Journal of Hydrology 527 (2015) 184–195

Contents lists available at ScienceDirect

Journal of Hydrology

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

Extreme Learning Machines: A new approach for prediction of reference evapotranspiration

Shafika Sultan Abdullah [a,e,*], M.A. Malek [a], Namiq Sultan Abdullah [b], Ozgur Kisi [c], Keem Siah Yap [d]

[a] Department of Civil Engineering, Universiti Tenaga Nasional, Malaysia
[b] Department of Electrical and Computer Engineering, University of Duhok, Iraq
[c] Civil Engineering Department, Architecture and Engineering Faculty, Canik Basari University, Samsun, Turkey
[d] Department of Electronics and Communication Engineering, and College of Graduate Studies, Universiti Tenaga Nasional, Malaysia
[e] Akre Technical Institute, Dohuk Polytechnic University, Dohuk, Iraq

ARTICLE INFO

Article history:
Received 4 February 2015
Received in revised form 27 April 2015
Accepted 29 April 2015
Available online 7 May 2015
This manuscript was handled by Geoff Syme, Editor-in-Chief, with the assistance of John W. Nicklow, Associate Editor

Keywords:
Reference evapotranspiration
Penman–Monteith equation
Extreme Learning Machines
Artificial Neural Networks

SUMMARY

Recognizing the importance of precise determination of reference evapotranspiration (ET_0) is a principal step in the attempts to reserve huge quantities of squandered water. This paper investigates the efficiency of Extreme Learning Machines (ELM) algorithm at predicting Penman–Monteith (P–M) ET_0 for Mosul, Baghdad, and Basrah meteorological stations, located at the north, mid, and southern part of Iraq. Data of weather parameters containing maximum air temperature ($T_{\max}$), minimum air temperature ($T_{\min}$), sunshine hours (R_n), relative humidity (R_h), and wind speed (U_2) for the period (2000–2013) are used as inputs to the ELM model by using four different input cases including complete and incomplete sets of meteorological data. The performance of ELM model is compared with the empirical P–M equation and with feedforward backpropagation (FFBP) model. The evaluation criteria used for comparison are the root of mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2). The statistical results of both models are found to be encouraging; particularly results of running the ELM model with incomplete sets of data, noticing that the sensitivity of the proposed model to missing data changes from one location to another, as well as along the year for certain study location. The R_n is found to be the most effective parameter in Mosul Station, while U_2 and R_h are found to act almost in parallel with R_n in Baghdad Station, and for conditions of Basrah Station; U_2 and R_h prove to be the dominant parameters. The minimum and maximum time intervals required for running ELM model for all stations, and in all applied conditions, are (4.64–6.19) seconds respectively, while the same order of timing required for running the FFBP model is (6.30–27.80) seconds. The maximum R^2 recorded for the ELM model is 0.991, while for the FFBP it is 0.985. The ELM proved to be efficient, simple in application, of high speed, and has very good generalization performance; therefore, this algorithm is highly recommended for locations similar to the geographical and meteorological conditions of Iraq that consists of both arid and semiarid regions.

1. Introduction

Water policy is one of the principal global concerns that aim to control water usage and determine proper ways to minimize the misuse and squandering of available water resources. Evapotranspiration (ET) is a significant variable in the hydrological cycle, and plays a key role in designing and operating irrigation projects; efforts to predict precise ET have been performed by many researchers for nearly a century now, and that is imputed to the early awareness of the importance of water as an essential substance for sustainability of life on this planet. The challenge is further escalated by evolution in urbanization and accelerating growth of population.

The need to study and explain the phenomena of evapotranspiration stimulated the need for another comprehensive expression called the reference evapotranspiration (ET_0). ET_0 is defined as “the rate of evapotranspiration from a hypothetical reference crop with an assumed height of 0.12 m, a fixed crop surface resistance of 70 sm^{-1}, and an albedo of 0.23, closely resembling the evapotranspiration from an extensive surface of green grass cover of uniform height, actively growing, and completely shading the ground with

---

* Corresponding author at: Department of Civil Engineering, Universiti Tenaga Nasional, IKRAM-UNITEN Road, 43000 Kajang, Selangor, Malaysia. Tel.: +60 149341736.

E-mail address: sha_akre@yahoo.com (S.S. Abdullah).

http://dx.doi.org/10.1016/j.jhydrol.2015.04.073
0022-1694/

adequate water". The proposed ET_0 could easily be linked to any kind of crop or at any growth stage through the crop coefficient (K_C) to find the crop evapotranspiration (ET_C) [i.e., $ET_C = ET_0 \times K_C$] (Allen et al., 1998).

Methods for determining ET values are either direct or indirect; direct methods like micrometeorological techniques (Landeras et al., 2008), and Lysimeters that are mostly considered for research purposes due to special construction and maintenance requirements associated with relatively high expenses (Rana and Katerji, 2000; Wright, 1988; Allen et al., 1998). The indirect methods are mathematical models imposed itself as the adequate alternative to the former indirect methods due to several advantages like time and cost saving in addition to ease of application, numerous conventional methods has been developed to simulate the reality of ET process. Thornthwaite method is proposed for calculating the potential evapotranspiration and water balance (Thornthwaite, 1948; Thornthwaite and Mather, 1955; Black, 2007); the Hargreaves equation (Hargreaves and Samani, 1985) is considered as one of the simplest methods, it depends mainly on air temperature and the extra-terrestrial radiation in predicting ET_0; Blaney–Criddle equation (Doorenbos and Pruitt, 1977) is suggested for predicting in locations where air temperature is the only available weather data. Penman (1948) derived an equation for prediction of evaporation from open surfaces by the combination of energy balance with mass transfer methods; despite the criticism of this attempt, the combination method has provided the basis for more advanced future studies, and introduced the resistance factors simulating the aerodynamic resistance, allowing the application of this equation for the estimation of the ET from crop surfaces (Monteith, 1985; Howell and Evett, 2004; Monteith and Unsworth, 2013). The current form of Penman–Monteith (P–M) equation proposed by Allen et al. (1998) represents the integrated combination of the original Penman combination method (Penman, 1948), the modified Penman equation (Doorenbos and Pruitt, 1977), and Monteith equation (Monteith, 1965; Jensen et al., 1990). The experts of Food and Agriculture Organization (FAO) of the United Nations thoroughly investigated the Penman–Monteith (P–M) energy balance/aerodynamic combination equation, and concluded that this equation covers all the factors governing the evapotranspiration process, and overcomes most of the deficiencies identified in the previous empirical equations; accordingly, this equation was adopted by the organization, under the title of FAO P–M equation, and presented as the sole standard method for estimation of ET_0 (Allen et al., 1998).

In the FAO P–M equation, the dominant factors affecting the calculation of ET_0 are the weather variables such as maximum and minimum air temperature; wind speed; sunshine duration; and air humidity measured or estimated at meteorological stations of a specific location, in addition to latitude and altitude that represent the geographic and climatic characteristics of the study location (Allen et al., 1998). Since 1998, the P–M equation has been widely investigated and applied in many parts of the world as a reliable equation for the prediction of ET_0 (Son et al., 2005; Cai et al., 2007; Adeboye et al., 2009; Bakhtiari et al., 2011; Labędzki et al., 2011; Shahrokhnia and Sepaskhah, 2012; Subedi et al., 2013; Kisi, 2014). However, the main identified drawback in the application of FAO P–M equation is the comprehensive weather data required in the estimation process; and the necessity of providing rich historical records of these weather data over a sufficient period, for each study location, to obtain reliable estimates of P–M ET_0, which could be a real hindrance, especially in developing countries; where the number of meteorological stations is limited and weather data records could be scarce (Droogers and Allen, 2002; Shiri et al., 2012; Tabari and Hosseinzadeh Talaee, 2013).

Artificial neural Networks (ANNs) have signified itself as a preferable technique in many fields in this era of rapid technology, and prediction of ET_0 is not exceptional. Many researchers have thoroughly investigated and proved the efficiency of several ANN techniques in predicting ET_0 to overcome the identified limitations and difficulties of traditional methods. Kumar et al. (2002) investigated the application of Multilayer Perception (MLP) of ANNs in estimating ET_0, the standard backpropagation algorithm with learning rates of 0.2 and 0.8, and by employing two different sets of climatic data with different architecture of inputs, hidden, and output layers; the performance of the proposed model in estimating ET_0 is better than the PM equation, in comparison with lysimeter measurements for Davis, California. The potential of using Generalized Regression Neural Network (GRNN) technique for modeling P–M ET_0 is successfully investigated by Kisi (2006), with inputs of daily weather parameters. In the Republic of Korea (Kim and Kim, 2008) developed three types of generalized regression neural networks (GRNNM) models with genetic algorithm (GA), to predict ET_0, the employed statistical coefficients show that the Optimal Combine-GRNNM-GA type provided better results than those of the Extreme-GRNNM-GA and the Average-GRNNM-GA types. Eslamian et al. (2012) examined the capability of ANN and hybrid of ANN with genetic algorithm (ANN + GA) models as tools for modeling P–M ET_0, the performance of the ANN + GA model is proved to be more efficient than the ANN model. The study of Kumar et al. (2010) and Abdulla and Abdul Malek (in press) represent a comprehensive review of most of the researches dealing with predicting ET_0 using ANNs techniques.

The Extreme Learning Machines (ELM) is a fast learning technique with high generalization performance that basically uses Single-hidden Layer Feedforward Neural Networks (SLFNNs) (Huang et al., 2004; Cambria and Huang, 2013), this technique has been successfully applied in various fields of research (Extreme Learning Machines, 2013); but, to the knowledge of the authors, there is no former attempt to apply ELM in the prediction of ET_0, this paper studies the suitability of employing ELM in predicting ET_0 in Iraq, with complete and incomplete weather data, in an attempt to overcome the identified restrictions in the application of P–M equation. Meteorological data records from Mosul, Baghdad, and Basrah Stations in Iraq, over the period 2000–2013, are used in P–M equation and subsequently in the ELM technique; the records include monthly averages of maximum air temperature (T_{max}), minimum air temperature (T_{min}), sunshine hours (R_n), relative humidity (R_h), and wind speed (U_2). The altitude and latitude of each station are also recorded for the identification of study location. The performance of the proposed technique is evaluated against feedforward backpropagation (FFBP) model for assessment of the newly proposed model; using values of ET_0 obtained by traditional P–M equation as reference (Bench Mark). The evaluation criteria are the root of mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2).

2. Study area

The researchers were keen to choose the study area so that it represents the general atmospheric and geographic conditions of Iraq, The study area includes Mosul, Baghdad, and Basrah Stations that are located at the north, mid, and southern part of Iraq respectively. Mosul Station (Ref. No. 680) is located between latitude 36° 19′ 00″ N, longitude 43° 09′ 00″ E, and at altitude of 223.0 m above sea level; Baghdad Station (Ref. No. 650) is located between latitude 33° 18′ 00″ N, longitude 44° 24′ 00″ E, and at altitude of 31.7 m above sea level; and Basrah Station (Ref. No.689) which is located between latitude 30° 31′ 00″ N, longitude 47° 47′ 00″ E, and at altitude of 2.0 m above sea level; as shown in Fig. 1. The weather data consist of T_{max}, T_{min}, U_2, R_h, and R_n of the three main stations for the period (2000–2013) (Ministry of Transportation, 2014).