Computers and Electronics in Agriculture 136 (2017) 71–78

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data

Yu Feng [a,1], Yong Peng [a,1], Ningbo Cui [a,*], Daozhi Gong [b], Kuandi Zhang [c,**]

[a] State Key Laboratory of Hydraulics and Mountain River Engineering & College of Water Resource and Hydropower, Sichuan University, Chengdu, China
[b] State Engineering Laboratory of Efficient Water Use of Crops and Disaster Loss Mitigation/MOA Key Laboratory of Dryland Agriculture, Institute of Environment and Sustainable Development in Agriculture, Chinese Academy of Agriculture Sciences, Beijing, China
[c] College of Water Resources and Architectural Engineering, Northwest A&F University, Yangling, China

ARTICLE INFO

Article history:
Received 16 August 2016
Received in revised form 3 January 2017
Accepted 23 January 2017
Available online 14 March 2017

Keywords:
Reference evapotranspiration
Temperature data
Extreme learning machine
Generalized regression neural network
Hargreaves model

ABSTRACT

Accurate estimation of reference evapotranspiration (ET_0) is essential to agricultural water management. The present study developed two artificial intelligence models for daily ET_0 estimation only with temperature data, including extreme learning machine (ELM) and generalized regression neural network (GRNN) in 6 meteorological stations of Sichuan basin, southwest China, and compared the proposed ELM and GRNN with the corresponding temperature-based Hargreaves (HG) model and its calibrated version considering FAO-56 Penman-Monteith ET_0 as benchmark. Two data management scenarios were evaluated for estimation of ET_0: (1) the models were trained/calibrated and tested using the local data of each station; and (2) the models were trained/calibrated using the pooled data from all the stations and tested in each station. In the first scenario, the results showed that the temperature-based ELM model provided the better estimation than the GRNN, HG and calibrated HG models, with average relative root mean square error (RRMSE) of 0.198, mean absolute error (MAE) of 0.267 mm/d and Nash-Sutcliffe coefficient (NS) of 0.891, respectively. In the second scenario, GRNN model provided the most accurate results among the considered models, with average RRMSE of 0.194, MAE of 0.263 mm/d and NS of 0.895, respectively. Both of the temperature-based GRNN and ELM performed much better than the HG and calibrated HG models for the two scenarios, and the temperature-based GRNN and ELM models are appropriate alternatives for accurate estimation of ET_0 for Sichuan basin of southwest China, which is very helpful for farmers or irrigation system operators to improve their irrigation scheduling.

1. Introduction

As the only term that appears in both water balance and surface energy balance equation (Xu and Singh, 2005), evapotranspiration (ET) is of importance for ecological and hydrological processes, and plays a key role in designing and operating irrigation projects (Abdullah et al., 2015). Its accurate estimation provides valuable information for computation of crop water requirement, development of irrigation scheduling, management of water resources and determination of the water budget (Shiri et al., 2012).

ET can be measured directly by experimental techniques, e.g. eddy covariance systems, lysimeters and Bowen ratio energy balance (Zhang et al., 2013; Kool et al., 2014; Martí et al., 2015), but these methods are complex, costly and not available in many regions (Allen et al., 1998; Ding et al., 2013). Therefore, development of mathematical models for ET estimation is highly needed, which usually relies on reference evapotranspiration (ET_0). The FAO-56 Penman-Monteith (PM) model is recommended as the sole standard method for estimating ET_0 and validating other models (Allen et al., 1998), which requires a number of meteorological variables, including maximum and minimum air temperature, solar radiation, relative humidity and wind speed. However, these meteorological inputs are not commonly available or unreliable, especially in developing countries (Droogers and Allen, 2002; Almorox et al., 2015). According to Shih (1984) and Traore et al. (2010), an ideal method for estimating of ET_0 should be selected based on minimal input data variables without affecting the accuracy of estimation. Thus temperature-based Hargreaves (HG)

* Corresponding author at: College of Water Resource and Hydropower, Sichuan University, 610065 Chengdu, China.
** Corresponding author.
E-mail addresses: fengyu272@163.com (Y. Feng), pengyongscu@foxmail.com (Y. Peng), cuiningbo@126.com (N. Cui), zhangkuandi428@126.com (K. Zhang).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.compag.2017.01.027
0168-1699/© 2017 Elsevier B.V. All rights reserved.

model is an alternative due to its simplicity and high accuracy (Jensen et al., 1997), and can be applied for future estimation of ET_0 using the temperature forecasts (Luo et al., 2014). Allen et al. (1998) recommended the HG model as PM alternative method for ET_0 estimation when the data set of PM model required are not fully available. Almorox et al. (2015) assessed 11 representative temperature-based methods for estimating ET_0, HG model provided the most accurate global average performance in arid, semiarid, temperate, cold and polar climates. But the model usually underestimates ET_0 under high wind conditions (wind speed > 3 m/s) and overestimates under conditions of high relative humidity or at low evapotranspiration rates (Allen et al., 1998; Droogers and Allen, 2002; Xu and Singh, 2002), so a local calibration of HG model is very necessary.

In the last years, artificial intelligence (AI) models have been successfully applied to estimate ET_0 with limited meteorological data. The implementation of AI models in ET_0 estimation was first investigated by Kumar et al. (2002) using artificial neural network (ANN). Later, AAN in modeling ET_0 received much attention from researchers (Trajkovic et al., 2003; Kisi, 2006, 2008; Zanetti et al., 2007; Kim and Kim, 2008; Landeras et al., 2008; Traore et al., 2010; Martí et al., 2011). Kumar et al. (2011) discussed ANN architecture, development, selection of training algorithm and performance criteria for ET_0 estimation. However, ANN requires many data for training, and is easily getting stuck in a local minimum. Some new AI models have been proposed for ET_0 estimation, e.g. support vector machines (Kisi and Cimen, 2009; Tabari et al., 2012), adaptive neuro-fuzzy inference system (Tabari et al., 2012; Pour Ali Baba et al., 2013; Shiri et al., 2011, 2014), generalized neurofuzzy models (Kisi et al., 2012), M5 Model Tree (Kisi, 2016), fuzzy genetic approaches (Kisi and Cengiz, 2013; Kisi, 2013), gene expression programming (Shiri et al., 2014; Martí et al., 2015), and extreme learning machine (ELM) (Abdullah et al., 2015; Patil and Deka, 2016; Feng et al., 2016a).

Sichuan basin is one of the major agricultural regions in China, but seasonal drought happens frequently in this area. Scarcity of water and growing demand for food supplies emphasize on developing improved methods for crop-water estimation (Patil and Deka, 2016). Moreover, real-time irrigation management and water resources allocation is highly need in China for development of precision agriculture. Accurate estimation of ET_0 is crucial to enhance precision irrigation level and increase water use efficiency. The present study aims to investigate the ability of ELM and generalized regression neural network (GRNN) for ET_0 estimation only with temperature inputs, considering two data management scenarios (I) the models were trained and tested using the local data of each station; and (II) the models were trained using the pooled data from all the stations and tested in each station. Further, the ELM and GRNN models were compared against to the well-known empirical Hargreaves models.

2. Material and methods

2.1. Study area and data set

The study area is located in Sichuan basin, with an area of about 0.26 million km^2, a population of 90 million. The well-known Dujiangyan Irrigation Project is located in the centre of Sichuan basin, supplying irrigation water for 0.7 million hm^2 irrigated farmland. The study area has a warm and humid climate, with mean annual air temperature of 17.4 °C and mean annual relative humidity of 79.5%. The meteorological stations and the statistical properties of climatic variables are shown in Table 1.

Daily meteorological variables, including maximum (T_{max}), and minimum (T_{min}) air temperature at 2 m height, mean relative humidity (RH), wind speed at 10 m height (U_{10}), sunshine duration, were obtained at 6 meteorological stations in Sichuan basin (Fig. 1) during 1961–2014. The data with good quality were obtained from the National Climatic Centre of the China Meteorological Administration. Missing daily data which account for about 0.02% of the 6 stations were reconstructed using linear interpolation method.

2.2. FAO-56. Penman-Monteith model

Experimental ET_0 data were unavailable at the study area, therefore ET_0 values estimated by FAO-56 PM model were considered as the targets for ELM, GRNN and Hargreaves models, which is an accepted and very common practice in this situation (Allen et al., 1998; Ladlani et al., 2012; Shiri et al., 2015; Feng et al., 2017). The FAO-56 PM model is expressed as:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma\frac{900}{T_{mean}+273}U_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34U_2)} \quad (1)$$

where ET_0 is reference evapotranspiration (mm/d), R_n is net radiation (MJ/m^2 d), G is soil heat flux density (MJ/m^2 d), T_{mean} is mean air temperature (°C), e_s is saturation vapor pressure, kPa; e_a is actual vapor pressure, kPa; Δ is slope of the saturation vapor pressure function (kPa/°C), γ is psychometric constant (kPa/°C), U_2 is wind speed at 2 m height (m/s).

Due to lack of U_2 and R_s data, these two parameters were estimated by U_{10} and sunshine duration data, respectively (Allen et al., 1998):

$$U_2 = U_{10}\frac{4.87}{\ln(67.8z - 5.42)} \quad (2)$$

$$R_s = \left(a_s + b_s\frac{n}{N}\right)R_a \quad (3)$$

where U_2 is wind speed at 2 m height (m/s), U_{10} is measured wind speed at 10 m height (m/s), z is the height of measurement (10 m), R_s is solar or shortwave radiation (MJ/m^2 d), n is sunshine duration (h), N is maximum possible duration of sunshine or daylight hours (h), R_a is extraterrestrial radiation (MJ/m^2 d), a_s and b_s are constant, with value of 0.25 and 0.50 recommended by FAO-56 (Allen et al., 1998).

2.3. Extremes learning machine

Extreme learning machine (ELM) was first proposed by Huang et al. (2006), and its learning speed can be thousands of times faster than traditional feedforward neural network (FFNN) learning algorithms like back-propagation algorithm while obtaining better generalization performance. For traditional FFNN, all the parameters need to be tuned and thus there exists the dependency between different layers of parameters (weights and biases) (Huang et al., 2006). The main advantage of ELM is that the hidden layer does not need to be tuned and the learning speed is faster than traditional FFNN, in addition to better generalization performance (Abdullah et al., 2015). The ELM model consists of input layer (input variables), hidden layer (neurons), and output layer (ET_0). Different hidden nodes were used to estimate ET_0 (e.g. 10, 20, 30, ..., 100), and we found that with an increase in the number of hidden nodes, the training error reduced significantly in the initial stages and the reduction was not so significant in the later stage when hidden nodes exceeded 50. Thus, 50 hidden nodes have been proved efficient to estimate ET_0. Further details about ELM may be found in Huang et al. (2006), Abdullah et al. (2015), Feng et al., 2016a and Gocic et al. (2016).