Available online at www.sciencedirect.com

ScienceDirect

Solar Energy 115 (2015) 632–644

SOLAR ENERGY

www.elsevier.com/locate/solener

A support vector machine–firefly algorithm-based model for global solar radiation prediction

Lanre Olatomiwa [a,c], Saad Mekhilef [a,*], Shahaboddin Shamshirband [b,*], Kasra Mohammadi [d], Dalibor Petković [e], Ch Sudheer [f]

[a] Power Electronics and Renewable Energy Research Laboratory (PEARL), Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia

[b] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

[c] Department of Electrical & Electronic Engineering, Federal University of Technology, PMB 65, Minna, Nigeria

[d] Faculty of Mechanical Engineering, University of Kashan, Kashan, Iran

[e] University of Niš, Faculty of Mechanical Engineering, Department for Mechatronics and Control, Aleksandra Medvedeva 14, 18000 Niš, Serbia

[f] Department of Civil and Environmental Engineering, ITM University, Gurgaon, Haryana 122017, India

Received 18 September 2014; received in revised form 16 December 2014; accepted 9 March 2015

Communicated by: Associate Editor Christian A. Gueymard

Abstract

In this paper, the accuracy of a hybrid machine learning technique for solar radiation prediction based on some meteorological data is examined. For this aim, a novel method named as SVM–FFA is developed by hybridizing the Support Vector Machines (SVMs) with Firefly Algorithm (FFA) to predict the monthly mean horizontal global solar radiation using three meteorological parameters of sunshine duration ($\bar{n}$), maximum temperature (T_{max}) and minimum temperature (T_{min}) as inputs. The predictions accuracy of the proposed SVM–FFA model is validated compared to those of Artificial Neural Networks (ANN) and Genetic Programming (GP) models. The root mean square (RMSE), coefficient of determination (R^2), correlation coefficient (r) and mean absolute percentage error (MAPE) are used as reliable indicators to assess the models' performance. The attained results show that the developed SVM–FFA model provides more precise predictions compared to ANN and GP models, with RMSE of 0.6988, R^2 of 0.8024, r of 0.8956 and MAPE of 6.1768 in training phase while, RMSE value of 1.8661, R^2 value of 0.7280, r value of 0.8532 and MAPE value of 11.5192 are obtained in the testing phase. The results specify that the developed SVM–FFA model can be adjudged as an efficient machine learning technique for accurate prediction of horizontal global solar radiation.

Keywords: Support vector machine; Firefly algorithm; Hybrid model; Global solar radiation prediction; Meteorological parameters

1. Introduction

The long-term knowledge of solar radiation at any particular locations is essential for variety of areas such as agricultural, hydrological, ecological as well as solar energy applications. It has been proved that the abundant potential of solar energy can play an important role to meet the ever-growing energy demand of the world (Ming et al., 2014; Akikur et al., 2013; Azoumah et al., 2011; Bajpai and Dash, 2012; Hasan et al., 2012). Among different types of renewable resources, solar energy has attracted enormous attention because not only it is

* Corresponding authors. Tel.: +60 146266763 (S. Shamshirband).
E-mail addresses: saad@um.edu.my (S. Mekhilef), shamshirband@um.edu.my (S. Shamshirband).

http://dx.doi.org/10.1016/j.solener.2015.03.015
0038-092X/

Nomenclature

ANFIS	adaptive neuro fuzzy inference system
ANN	Artificial Neural Network
FFA	Firefly Algorithm
FG	fuzzy genetic
GP	Genetic Programming
$\overline{H}$	monthly mean global solar radiation ($MJ/m^2/day$)
$\bar{n}$	monthly mean sunshine duration hour (h)
RBF	radial basis function
RMSE	root-mean-square error
R^2	coefficient of determination
SVM	Support Vector Machine
$\overline{T}_{max}$	monthly mean maximum temperature (°C)
$\overline{T}_{min}$	monthly mean minimum temperature (°C)

sustainable, but also it is abundant and environmental friendly (Akikur et al., 2013). Solar energy exploitation is beneficial in abatement of prevalent global warming, since it does not emit CO_2 or hazardous greenhouse gases. In electricity production, solar radiation study is a prerequisite for design and prediction of energy output of solar conversion system. The best way to obtain solar radiation data is from measurements taken remotely at a particular location using designated measuring instruments; due to required high cost for calibration and maintenance of the instruments, solar radiation data are limited in many meteorological stations around the world (Hunt et al., 1998). The difficulties and uncertainty involve in the measurement of global solar radiation have resulted in development of so many models and algorithms for its estimation from some routinely measured meteorological variables consisting sunshine hour, maximum, minimum and average air temperature, relative humidity, cloud factor, etc. In Nigeria, numerous of the government owned meteorological stations have no record of solar radiation data, even where the record are available there are some missing days or month without record possibly due to improper calibration of measuring equipment employed. Over the past years, a vast number of methods including the empirical models (Angstrom, 1924; Hargreaves and Samani, 1982; Bristow and Campbell, 1984; Besharat et al., 2013; Halawa et al., 2014), satellite-derived model (Pinker et al., 1995; Viana et al., 2011) and stochastic algorithm model (Markov chain) (Hocaoğlu, 2011; Amato et al., 1986; Aguiar et al., 1988) have been developed for estimating the global solar radiation on a horizontal surface. Empirical models have been widely developed and used to correlate the global solar radiation with various routinely measured meteorological and geographical parameters. In many researches, the parameters such as sunshine duration, maximum and minimum temperatures have been recognized as the most proper elements for solar radiation prediction (Besharat et al., 2013; Trnka et al., 2005; Chen and Li, 2013; Wu et al., 2007). However, due to inaccessibility of sunshine duration data in some locations, some studies have proved that good estimations can be attained by using measured maximum and minimum temperature as inputs (Hargreaves and Samani, 1982; Bristow and Campbell, 1984; Liu et al., 2009). Although, application of satellite-based methods seems promising for estimation of solar radiation over a large region, their main drawbacks are the required cost and lack of sufficient historical data because it is relatively new. These methodologies have shown low performance when forecasting/modeling data on long term basis; they are also not suitable when there are some missing data in the database. However, one way to overcome these problems is utilization of artificial intelligence techniques.

In Nigeria, several works have been carried out on predictions of solar radiation using the conventional empirical models (Ezekwe and Ezeilo, 1981; Sambo, 1986; Akpabio and Etuk, 2003; Layi Fagbenle, 1993; Ajayi et al., 2014). Nevertheless, due to necessity of accurate and reliable solar radiation, artificial and computational intelligence techniques have been broadly applied to estimate solar radiation in many regions around the word. Al-Alawi and Al-Hinai (1998) predicted solar radiation for a location with no availability of measured data. They used monthly mean daily values of temperature, pressure, relative humidity, sunshine duration hours and wind speed as inputs for Artificial Neural Networks (ANN) technique to predict global solar radiation. They compared the results with empirical methods model and found more accuracy for ANN-based model. Mellit et al. (2006) employed the combination of neural and wavelet network to forecast daily solar radiation for photovoltaic (PV) sizing application. In their study, wavelets served as activation function. Their results of the forecast demonstrated the more favorable performance of the approach compared to other neural network models. In Jiang (2009), a ANN model was developed to estimate monthly mean daily solar radiation for eight typical cities in China. The achieved results were compared to those of conventional empirical models. The statistical analysis results indicated a good correlation between estimated values by the ANN model and the actual data with higher accuracy than other empirical models.

Behrang et al. (2011) applied particle swarm optimization (PSO) technique to estimate monthly mean daily global solar radiation on a horizontal surface for 17 cities in different regions of Iran. Their results showed better performance of PSO-based models compared to the traditional empirical models. Mohandes (2012) employed PSO