model tree
Modeling

ABSTRACT

In this study, four different heuristic regression methods including Kriging, response surface method (RSM), multivariate adaptive regression (MARS) and M5 model tree (M5Tree) have been investigated for accurate estimating of solar radiation with different input data. Monthly solar radiation (SR) from Adana and Antakya stations, which are located in Eastern Mediterranean Region of Turkey is estimated based on the input data of maximum temperature (T_{max}), minimum temperature (T_{min}), sunshine hours (H_s), wind speed (W_s), and relative humidity (RH). In Adana station, the best MARS model provided slightly better accuracy than the Kriging, RSM and M5Tree while the Kriging was found to be the better than the MARS, RSM and M5Tree in Antakya station. The predictions of M5Tree model are shown inaccurate results for both maximum errors and minimum agreement compared to another models. The effect of periodicity input is examined to obtain the accurate predictions of solar radiation for these stations based on the four heuristic –based modeling Kriging, MARS, RSM, M5Tree approaches. Periodicity input data improved the root mean square errors of the best MARS, RSM, M5Tree and Kriging models as 34%, 37%, 46% and 39% for Adana station and by 51%, 47%, 38% and 49% for Antakya station, respectively. The periodic Kriging models performed superior to the periodic MARS, RSM and M5Tree models.

1. Introduction

Energy with its different forms plays vital role in many events. With the increasing population and new technologies, it becomes more vital. Every day in almost every sector we are becoming more technology dependent and consequently this increases our dependency to the energy. Moreover, the exhaustion of traditional energy resources and the environmental issues that caused by them lead many researches to find alternative energy resources. One of the most important alternative energy resources is solar energy. It is important because it is sustainable, renewable, abundant and has unpolluted nature. Therefore, there must be well-equipped correctly structured solar power systems to get more benefit from solar energy. Since solar radiation gives solar availability solar radiation became very important. Due to high cost of the measurement equipment and their maintenance and operation requirements many researches have been conducted based on several developed models. Data availability, accuracy of estimations and simplicity of modeling are the main factors to developer the model approaches for estimating the solar energy.

Artificial Neural Networks (ANNs) have been successfully used in a variety of fields such as pattern recognition, pattern association, function approximation, associative memories and generation of new meaningful pattern in solving complicated problems. Since ANNs are known as an influential correlation technique in several fields, ANNs are mostly preferred and were used by many researchers in different configurations [1–11] to estimate SR and they found promising results. The ability of the ANN was reviewed by Yadav and Chandel [11] and resulted that the accuracy of ANN models was depended on the input parameters. The prediction performance of the ANN model was optimized to obtain the efficient estimation of SR based on learning algorithms and number of hidden neurons for Eastern Mediterranean Region of Turkey [7]. Vakili et al., [12] applied the ANN model for prediction of daily global solar radiation based on the input parameters of particulate matter, relative humidity, wind speed and daily temperature one-year period of Tehran in Iran. Three types of ANN models including Multilayer Perceptron (MLP), Generalized Regression NN (GRNN) and Radial Basis NN (RBNN) were applied for predicting the daily global solar radiation based on meteorological variables that the MLP and RBNN models showed better accuracy than the GRNN [13]. The comparison between ANNs and other modeling methods have

* Corresponding author.
E-mail address: bkeshtegar@uoz.ac.ir (B. Keshtegar).

http://dx.doi.org/10.1016/j.rser.2017.07.054
Received 20 February 2017; Received in revised form 5 July 2017; Accepted 16 July 2017

been investigated by some researches. They tested the accuracy of ANN with the historical methods for real system and concluded that ANN models were better an alternative method [14–19]. The linear model and ANN-based multilayer perceptron were compared using input variables of relative sunshine duration, maximum and minimum temperature, and rainfall for daily global solar radiation in the Province of Salta, Argentina [17]. Modeling procedures including Gene Expression Programming (GEP), Artificial Neural Networks (ANNs), and Adaptive Neuro-Fuzzy were compared for prediction of daily incoming solar radiation in the Basque region of Alava situated in Northern Spain [15]. Support vector regression (SVR) method was also assessed and used to analyze its feasibility and detect its capacity for prediction of SR [20–22]. SVR technique using polynomial and radial basis functions was applied for solar radiation prediction based on meteorological data of Iseyin in Nigeria [23] and Tehran province of Iran [24] while solar radiation was estimated using SVM based on sunshine ratio, maximum and minimum air temperature, relative humidity, and atmospheric water vapor pressure in China [22]. The daily SR of Kerman station in a sunny part of Iran was estimated using SVM with wavelet transform [25]. Olatomiwa et al. [23] showed that the SVR with polynomial basis function can be predicted accurate estimations for SR compared to radial basis functions.

In some mathematical regression models, Ibrahim et al. [26] investigated the relationship between air temperature and solar radiation using linear regression model for Perlis, Northern Malaysia. Jemaa et al. [27] used the linear, quadratic and cubic empirical equations for predictions of monthly and annual average global SR of Troyes-Barberey city in France and found that the linear model gives better predictions for this station. Teke and Yildirim [28] generated linear, quadratic and cubic for predicting global SR in Eastern Mediterranean Region of Turkey and recommended the use of cubic model. Four empirical daily and monthly global SR using meteorological, climatological, and geographical parameters as model input including sunshine-based, cloud-based, temperature-based, and other meteorological parameter were reviewed for Yazd city, Iran and conducted that the El-Metwally sunshine-based model provided a higher accuracy for monthly global SR [29]. However, thirteen empirical global SR models were compared using use daily meteorological data for six stations of Mexico by Quej et al. [30]and showed that Bristow and Campbell model produced the best results.

The adaptive neuro fuzzy inference system (ANFIS) was applied for estimating the SR [31–33]. Mohammadi et al. [32,33] developed a model by ANFIS for prediction of daily global SR based on the most relevant variable input variables for Kerman, Iran [33]. Olatomiwa et al. [34] applied ANFIS methodology for global SR prediction based on the meteorological data of Iseyin, Nigeria. Ramedani et al. [35] compared the SVR and fuzzy linear regression (FLR) for prediction of global SR of Tehran Province, Iran and they found that the SVR performed better than the FLR. ANFIS, ANN, and support vector machine (SVM) were used to predict the daily SR based on measured meteorological variables for the Yucatán Peninsula, México and conducted that SVM technique with model provided a better performance [36]. ANFIS with grid partition and subtractive clustering modeling approaches are applied for predicting the daily global SR radiation in China and are compared with M5 Tree model that the best modeling method was obtained different foe each stations [37]

Mousavi et al. [38] used genetic programming (GP) for prediction of the daily solar radiation with atmospheric pressure, relative humidity, daily air temperature, earth temperature and wind speed as inputs and found that the GP model significantly outperformed the traditional models. Since solar radiation series has linear and nonlinear components, Benmouiza and Cheknane [39] used autoregressive–moving-average (ARMA) and nonlinear autoregressive (NAR) methods and developed a hybrid method combining ARMA model and NAR neural network model to predict global SR. Olatomiwa et al. [40] developed and used a novel hybrid method named SVR with firefly algorithm for global SR prediction using minimum temperature, sunshine duration and maximum temperature as inputs. The results showed that the developed SVR with firefly algorithm provided more accurate predictions than GP and ANN models. Gairaa et al. [41] proposed a hybrid model of ANN and the linear ARMA model in predicting daily global SR of two different climate sites in Algeria and found that the proposed model showed better results than the ARMA and ANN. The extreme learning machine (ELM), genetic programming (GP), SVM, and ANN were employed and compared for global SR using sunshine duration-based, difference temperature-based and multiple parameters-based input data in Shiraz of Iran [42] and conducted the ELM showed accurate predictions. There are several general review papers, which compare existing methods exhaustively [43–45]. It is evident from the related reviews and literature that the use of Kriging and response surface method (RSM) in modeling solar radiation and their comparison with the multivariate adaptive regression spline (MARS) and M5 model tree (M5Tree) methods have not been studied yet.

The purposes of this paper are to i) to investigate the ability of Kriging and RSM for estimating solar radiation, ii) to compare the results with those of the MARS and M5Tree methods and iii) to examine the effect of periodicity component on models' estimation accuracy as input variable. Consequently, this paper is structured to achieve these proposes by the following sections; The monthly input databases in test and train sets are described for two Eastern Mediterranean Regions of Turkey in Section 2. Then, the modeling approaches using the M5Tree, MARS, RSM and Kriging are given in Section 3. The comparative results of different methods for predictions of SR are shown for different input databases in Section 4 that the root mean square error (RMSE), mean absolute error (MAE), mean bias error (MBE), agreement index (d) and Nash Sutcliffe efficiency (NSE) are used to evaluate the accuracy of the four heuristic regression techniques. Finally, the conclusion is presented that the Kriging can be provided accurate predictions compared to other methods and the RSM is more accurate than the M5Tree. The periodicity is remarkably affected on models' accuracy for estimation of SR.

2. Case study

The present study used monthly maximum temperature (T_{max}), minimum temperature (T_{min}), sunshine hours (H_s), wind speed (W_s), relative humidity (RH) and solar radiation (SR) data from Adana (latitude 37.22 °N, longitude 35.40 °E, altitude 20 m) and Antakya (latitude 36.33 °N, longitude 36.30 °E, altitude 100 m) stations which are located in Eastern Mediterranean Region of Turkey. The locations of the stations are shown in Fig. 1. Mediterranean climate has cool, rainy winters and moderately dry and hot summers. With respect to Köppen criteria, the Mediterranean climate is described by winter rainy season (Csa, Csb) and hot, dry, sunny summers and actually, this is exactly the opposite of a monsoon climate [46]. Yearly rainfall in this region ranges from 580 to 1300 mm. According to the distribution of average total radiation over Turkey, Mediterranean

Fig. 1. The location of the Antalya and Antakya stations in Mediterranean Region of Turkey.