Agricultural Water Management 210 (2018) 59–69

Contents lists available at ScienceDirect

Agricultural Water Management

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

Accessible remote sensing data based reference evapotranspiration estimation modelling

Zixiong Zhang[a], Yicheng Gong[a,b], Zhongjing Wang[a,b,c,*]

[a] Department of Hydraulic Engineering, Tsinghua University, Beijing, 100084, China
[b] State Key Lab of Hydroscience and Engineering, Tsinghua University, Beijing, 100084, China
[c] State Key Lab of Plateau Ecology and Agriculture, Qinghai University, Xining, 810016, China

ARTICLE INFO

Keywords:
Reference evapotranspiration
MODIS
Machine learning
Northwest China

ABSTRACT

Estimating reference evapotranspiration (ET_0) is a fundamental requirement of agricultural water management. The FAO Penman–Monteith (FAO-PM) equation has been used as the standard for ET_0 estimation. However, the lack of necessary meteorological data makes it difficult to estimate spatially distributed ET_0 using the FAO-PM method in the wider ungauged areas. In this study, the aim is to explore the methodology for estimating reference evapotranspiration based on remote sensing data. In this method, remote sensing data are combined with machine learning algorithms to establish a model for spatially distributed ET_0 estimation. Three machine learning algorithms were tested, including support vector machine (SVM), back-propagation neural network (BP), and adaptive neuro fuzzy inference system (ANFIS). Results showed this method had good ability in estimating ET_0. Application of the method in Northwest China indicated that the land surface temperature (LST) can be used to accurately estimate ET_0 with high correlation coefficients (r^2 of 0.897–0.915). The surface reflectance has potential for estimating ET_0 with LST and can slightly improve model accuracy based on LST. Evaluation showed LST was more essential than surface reflectance and the model only based on LST had satisfactory performance. This method could be applicability in worldwide with available remote sensing and meteorological data due to the relationship between LST and ET_0.

1. Introduction

Reference evapotranspiration (ET_0) is a major research area of both hydrology and water resources management (Valipour, 2015), especially in irrigation agriculture. ET_0 is defined as an evapotranspiration rate from a hypothetical reference surface that was specific characteristics. The aim is to determine a specific evapotranspiration demand depending on the atmosphere instead of crop type, development and management practices (Allen et al., 1998).

The most direct and important application of ET_0 is in the field of irrigation. ET_0 is the key parameter used to calculate crop evapotranspiration under field conditions. Many studies have shown that a detailed and rational irrigation system can improve irrigation water efficiency (Kang et al., 2002; Cifre et al., 2005; Du et al., 2010). The prerequisite for estimating the optimal water requirement of a crop is to establish a relationship between climatic conditions and evapotranspiration (Maeda et al., 2011). It is generally achieved by multiplying the ET_0 with a crop coefficient k_c, which is specific to a particular crop during a certain growth period, such as sowing, mid-season and harvest (Strong et al., 2017). Therefore, the accuracy of evapotranspiration estimation is the basis of efficient irrigation systems.

In past decades, a variety of approaches have been developed for calculating ET_0 based on simulating fundamental physical principles, such as energy balance and mass transfer (Maeda et al., 2011). Since Allen et al. (1998) established the first version of FAO-56 with the Penman-Monteith equation (FAO-PM), FAO-PM became a standard for calculating reference evapotranspiration (Jabloun and Sahli, 2008; Sentelhas et al., 2010; Djaman et al., 2015). FAO-PM has also been widely used due to its satisfactory results under various climate conditions around the world (Garcia et al., 2004; Cai et al., 2007; Bodner et al., 2007; Liu and Luo, 2010).

However, application of FAO-PM remains inconvenient because it requires a large amount of meteorological data (Hobbins, 2016; Gavilán et al., 2006), which originates from standard metrological observation stations. Reportedly, the number and distribution of metrological stations depends on national or local economic development and the establishment of meteorological observation networks and national data access policies (Hijmans et al., 2005). The FAO-PM method requires air

* Corresponding author at: State Key Lab of Hydroscience and Engineering, Tsinghua University, Beijing, 100084, China.
E-mail address: zj.wang@tsinghua.edu.cn (Z. Wang).

https://doi.org/10.1016/j.agwat.2018.07.039
Received 26 November 2017; Received in revised form 18 July 2018; Accepted 29 July 2018
0378-3774/

temperature, wind speed, relative humidity, solar radiation, etc. However, observations of these parameters at meteorological stations are limited in developing countries (Droogers and Allen, 2002). Even if there is a meteorological station in a specific area, the data collected may be insufficient to represent the regional value of ET_0 due to terrain complexity. As a result, the optimal status of irrigation management in such areas is difficult to achieve, increasing the risk of water shortages and water conflicts (Maeda et al., 2011).

To overcome the existing limits of the FAO-PM model, various attempts aiming to estimate ET_0 with limited observed data have been conducted. A large number of studies have focused on estimating ET_0 using limited ground data. In those studies, a number of methods such as the Hargreaves and Samani equation, Priestley–Taylor equation, and Thornthwaite equation were produced for estimating ET_0 (Jabloun and Sahli, 2008; Sentelhas et al., 2010; Djaman et al., 2015; Tomas-Burguera et al., 2017). The results of these studies showed that these equations, which were based on limited meteorological data, could be used for reference evapotranspiration estimation, where reasonable correlation coefficient and root mean square error values could be obtained. However, the diversity of methods can achieve vastly different accuracies in specific regions (Djaman et al., 2015), which results in a large basis for improper equations. In addition, simple equations are more likely to misestimate, and thus, adjustment of the method coefficient and selection must be carefully conducted before application.

In addition to the improvements in ET_0 equations based on reducing the dependency of climatic data, artificial intelligence approaches were also introduced to develop ET_0 models in a new pattern. Shiri et al. (2012) set up an ET_0 model with limited meteorological data through a genetic programming approach in the Basque Country and attained good model performance. Falamarzi et al. (2014) proposed an ET_0 model based on only temperature and wind speed data using artificial and wavelet neural networks in Australia, which were successful in modelling ET_0. Feng et al. (2016) compared the extreme learning machine (ELM), artificial neural networks optimized by genetic algorithm (GANN), wavelet neural networks (WNN) and empirical models in ET_0 estimation using few meteorological data, and the results showed that the ELM and GANN models achieved acceptable levels of accuracy.

Furthermore, efforts have been aimed at overcoming the shortcomings of a lack of meteorological stations, and thus, a number of valuable approaches use remote sensing data to estimate ET_0, rather than meteorological data. Maeda et al. (2011) used MODIS (Moderate Resolution Imaging Spectroradiometer) LST to replace the atmospheric temperature in temperature-based ET_0 models and achieved a correlation coefficient of 0.67. Alipour et al. (2014) compared M5 model tree and artificial neural network (ANN) for estimating ET_0 using MODIS products, and the coefficient of determination values of the ANN and M5 tree models were over 0.79 and 0.80, respectively. Based on remote sensing data, several algorithms were also developed to estimate the actual evapotranspiration combined with meteorological data, such as SEBAL (Bastiaanssen et al., 1998), S-SEBI (Roerink et al., 2000), SEBS (Su, 2002), METRIC (Allen et al., 2007), TTME (Long and Singh, 2012) and REDRAW (Feng and Wang, 2013). Those studies also hint at the availability of remote sensing data in modelling ET_0.

In this study, we aimed to 1) use machine learning algorithms to establish a model for estimating spatially distributed reference evapotranspiration; 2) study the ways in which accessible remote sensing data can be introduced into the ET_0 estimation model; 3) evaluate the accuracy of proposed ET_0 estimation models based on different combinations of remote sensing data; and 4) analyse model applicability and limitations. Remote sensing data were retrieved from MODIS, which was validated over a large range of land surface conditions and satisfactorily used in scientific studies. The meteorological data used were from nine sparsely distributed national meteorological stations and two eco-hydrological monitoring stations in Northwest China were used.

2. Materials and methodology

2.1. Materials and availability

FAO-PM based ET_0 estimation depends on the parameters of solar radiation, minimum, maximum and mean air temperature, relative humidity, and wind speed, as shown in Eq. (1). These parameters originate from meteorological station observations, which are equipped to measure these parameters. Meteorological station observed data has the advantages of long-term and high time resolution. However, establishing and maintaining these stations is costly, thus, the station distribution depends on population and economic conditions, which are usually inadequate.

Satellite observations have been an essential method for monitoring land surface conditions. These observations have the advantages of wide coverage and providing accessible data for the regions lacking land observations. Frequently used remote sensing products such as MODIS and Landsat could be easily accessed on their respective websites. Although remote sensing data has the disadvantages of a lower temporal resolution than meteorological station data, the daily datasets from MODIS could be applied for ET_0 estimations. MODIS-produced LST data and surface reflectance data are available daily and easy to access. In this study, limited meteorological data were used to calibrate and validate the new ET_0 estimation equation, which is based on MODIS data.

2.2. Modeling structure and approach

2.2.1. Modelling structure

The remote sensing based ET_0 model is designed to select the possible accessible data from remote sensing datasets to determine a relationship between the remote sensing data and ET_0 (calculated by FAO-PM) using machine learning algorithms. The structure of the model is shown in Fig. 1.

Remote sensing data is spatially distributed information, which expresses land surface features in raster format. Remote sensing data is utilized as input in both training and validation sets. Several kinds of remote sensing data were selected from MODIS to establish the ET_0 estimation model, similar to the studies by Maeda et al. (2011) and Alipour et al. (2014). In this study, a wider variety of remote sensing data and machine learning algorithms were analysed to determine better relationships between remote sensing data and ET_0.

Normally, the meteorological data are point based. As the standard of meteorological data based ET_0 estimation, the FAO-PM is expressed as follows and is used to train the remote sensing based ET_0 estimation models as output datasets of the training set.

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma\frac{900}{T+273}u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \tag{1}$$

where ET_0 is the reference evapotranspiration (mm day^{-1}), R_n is net radiation at the crop surface (MJ m^{-2} day^{-1}), G is the ground heat flux density (MJ m^{-2} day^{-1}), T is the mean daily air temperature at 2 m height (°C), u_2 is the wind speed at 2 m height (m s^{-1}), e_s is the saturation vapour pressure (kPa), e_a is the actual vapour pressure (kPa), Δ is the slope vapour pressure curve (kPa °C^{-1}), and γ is the psychrometric constant (kPa °C^{-1}).

The remote sensing data in the training set includes accessible LST data and surface reflectance data, which were retrieved from the point with the same geographical location as the ground meteorological station. Thus, the remote sensing and meteorological data match each other and there is a one-to-one correspondence between the remote sensing and meteorological data.

The input and output datasets of the training set are prepared, and machine learning is conducted to establish a remote sensing based ET_0 estimation model. Machine learning may be used to determine the