Computers and Electronics in Agriculture 118 (2015) 28–37

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Comparison of artificial intelligence techniques via empirical equations for prediction of solar radiation

Hatice Citakoglu *

Faculty of Engineering, Civil Engineering Department, Erciyes University, Kayseri, Turkey

ARTICLE INFO

Article history:
Received 25 April 2015
Received in revised form 18 August 2015
Accepted 19 August 2015
Available online 2 September 2015

Keywords:
Solar radiation
ANN
ANFIS
Empirical equations

ABSTRACT

The most important meteorological variables influencing plant growth are temperature, moisture, and solar radiation. Because it is scarcely gauged at meteorological stations in Turkey, solar radiation is commonly estimated by artificial neural networks (ANN), adaptive network-fuzzy inference system (ANFIS), multiple linear regression (MLR) models, or by empirical equations relating it to available meteorological data at monthly periods composed differently for each one of 12 months. In general, the explanatory meteorological data comprise month number, extraterrestrial radiation, average air temperature, average relative humidity, average sunshine duration, and daylight hours. Such data together with solar radiation measured by the Turkish State Meteorological Service (MGM) at 163 stations having records of at least 20 years are used in monthly units in developing the models. First, as a result of a variance inflation factors analysis, calendar month number (M), extraterrestrial radiation (R_a), average air temperature (T_{mean}), and average relative humidity (RH_{mean}) are determined to be meaningful explanatory variables for estimation of solar radiation. Second, various combinations of input variables are dissected using the ANFIS and ANN models. Next, the accuracies of the ANFIS, ANN, and MLR models, and of Angstrom, Abdalla, Bahel, and Hargreaves–Samani empirical equations are compared with each other. The final results show that the ANN model performs better than the ANFIS and MLR models and the empirical equations in estimating solar radiation in Turkey.

1. Introduction

Solar radiation is the energy source of the hydrologic cycle and is an important meteorological data because it directly affects the growth period and development of plants. Evapotranspiration, which is very important for applications in agriculture and environment, can be calculated as a function of several meteorological data including solar radiation. Data such as solar radiation, air temperature, and relative humidity are necessary for the renewable energy industry, also (Kalogirou et al., 1999). Quantifying regional distribution of solar radiation over a country is valuable for national policies regarding exploitation of solar energy (Fotiadi et al., 2006). Evapotranspiration (ET) is the sum of evaporation and plant transpiration from soil surface to the atmosphere. Reference evapotranspiration (ET_0) is defined as the evapotranspiration of grass of uniform height on a completely shaded ground while having adequate water in the soil. ET_0 is estimated by many empirical or semi-empirical equations using maximum (T_{max}), minimum (T_{min}), and mean air temperatures (T_{mean}), solar radiation (R_s), relative humidity (RH), and wind speed (U_2) as explanatory variables. Because both the Food and Agriculture Organization (FAO) and the World Meteorological Organization (WMO) recommend the Penman–Monteith model for calculating ET_0, it is the most widely used method known by the symbol FAO-56 PM.

Solar radiation strongly controls evaporation from the land surface. According to Llasat and Snyder (1998) small changes in solar radiation have considerable effects on ET_0. Bois et al. (2008) reported that ET_0 calculated by FAO-56 PM in south of France was highly sensitive to solar radiation. Cobaner (2011) used ANFIS models in estimation of ET_0 using solar radiation and air temperature as input data. Citakoglu et al. (2014) found that using only the solar radiation as the input variable gave much better ET_0 estimates than using the wind speed, relative humidity, and air temperature altogether.

There are meaningful relationships depicting solar radiation in terms of some independent variables. For example, Guan et al. (2007) noted that solar energy (as solar radiation), humidity, atmospheric pressure, wind speed, and air temperature are all interrelated. Ododo et al. (1995) showed that solar energy depends, among others, on maximum air temperature. Bandyopadhyay et al. (2008) defined a relationship between solar radiation and

* Address: Faculty of Engineering, Civil Engineering Department, Erciyes University, 38039 Kayseri, Turkey. Tel.: +90 352 43700800; fax: +90 352 4375784.
E-mail address: hcitakoglu@erciyes.edu.tr

http://dx.doi.org/10.1016/j.compag.2015.08.020
0168-1699/

maximum and minimum air temperatures. Hargreaves and Samani (1982) presented a simple equation expressing solar energy as a function of air temperature and position of the sun to take into account extraterrestrial radiation. Their equation, which is Eq. (15) herein, has only one empirical coefficient denoted by a, suggested by them to be in the range: $0.19 \leqslant a \leqslant 0.17$. ElNesr et al. (2015) calculated specific values for a using relevant data at 29 stations in Saudi Arabia and showed that a was inversely proportional to the distance from the coast and to the altitude. Rahimikhoob (2010) used the mean air temperature as the input variable to calculate the solar radiation by ANN models.

Various empirical models have been developed to estimate solar radiation as a function of such meteorological data as sunshine hours, maximum and minimum air temperatures, and relative humidity. Angstrom (1924) proposed the earliest model for estimating solar radiation as a function of sunshine hours. Further, by ANN models, researchers have estimated solar radiation using relevant meteorological data as explanatory variables (e.g., Mubiru and Banda, 2008; Tymvios et al., 2006). Rehman and Mohandes (2008) developed three ANN models for Saudi Arabia using various input combinations, which involved day of the year, air temperature, and relative humidity. They concluded that air temperature and relative humidity are sufficient to estimate solar radiation. Alam et al. (2009) built an ANN model to estimate daily solar radiation in India. Dastorani et al. (2010) used multiple linear regression, multi-layer perceptron (MLP), Elman neural network, and neural network auto-regressive models with exogenous inputs (NNARX), and an adaptive network-fuzzy inference system (ANFIS) models for estimating solar radiation in UK. Dastorani et al. (2010) stated that estimates by the artificial intelligence models showed better results than the empirical formulas. Because ET_0 varies both spatially and temporally, the latitude, longitude, and altitude also have been used as additional input data in the ANN models for estimation of solar radiation (Mohandes et al., 1998; Rehman and Mohandes, 2008; Hontoria et al., 2005; Siqueira et al., 2010). Linares-Rodríguez et al. (2011) selected latitude and longitude, day number, total cloud cover, air temperature, total water vapor, and total ozone column as ANN input variables to estimate solar radiation. They found that prediction by ANN was distance dependent. Fadare et al. (2010) designed multi-layered, feed-forward, back propagation ANN models with different architectures for estimating solar radiation using latitude, longitude, altitude, month, average sunshine duration, average air temperature, and relative humidity as input data. The results of their study indicated that the Fuzzy Genetic model was better than the ANN and ANFIS models. Olatomiwa et al. (2015) used sunshine duration, maximum and minimum air temperatures as input variables to develop Support Vector Machines Firefly Algorithm (SVM–FFA), artificial neural networks (ANN) and Genetic Programming (GP) models for estimating solar radiation. The result of their study showed that the SVM–FFA model was an efficient machine learning approach for estimating solar radiation at the Iranian city of Tabass. Mohammadi et al. (2015a) used ANFIS to estimate solar radiation by day of the year as the only input. They defined that ANFIS model would play a notable role to estimate solar radiation. Mohammadi et al. (2015b) developed Support Vector Machine with Wavelet Transform algorithm (SVM–WT), ANN, and GP to estimate daily and monthly solar radiation in Iranian coastal areas using relative sunshine duration, air temperatures, relative humidity, average temperature and extraterrestrial solar radiation as inputs. Their results indicated that the SVM-WT was the best model. The support vector regression (SVR) methodology was used by Mohammadi et al. (2015c) to estimate solar radiation based on sunshine hours and daylight hours as inputs in Iran. Their results showed that the SVR model was successful for estimation of solar radiation using sunshine hours and daylight hours. Zhang et al. (2015) developed an integrated algorithm (IA) to estimate solar radiation using aerosol optical depth (AOD), and precipitable water (PW), which performed well.

To identify the incorrect global solar radiation values, the clearness index (K_T) was computed and the values which were out of range of $0.015 < K_T < 1$ were eliminated (Jiang, 2009; Mohammadi et al., 2015a,c). It is worth mentioning that clearness index (K_T) is the ratio between global solar radiation incidents on a horizontal surface (R_s) to extraterrestrial horizontal solar radiation on a horizontal surface (R_a) (Mohammadi et al., 2015c). R_a is the extraterrestrial radiation in units of calories on cubic centimeter per day in any geographical location. However, solar attenuation occurs as radiation passes through the atmosphere due to some atmospheric phenomenon such as aerosol extinction, cloud extinction, and Rayleigh scattering. Therefore, in the available solar radiation data all values of R_s should be smaller than R_a, which means $K_T < 1$.

The objective of this study was to develop ANN, ANFIS, and MLR models to estimate solar radiation in Turkey on a monthly basis as a function of month number (M), extraterrestrial radiation (R_a), average temperature (T_{mean}), and average relative humidity (RH_{mean}) using measured data including solar radiation gauged at 163 meteorological stations in Turkey, and further to compare their accuracies in estimating the measured solar radiations with those of four empirical equations.

2. Materials and methods

2.1. Materials

The monthly mean meteorological data used in this study are those measured at 163 stations in Turkey by the Turkish State Meteorological Services (MGM). The record lengths of these measurements vary between 20 and 45 years. The 163 stations are almost uniformly dispersed all over Turkey, which is positioned between the longitudes of 26°E and 45°E and the latitudes of 36°N and 42°N, while the elevations of the stations vary from 3 m to 2400 m above MSL. Fig. 1 shows the locations of these gauging stations.

The total area of Turkey is about 800 thousand km^2 and 36% of it comprises agricultural lands. Due to the versatile and abounding agricultural activities, meteorological data, especially solar radiation data, are necessarily needed for sound estimates of crop water requirements and hence for irrigation water demands. Similar to other meteorological data, solar radiation also reveals distinct variations over seven geographical regions of Turkey. Here, the data measured in these 163 stations are split in two parts. The first part is used in developing the models, while the second part is used for testing the accuracy of the developed models. In Fig. 1, the stations used for the training phases are shown as hollow circles and those used for testing the models are plotted as solid circles.

Statistical parameters of the used data are shown in Table 1. Training and testing segments have similar features. Distributions of daylight hours, sunshine duration, and extraterrestrial radiation data are observed to have kurtosis coefficients of close magnitudes both for training and testing phases. The minimum temperature has a much wider variance than the other variables, which reveals itself with appreciably high variation coefficients both for training and testing data.

2.2. Variable selection

In research studies of similar theme, quite a few potential explanatory variables are used to account for all the influential factors. But, not all studies use so many variables since there are