Measurement 50 (2014) 255–262

Contents lists available at ScienceDirect

Measurement

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

A least squares support vector machine model for prediction of the next day solar insolation for effective use of PV systems

Betul Bektas Ekici *

Department of Construction Education, Firat University, Elazığ 23119, Turkey

ARTICLE INFO

Article history:
Received 25 September 2013
Received in revised form 27 November 2013
Accepted 6 January 2014
Available online 18 January 2014

Keywords:
Least squares support vector machines
Regression
Prediction
Solar insolation
Temperature

ABSTRACT

Accurate prediction of daily solar insolation has been one of the most important issues of solar engineering. The amount of solar insolation on a given location is a vital data for photovoltaic plants. Systems efficiency is easily affected by the changes in solar radiation so, this study is aimed to develop a Least Squares Support Vector Machine (LS-SVM) based intelligent model to predict the next day's solar insolation for taking measures. Daily temperature and insolation data measured by Turkish State Meteorological Service for three years (2000–2002) were used as training data and the values of 2003 used as testing data. Numbers of the days from 1st January, daily mean temperature, daily maximum temperature, sunshine duration and the solar insolation of the day before parameters have been used as inputs to predict the daily solar insolation. The simulations were carried out with SVM Toolbox of MATLAB software. As a conclusion the results show that LS-SVM is a good method in estimating the amount of solar insolation of a given location with 99.294% accuracy.

1. Introduction

The studies concerned with renewable energy in recent years gives clues of the solution of the problems of global warming and extinction of fossil fuels with solar energy. The sun is the widest source with its huge amount of expansive energy. It is a current working area of recent years because of it advantages like easily accessibility, local applicability, usability without complex technology and its cleanliness. Due to its geographical position between 36° and 42° latitudes, Turkey is abundant with solar energy and has an opportunity to benefit from this endless energy source in building design, developing renewable energy technologies, in agriculture and many other applications [1]. Turkey is divided into four regions according to its solar potential as shown in Fig. 1 [2]. By the way, the amount of total solar radiation and sunshine duration of the seven geographical regions of Turkey are given in Table 1 [2].

Especially photovoltaic systems are one of the most beneficial plants in clean electricity production. The system is directly converts sunlight into electricity so it is easily affected with the changes in the intensity of solar radiation. These fluctuations cause troubles between demand and supply and reduce the power quality. To overcome this important problem the daily solar radiation data of the next day is vital for continuing the systems efficient working and storage the solar power.

The efficient usage of solar energy in a region is directly proportional to the determination of the potential of the region. For solar applications it is hard to predict the same value with empirical methods. Because there are many factors that affect the amount of solar radiation (cloud cover, moisture, etc.) which are generally neglected in most of the solar radiation calculation methods. In some cases the measured values of solar radiation will differ from each other because of the sensibility of the measuring devices. Accurate solar radiation data is required for modeling and designing of solar energy systems like photovoltaic, solar thermal systems and passive solar design applications.

* Tel.: +90 424 2370000x4316
E-mail address: bbektas@firat.edu.tr

0263-2241/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.measurement.2014.01.010

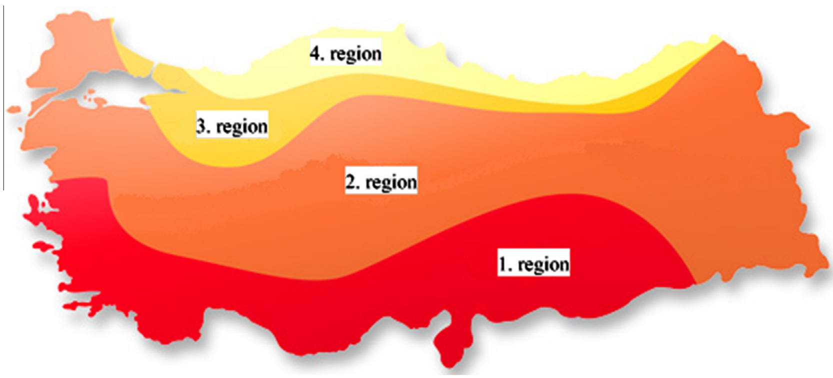

Fig. 1. The solar map of Turkey [2].

Table 1
The solar potential of Turkey's geographical regions [2].

Region	Total solar radiation (kW h/m^2 year)	Sunshine duration (hour/year)
South-Eastern Anatolia	1460	2993
Mediterranean	1390	2956
Eastern Anatolia	1365	2664
Central Anatolia	1314	2628
Aegean	1304	2738
Marmara	1168	2409
Black Sea	1120	1971

For years a great number of studies have been carried out for the estimation of solar energy potential in various locations which are based on conventional physical models or some statistical assumptions [3–16]. However with the development in computer technology, artificial intelligence techniques started to be used for prediction problems of many engineering areas. Several methods have been presented, for estimating the amount of solar radiation with artificial intelligence techniques on a given location [17–31].

In recent years, Support Vector Machines (SVM) has been a popular technique developed by Vapnik [32] and employed in many engineering studies [33–36]. Then, Suykens and Vandevalle [37] proposed a SVM based Least Squares Support Vector Machines (LS-SVM) model. In literature; Zhao et al. [38] proposed a new LS-SVM based prediction algorithm to forecast the actual gas emissions in a coal mine in Shanxi Province. After comparing with other related algorithm they found out that LS-SVM is very effective in gas prediction. Gencoglu and Uyar [39] developed a LS-SVM model regression method in order to form a flashover model of the polluted insulators. They claimed that their proposed method is a strong tool in determining the critical flashover voltage (FOV) and in selecting the insulator type of any region by using the detailed information of the region and electrical transmission system. Esen et al. [40] predicted the efficiency of solar air heater system with double flow aluminum cans absorber plate for a three type collector in Elazig, Turkey by using least squares support vector machines. They achieve 0.0024 RMSE and 0.9997 R^2 value. Baylar et al. [41] employed an intelligent LS-SVM tool for predicting the air entrainment rate and aeration efficiency of weirs. They have obtained a correlation of 0.99 between the predicted and measured values.

This study delineates a LS-SVM based model for predicting the amount of solar insolation values of Elazig city located in the east of Turkey by using the real climatic data obtained from the Turkish State Meteorological Service. The number of the day from 1st January, daily mean temperature $\left(T_{mean} = (\sum_{i=1}^{24}(T_{oi}))/24\right)$, daily maximum temperature, sunshine duration and the insolation of the previous day parameters were used as inputs and the daily insolation as output of the proposed model. MATLAB was employed for LS-SVM applications.

2. Least squares support vector machines

LS-SVM proposed by Suykens et al. [42], is a modificated version of SVM and a more simple technique than SVM [43]. The LS-SVM enables to deal with linear and non-linear multivariable calibration and solves multivariable calibration problems comparatively fast way.

The process of LS-SVM for regression is expressed below. In LS-SVM a linear estimation is done in kernel induced feature space. By considering a data set $\{x_i, y_i\}, \quad i = 1, 2, \ldots N$ with input data $x_i \in R$ and output data $y_i \in R$. While $\phi(.)$ denotes the feature map the regression model can be constituted as follows [37,46,47]:

$$y = \omega^T \cdot \phi(x) + b \tag{1}$$

where ω, is the weight vector of the target function and b is the bias term. As in SVM, it is necessary to minimize a cost function (C) containing a penalized regression error as shown below [48,49]:

$$C = \frac{1}{2}\omega^T \cdot \omega + \frac{1}{2}\gamma\sum_{i=1}^{N} e_i^2 \tag{2}$$

Such that:

$$y_i = \omega^T \cdot \phi(x_i) + b + e_i \quad i = 1, 2, \ldots N \tag{3}$$

The first part of this cost function is a weight decay which is used to regularize weight sizes and penaltize large weights. Due to this regularization, the weights converge