Agricultural Water Management 149 (2015) 81–90

Contents lists available at ScienceDirect

Agricultural Water Management

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

Modeling reference evapotranspiration with calculated targets. Assessment and implications

Pau Martí[a,*], Pablo González-Altozano[b], Ramón López-Urrea[c], Luis A. Mancha[d], Jalal Shiri[e]

[a] Programa d'Ús Eficient de l'Aigua, Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Edifici Fruitcentre, PCiTAL Parc de Gardeny 25003 Lleida, Spain
[b] Departament d'Enginyeria Rural i Agroalimentària, Universitat Politècnica de València, c/Camí de Vera s/n, 46022 València, Spain
[c] Instituto Técnico Agronómico Provincial (ITAP) and FUNDESCAM, Avda. Gregorio Arcos 19, 02005 Albacete, Spain
[d] Departamento de Hortofruticultura, Centro de Investigación "La Orden-Valdesequera", Junta de Extremadura, Autovía A-V, Km 372, Guadajira, 06187 Badajoz, Spain
[e] Water Engineering Department, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

ARTICLE INFO

Article history:
Received 19 July 2014
Accepted 28 October 2014

Keywords:
Reference evapotranspiration
Calculated targets
Data driven models
Calibration

ABSTRACT

Due to the absence of experimental reference evapotranspiration (ET_o) records, data-driven models consider in most cases calculated ET_o targets to train and test the models, usually according to the standard FAO56 Penman Monteith equation (FAO56-PM). This procedure is also adopted for calibrating more conventional empirical approaches like the well-known Hargreaves (HG) equation. This study aims at assessing the performance implications derived from using calculated targets instead of experimental measurements for training and testing data-driven models or calibrating empirical methods. Therefore an application of a gene expression programming (GEP) based approach for estimating ET_o is presented considering calculated and lysimetric targets fed with two different input combinations and assessed through *k*-fold testing. The same procedure is adopted to evaluate the calibration of the HG equation. Finally, the FAO56-PM and the HG equations are compared with their corresponding GEP models bearing in mind the type of targets used. The locally trained GEP4 and GEP6 models trained using the experimental lysimetric targets are more accurate than the corresponding HG and FAO56-PM equations, assessed using lysimetric benchmarks. The external performance accuracy of GEP4 and GEP6 models decreases considerably in the cross-station approach using experimental targets. In this case, the FAO56-PM and the HG equations might be preferable. The accuracy of the GEP models trained with calculated targets decreases considerably when the performance is assessed using experimental benchmarks. The conclusions drawn when only calculated benchmarks are used might be not sound or even false. The same applies for empirical equations calibrated with calculated targets. Four new GEP-based equations (one per input combination and station) are provided to estimate ET_o.

1. Introduction

The current and expected world population growth must be supported by increases in food production derived from expanded development in arable land, increased cropping intensity, new cropping systems, as well as yield increases and better water use efficiency (Schultz et al., 2005; Bachour et al., 2013). In contrast to the growing dependence on irrigated production, water availability for irrigation has been reduced due to a combination of frequent droughts and competition for water resources among agricultural, individual and urban users. Therefore, sophisticated irrigation water management will be required to optimize water use efficiency and maintain sufficient levels of crop productivity and quality (Ortega Farias et al., 2009). In order to achieve these targets, accurate estimation of evapotranspiration (ET) can be a viable tool to improve the design and management of irrigation programs. ET is a crucial component of the hydrological cycle in agricultural systems, particularly in irrigated systems. Reference evapotranspiration (ET_o) represents the ET from a hypothetical reference surface and was introduced to express the evaporative demand of the atmosphere independent of management practices, crop type and development. Knowledge of the spatiotemporal distribution of ET_o enables the calculation of the required amount of crop water using established crop coefficients.

* Corresponding author. Tel.: +34 902789449.
E-mail address: paumarpe@hotmail.es (P. Martí).

http://dx.doi.org/10.1016/j.agwat.2014.10.028
0378-3774/

ET_o can be measured using lysimeters, a soil water balance approach, eddy covariance methods, a Bowen ratio energy balance system, or a surface renewal method (Ortega Farias et al., 2009). One of the most accurate ways to measure ET_o is by using a lysimeter, which determines ET_o on the basis of the measurement of some of the components of the water balance in a controlled crop area (Gavilán et al., 2006). Nevertheless, lysimeters are unsuitable for monitoring ET as compared to direct climate-based measurement at weather stations. This is not only due to their cost and complexity, but also because the limited area of a typical weather station enclosure does not provide sufficient fetch from a representative surface for these measurements to be meaningful (Sentelhas et al., 2010).

As a result, measured ET_o records are not available in most cases, and the application of mathematical estimation models is the only alternative. This lack of experimental measurements and the increasing availability of more and more reliable standard climatic data through improving networks of meteorological stations have lead to the development of a wide variety of estimation models in the last decades. The choice of a model usually depends on the specific climatic inputs that are available in the test station and on the previous performance of a model in a similar climatic context. The FAO proposed the Penman–Monteith (PM) equation as the standard method for estimating ET_o some years ago (Allen et al., 1998). Nevertheless, this FAO56-PM method requires a lot of climatic inputs for its application which are commonly not available and/or reliable. Therefore, the development of models relying on fewer inputs has turned into an important task, especially in developing countries, where climatic data are scarce and sparse.

As an alternative to the FAO56-PM equation, there are different equations with fewer meteorological parameter requirements, like the well-known Hargreaves (HG) equation (Hargreaves and Samani, 1985). The HG equation requires only measured mean air temperature and temperature range, as well as calculated extraterrestrial radiation. Although accurate daily estimates have been reported, Hargreaves and Allen (2003) stated that the best estimates might be expected for five-day or longer periods, because daily estimates are subject to errors caused by the movement of weather fronts and by large variations in wind speed, and cloud cover. Shuttleworth (1993) even recommended not to use shorter periods than one month. Nevertheless, numerous agricultural and hydrological applications require daily ET_o data. According to Samani and Pessarakli (1986), Shuttleworth (1993) and Di Stefano and Ferro (1997), daily temperature range might be related to relative humidity and cloudiness. Further, advection depends on the interaction between temperature, relative humidity, vapor pressure and wind speed, which can be related to temperature range (Vanderlinden et al., 2004). HG estimates should not be overextended to different climatic conditions unless it has been previously locally calibrated (Samani, 2004). In this regard, the performance of the HG equation and its calibrated versions has been widely assessed in different climatic scenarios. Nevertheless, a complete review of such applications is beyond the scope of the present paper, and might be checked in recent publications, e.g. Mendicino and Senatore (2013), Bachour et al. (2013).

In the last years, an increasing number of data-driven modeling approaches (e.g. artificial neural networks, ANNs, neuro-fuzzy models, ANFIS, gene expression programming based models, GEP) have been successfully applied for estimating target variables in irrigation, including ET_o, as an alternative to conventional models. In particular, in recent years some GEP applications have addressed the modeling of ET_o, among others, Parasuraman et al. (2007), Guven et al. (2008), Shiri and Kisi (2011), Shiri et al. (2012, 2014a,b,c).

Using the principles of genetic programming (GP), GEP was developed by Ferreira (2001a,b). GP was introduced as a generalization of genetic algorithms (GA) (Koza, 1992). The fundamental difference between GP and GA lies in the nature of individuals. In GA individuals are linear strings of fixed length, such as chromosomes; whereas in GP individuals are non-linear entities of different sizes and shapes, such as parse trees. There are two main players in GEP (Ferreira, 2006): chromosomes (which are usually composed of more than one gene of equal length) and expression trees (programs) which are expressions of the genetic information encoded in chromosomes. The chromosomes are composed of multiple genes, each gene encoding a smaller subprogram. Furthermore, the structural and functional organization of linear chromosomes allows for unconstrained operation of important genetic operators, such as mutation, transposition and recombination. The most important advantages of GEP are (i) the chromosomes are simple entities (linear, compact, relatively small and easy to manipulate genetically), and (ii) the expression trees are exclusively the expression of their respective chromosomes. As a result, GEP surpasses the old GP system in 100–10,000 times (Ferreira, 2001a,b).

Due to the absence of experimental ET_o records, data-driven models consider in most cases calculated FAO56-PM ET_o targets to train and test the models. This procedure is also adopted for calibrating more conventional empirical approaches, like the previously mentioned HG equation. As the FAO56-PM equation is recommended for ET_o estimation and validation of other equations in absence of experimental measurements, studies considering FAO56-PM ET_o targets often omit the implications derived from this simplification. Although the soundness of these studies might be only partially affected by this simplification, conclusions should always be drawn bearing this in mind, which is omitted or forgotten in most cases. This study aims at assessing the performance implications derived from this simplification when using calculated targets for training and testing data-driven models or calibrating empirical methods. Therefore, the performance accuracy of two common approaches for ET_o estimation, namely the FAO56-PM and the HG equations, has been compared with their corresponding GEP based models (i.e. fed with the same inputs) considering calculated and experimental benchmarks. Finally, the accuracy of the models trained or calibrated using calculated targets has been also evaluated considering the corresponding actual experimental lysimetric benchmarks.

2. Methods

2.1. Data set

The data series of the climatic variables for this study were obtained from two lysimetric stations of Spain: Las Tiesas (Albacete) and La Orden (Badajoz), Fig. 1. The daily values of maximum (T_{max}), mean (T_{mean}) and minimum (T_{min}) temperature, average wind speed at 2 m height (u_2), maximum (RH_{max}) and minimum (RH_{min}) relative air humidity, solar radiation (R_s) and reference evapotranspiration (ET_o) were measured between January 2007 and December 2012. These years correspond to a climatologically normal period, without sharp or noticeable changes between years. A climatic characterization of the studied stations is given in Table 1 through the annual and global mean and standard deviation of the mentioned climatic parameters for the period 2007–2012.

2.1.1. Albacete station

In this case, the study was performed at "Las Tiesas" experimental farm near Albacete in the CastillaLa Mancha region (South-Eastern Spain), altitude 695 m above sea level, latitude 39°3′ North, longitude 2°5′ West. The surroundings are fully representative of the 110,000 ha of irrigated area in 'La Mancha