Measurement 76 (2015) 148–155

Contents lists available at ScienceDirect

Measurement

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

Using measured daily meteorological parameters to predict daily solar radiation

Seyyed Mohammad Mousavi [a], Elham S. Mostafavi [b,*], Alireza Jaafari [c], Arefeh Jaafari [d], Fariba Hosseinpour [c]

[a] Department of Geography and Urban Planning, Science and Research Branch, Islamic Azad University, Tehran, Iran
[b] Department of Industrial Engineering, Isfahan University of Technology, Isfahan, Iran
[c] Faculty of Economic and Accounting, Islamic Azad University, Central Tehran Branch (IAUCTB), Tehran, Iran
[d] Faculty of Technology and Engineering, Islamic Azad University, Central Tehran Branch (IAUCTB), Tehran, Iran

ARTICLE INFO

Article history:
Received 28 April 2015
Received in revised form 21 July 2015
Accepted 3 August 2015
Available online 10 August 2015

Keywords:
Daily solar radiation
Machine learning
Genetic programming
Sensitivity analysis
Prediction

ABSTRACT

A major factor for an efficient design of solar energy systems is to provide accurate estimations of the solar radiation. Many of the existing studies are focused on the analysis of monthly or annual solar radiation. This is while less attention has been paid to the determination of daily solar radiation. Accordingly, the main goal of this paper is to develop a robust machine learning approach, based on genetic programming (GP), for the estimation of the daily solar radiation. The solar radiation is formulated in terms of daily air temperature, relative humidity, atmospheric pressure, wind speed, and earth temperature. A comprehensive database containing about 7000 records collected for about 20 years (1995–2014) in a nominal city in Iran is used to develop the GP model. The performance of the derived model is verified using different criteria. A multiple linear regression analysis is performed to benchmark the GP model with a classical technique. The influences of the input variables on the solar energy are evaluated through a sensitivity analysis. The proposed model has a very good prediction performance and significantly outperforms the traditional regression model.

1. Introduction

Renewable energy sources such as solar have emerged as effective alternatives to fossil fuels. Optimal design of solar systems requires exact prediction of the solar energy [1–4]. To this aim, several empirical methods have been developed to avoid performing costly in-situ solar radiation measurements [5,6]. Some of the well-known methods in this area are auto-regression, Markov chain, or robust optimization techniques [7–9]. Among the empirical methods, machine learning has been widely used to solve real world problems [10–28]. Artificial neural networks (ANNs) are well-known machine learning systems that have been utilized to predict the solar radiation [2–4,29–40]. As typical examples in this context, Rahimikhoob [41] successfully applied ANN for predicting the global solar radiation. Alam et al. [42] employed ANN to predict beam solar radiation. Mellit et al. [43] conducted a comprehensive study on application multi-layer perceptron of ANN for estimating total solar radiation data. However, a limitation of ANNs is that they cannot always be converted into explicit forms to provide the details of the prediction process [44].

Genetic programming (GP) [45] is a new approach with notable simulation capabilities. It is a division of genetic algorithms (GA) that generates computer programs rather [44]. For the last decade, GP has been used to formulate

* Corresponding author.
E-mail address: elh.mostafavi@gmail.com (E.S. Mostafavi).

http://dx.doi.org/10.1016/j.measurement.2015.08.004

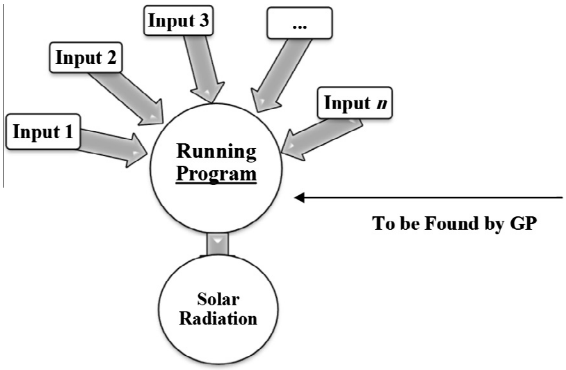

Fig. 1. Input and output relationship in GP process.

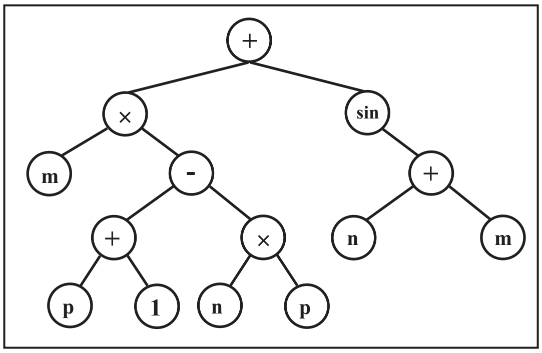

Fig. 2. A typical ET.

complicated engineering problems [44,46–52]. Gene expression programming (GEP) [53,54] is a recent branch of GP evolving programs of various sizes and shapes. Compared to other soft computing methods, it has been barely applied to energy related problems [2,3,55–57]. Although ANNs are used to predict the "daily" solar radiation [41,58–62], none of the existing studies have focused on GP-based analysis of daily solar radiation. Therefore, this study presents the GEP technique for predicting the daily solar radiation in Iran. A regression analysis was later preformed to benchmark the proposed model.

2. Methodology

GP is an optimization technique to create computer programs. The process is inspired by the biological evolution of living organisms [45]. Generally, in GP, the main goal is to find a program that connects inputs to outputs (see Fig. 1) [55]. The solutions derived by conventional GP are shown as tree structures [63]. However, there are other branches of GP that produce models in different shapes (see Fig. 2) [44,55]. Among different variants, linear variants have a higher speed permitting performing many simulations in less time [64–66].

GEP is a linear sub-division of GP creating programs with fixed length of character strings [53]. These solutions are presented as trees and called expression trees (ETs) [44,55]. A typical GEP gene can be as follows:

$$\underline{+}.\times.\sin.m.-.+.+.\times.n.m.p.1.n.p \qquad (1)$$

This equation is called K-expression [44,55]. m, n and p are variables and 1 is a constant. A K-expression can be as an ET as shown in Fig. 2.

More, this gene can be presented as a practical equation:

$$m((p+1)-(n\times p))+\sin(n+m) \qquad (2)$$

Further details about GEP can be found in [44,55].

3. The methodology for predicting daily solar radiation

The soft computing tools commonly follow similar steps to develop a prediction model for the daily solar radiation [3,55,67]. A similar approach was also considered herein.

3.1. GEP-Based formulation for daily solar global radiation

Based on an extensive literature review (e.g., [2–4,41,59,61], seven effective parameters were considered as inputs of the GEP daily solar radiation model. The GEP-based formulation of the daily solar radiation (DS) (kW h/m^2/d) was defined as given below:

$$DS = f(T_{ave}, T_{min}, T_{max}, H, P, W, E) \qquad (3)$$

where,

T_{ave} (°C) = Average of air temperature
T_{min} (°C) = Minimum of air temperature
T_{max} (°C) = Maximum of air temperature
H (%) = Relative humidity
P (kPa) = Atmospheric pressure
W (m/s) = Wind speed
E (°C) = Earth temperature

Thousands of records containing daily meteorological parameters for about 20 years from 1995 to January 30 of 2014 in a nominal city in Iran (Mashhad) was used for the modeling process. This database was obtained from National Meteorological Organization of Iran. The ranges

Table 1
Descriptive statistics of the variables.

Parameter	T_{ave} (°C)	T_{min} (°C)	T_{max} (°C)	H (%)	P (kPa)	W (m/s)	E (°C)	DS (kW h/m^2/d)
Mean	13.42	7.61	19.94	0.45	93.81	3.27	15.53	4.90
Standard Deviation	9.93	8.70	11.07	0.24	2.51	1.36	11.33	2.10
Range	50.52	52.63	53.51	0.93	8.92	8.37	60.62	8.70
Minimum	-16.56	-25.22	-10.20	0.04	88.36	0.49	−22.27	0.10
Maximum	33.96	27.41	43.31	0.97	97.28	8.86	38.35	8.80