Available online at www.sciencedirect.com

ScienceDirect

Solar Energy 109 (2014) 135–143

SOLAR ENERGY

www.elsevier.com/locate/solener

A comparative study between fuzzy linear regression and support vector regression for global solar radiation prediction in Iran

Zeynab Ramedani [a], Mahmoud Omid [b,*], Alireza Keyhani [b], Benyamin Khoshnevisan [b], Hadi Saboohi [c]

[a] Department of Agricultural Machinery Engineering, Faculty of Agricultural, University of Tabriz, Tabriz, Iran
[b] Department of Agricultural Machinery Engineering, Faculty of Agricultural Engineering and Technology, University of Tehran, Karaj, Iran
[c] Department of Information System, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

Received 23 January 2014; received in revised form 14 July 2014; accepted 13 August 2014
Available online 14 September 2014

Communicated by: Associate Editor Frank Vignola

Abstract

Energy is fundamental to, and plays a prominent role in the quality of life. Sustainable energy is important for the benefits it yields. Sustainable energy technologies are clean sources of energy that have a much lower environmental impact than conventional energy technologies. Among the different forms of clean energy, solar energy has attracted a lot of attention as it is not only sustainable, but is also renewable. Because the number of meteorological stations where global solar radiation (GSR) is recorded is limited in Iran, the aim was to develop three distinctive models in order to prognosticate GSR in Tehran Province, Iran. Accordingly, the fuzzy linear regression (FLR), polynomial and radial basis function (RBF) were applied as the kernel function of support vector regression (SVR). Input energies from different meteorological data obtained from the only station in the study region were selected as the model inputs while GSR was chosen as the model output. Instead of minimizing the observed training error, SVR_poly and SVR_rbf attempted to minimize the generalization error bounds so as to achieve generalized performance. The experimental results show that it is possible to achieve enhanced predictive accuracy and capability of generalization via the proposed approach. The calculated root mean square error and correlation coefficient disclosed that SVR_rbf performed well in predicting GSR compared with FLR.

Keywords: Renewable energy; Solar energy; Artificial intelligence; Support vector machine; Fuzzy modeling

1. Introduction

Energy is fundamental to, and plays a prominent role in the quality of life. With the increasing world population and rising living standards, the global energy demand is steadily increasing (Shamshirband et al., 2014). According to the United Nations, the world population in 2025 could reach 8.5 billion, which is almost five times what it was at the beginning of this century (Kitani, 1999). Nowadays, there is great dependency on abundant and uninterrupted energy supply for living and working, and by 2050, the worldwide energy demand is projected to be at least twice today's level. Global environmental issues and fossil resource exhaustion pose serious problems related to energy consumption. For decades, there has been positive correlation between urban air quality and fossil fuels such as oil, gas, and coal. In other words, the production and consumption of different types of fossil fuels for energy cause significant environmental challenges.

Sustainable energy is important for the benefits it yields. Sustainable energy technologies are clean sources of energy

* Corresponding author.
E-mail address: omid@ut.ac.ir (M. Omid).

http://dx.doi.org/10.1016/j.solener.2014.08.023
0038-092X/

that have a much lower environmental impact than conventional energy technologies. This is because they do not emit any greenhouse gases, making the world a cleaner and safer place, and they can prevent and reduce air emissions as well as water consumption, waste, noise and adverse land-use impacts. Among the different forms of renewable energy regarded as sustainable sources, solar energy has attracted a great deal of attention for it is not only sustainable, but it is also renewable, meaning there is an endless supply (Hove et al., 2014; Janjai et al., 2014; Ming et al., 2014). It is obvious that technical feasibility and economical operation of required technologies for implementing solar energy at a specific location depends on the available solar radiation or solar resource accessibility.

The yearly average solar radiation in Tehran Province – a province in the center of Iran and also the capital – is about 4.92 kW h/m^2/day (RETScreen, 2014). Such abundance of solar energy helps energy policy makers develop solar energy systems (solar power plants and solar heating systems), which are attractive alternatives to traditional power plants that burn fossil fuels such as oil and coal. Also, agricultural systems such as greenhouses, chicken farms, and the dairy industry, are among the largest consumers of heating energy. But to design and implement a solar power system, there is a need for accurate, detailed long-term knowledge of available global solar radiation (GSR) data in various forms (Zeng and Qiao, 2013). In Iran, there is a limited number of meteorological stations where global solar radiation (GSR) is recorded (IRIMO, 2000). Moreover, even at these stations there may be several days when GSR data are missing or lie outside the expected range. On the other hand, GSR and meteorological field analysis can be done experimentally because it is very difficult and time consuming to measure GSR at meteorological stations. Thus, soft programming techniques (Artificial neural Network, Fuzzy-logic, Adaptive-Network-Based Fuzzy Inference System, etc.) may act as powerful tools to analyze and predict GSR. A literature review shows that many researchers have focused on accurate GSR prediction using artificial intelligence or mathematical models (Benghanem and Mellit, 2010; Jiang, 2009; Mostafavi et al., 2013). (Voyant et al., 2014) tested three single methodologies including multi-layer perceptron (MLP), auto-regressive and moving average (ARMA), as well as persistence models in order to forecast GSR. They concluded that the hybridization of the three predictors (ARMA, MLP and persistence) produced better results. In another research study conducted by Wu et al. (2014), it was proposed that a genetic approach combing a multi-model framework be used for solar radiation time series prediction.

Among the popular methodologies in mathematical modeling applied in different engineering fields, fuzzy regression model applications have attracted great interest for prediction problems under uncertainty conditions (Shakouri et al., 2009). In 2004 a fuzzy linear regression model was developed by Al-Kandari et al. for electric load forecasting (Al-Kandari et al., 2004). In 2009, Azadeh et al. formulated a flexible fuzzy regression model to forecast oil consumption based on standard economic indicators (Azadeh et al., 2009). Thus, it was decided to explore the possibility of Fuzzy linear regression for developing a unified correlation to predict GSR in Tehran Province. However, in order to evaluate the power of the developed model, it was decided to compare the obtained results with the predicted results via support vector (SV) regression. A literature review clearly demonstrates that some researchers have already used support vector machine (SVM) and support vector regression (SVR) for developing GSR predictor models (Chen et al., 2013, 2011). Zeng and Qiao (2013) proposed a least-square (LS) support vector machine (SVM)-based model for short-term solar power prediction (SPP) in the USA. The model inputs were historical data on atmospheric transmissivity, sky cover, relative humidity, and wind speed. The model output was the predicted atmospheric transmissivity, which was then converted into solar power according to the site's latitude and time of day. Their results demonstrated that the proposed model not only significantly outperformed a reference autoregressive (AR) model but also achieved superior results to a radial basis function neural network (RBFNN)-based model in terms of prediction accuracy.

Based on the importance of GSR prediction, the aim in this paper is to present the feasibility of applying fuzzy linear regression and radial basis support vector regression to predict global solar radiation based on some simple meteorological data.

2. Material and methods

2.1. Study area and data set

Difficulties in with measuring GSR and the uncertainties of in the measured data have been caused considerable effort to be undertaken into developing procedures and software for the prediction and quality assessment of these such data. Such assessments are needed. This is necessary to ensure that the data selected for various applications are of the highest quality available. Except GSR, other meteorological data are parameters that are routinely recorded at a large number of climatological stations (manned and automatic), due to the low cost of the respective recording instrumentation and the ease of data acquisition.

Tehran province with has an area of 730 square km^2 and was selected as the study area. This province is located in at 35°44′35° latitude North latitudes and 51°30′ longitude West longitude, in the North-central of Iran (Ramedani et al., 2014). Measured daily data from belonging to a seven-year period (1994–2000) was collected from the Islamic Republic of Iran's Meteorological Office data center (IRIMO, 2000). This is the only station in Tehran Province. The yearly average solar radiation in the studied region is 4.92 kW h/m^2/day (RETScreen, 2014). The