Computers and Electronics in Agriculture 150 (2018) 1–4

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Evolutionary algorithm for reference evapotranspiration analysis

Srdjan Jovic*, Blagoje Nedeljkovic, Zoran Golubovic, Nikola Kostic

Faculty of Technical Sciences, Kosovska Mitrovica, University of Priština, 38 220 Kosovska Mitrovica, Kneza Milosa 7, Serbia

ARTICLE INFO

Keywords:
Evolutionary algorithm
Evapotranspiration
Estimation

ABSTRACT

Evapotranspiration of important indicator for management and planning of water resources. It is essential to analyze the evapotranspiration in order to improve water resources planning. The main goal of the study was to analyze the evapotranspiration based on several input parameters. It is important to estimate the influence of the input parameters on the evapotranspiration. For such a purpose evolutionary algorithm was applied. The algorithm applied in this article has space solution of genetic programs. Therefore this methodology is known as genetic programming. The input parameters in the model are monthly minimum and maximum air temperatures, sunshine hours, actual vapour pressure, minimum and maximum relative humidity and wind speed. Results presented in this study could be used for practical application of water resources planning and management based on the input parameters influence on the evapotranspiration.

1. Introduction

Evapotranspiration (ET) represents a complex and non-linear process natural which represent the vital component of the hydrologic cycle and therefore accurate estimation of the ET is very necessary for water resources management, catchment water balance and irrigation systems. Reference evapotranspiration (ET_0) is often used to estimate actual ET. There are different methods and approaches for the reference evapotranspiration. Each of the method has merits and drawbacks.

The single layer Penman-Monteith (PM) method is widely used method for the estimation of evapotranspiration (Srivastava et al., 2018). The accuracy of evapotranspiration estimate relies upon the quality of input weather data. According the results in article (Poon and Kinoshita, 2018) the largest decrease in ET is approximately 13–57 mm per month and is most prominent during the summer (April to September). The observed decrease in ET in article (Poon and Kinoshita, 2018) contributes to understanding of changes in water yield following wildfires, which is of interest for accurately modeling and predicting hydrologic processes in semi-arid landscapes. Optimal Interpolation scheme to generate a reference evapotranspiration was used in article (Tomas-Burguera et al., 2018), forcing meteorological variables, and their respective error variance where, a sensitivity analysis of observational uncertainties and network densification suggests the existence of a trade-off between quantity and quality of observations. Precipitation and reference evapotranspiration are key parameters in hydro-meteorological studies and used for agricultural planning, irrigation system design and management. Precipitation and evaporative demand are expected to be alter under climate change and affect the sustainable development (Pandey and Khare, 2017). Regression analysis in article (Straatmann et al., 2018) showed a close relationship between ET_o measured with atmometers and weather station across locations and years. The Penman-Monteith FAO-56 equation requires the complete climatic records for estimating reference evapotranspiration (ET_o) Mattar, 2018. Reference evapotranspiration has a significant role in agricultural water management. Commonly, FAO56 Penman-Monteith as standard method is used to calculate ET_0 (Khanmohammadi et al., 2017). The main objective of the research (Mehdizadeh et al., 2017) was to investigate the performance of empirical equations and soft computing approaches including gene expression programming (GEP), two types of support vector machine (SVM) namely SVM-polynomial (SVM-Poly) and SVM-radial basis function (SVM-RBF), as well as multivariate adaptive regression splines (MARS) in estimating monthly mean reference evapotranspiration (ET_o) in Iran and it was found that the MARS and SVM-RBF methods generally performed better than GEP and SVM-Poly. Overall, the performance of the MARS and SVM-RBF was better than used empirical equations (Mehdizadeh et al., 2017). The results in article (Feng et al., 2017) showed that during the past 60 years, annual sunshine hours, relative humidity, wind speed and precipitation decreased while temperature increased. ET_0 decreased at a rate of −1.5 mm per decade, or −0.4, −0.7, −0.3, and −0.1 mm per decade in spring, summer, autumn, and winter, respectively (Feng et al., 2017). The main objective of study (Čadro et al., 2017) was to validate and determine, compared to the FAO-PM method, a suitable and reliable

* Corresponding author.
E-mail address: srdjanjovic2016@hotmail.com (S. Jovic).

https://doi.org/10.1016/j.compag.2018.04.003
Received 14 March 2018; Received in revised form 18 March 2018; Accepted 3 April 2018
0168-1699/

alternative ET_0 equations that are requiring less input data and have a simple calculation procedure, with a special focus on Thornthwaite and Turc as methods previously often used in BiH. Relative contributions of climatic variables to ET_0 were temporally unstable and varied considerably in the nine agricultural regions and the whole China; spatial distribution for relative contribution of various climatic variables showed significant diversity among various agricultural regions (Wang et al., 2017). In study (Gao et al., 2017) the daily reference evapotranspiration (ET_{ref}) was estimated through the Penman-Monteith method at 15 meteorological station where it was found that in annual time scale, solar radiation to be the most dominant variable influencing ET_{ref}, however, in seasonal time scale, average air temperature, maximum air temperature, relative humidity were the most dominant factors in spring, summer and winter, and autumn, respectively. Accurate estimation of reference evapotranspiration is vital to hydrological and ecological processes. The FAO-56 Penman–Monteith (PM) model has the higher accuracy for ET_0 estimation, but it requires many meteorological inputs that are not commonly available. Therefore an ideal method is needed requiring as minimal as possible input data variables without affecting the accuracy of estimation (Feng et al., 2017). The study (Pandey et al., 2016) planned to identify a suitable alternative to the FAO-56 Penman-Monteith (FAO56PM) equation for calculating reference evapotranspiration (ET_0) from chosen temperature and radiation based models utilizing monthly meteorological data from 30 destinations. The artificial neural networks (ANN) and the empirical methods of Priestley-Taylor, Makkink, Hargreaves and mass transfer were used to estimate the reference evapotranspiration with daily meteorological data (Antonopoulos and Antonopoulos, 2017). Artificial intelligence methods have potential for reference evapotranspiration estimation since the methods could handle complex and nonlinear data.

In this study reference evapotranspiration is analyzed based on several input parameters which represent weather parameters. There is need to estimated which weather parameters has the highest influence for the reference evapotranspiration. Also there is need to analyze the influence of the combinations of the input parameters on the reference evapotranspiration. The algorithm applied in this article represent one type of artificial intelligence method. The algorithm is based on evolutionary computation or genetic computation. The algorithm in this study space solution of genetic programs hence the algorithm is known as genetic programming (GP). The main aim of this study is to investigate the accuracy of GP approaches for ET_0 modeling.

2. Materials and methods

2.1. Data used

In the investigation meteorological data were used for the evapotranspiration analysis. The data include minimum (T_{min}) and maximum (T_{max}) air temperatures, sunshine hours (n), actual vapour pressure (VP), minimum (RH_{min}) and maximum (RH_{max}) relative humidity and wind speed. Since the evapotranspiration and humidity have pronounced inertia additional inputs are considered as well. The inputs are time lagged evapotranspiration, and time lagged humidity. Data were obtained from the World Bank Database for European Union. The used units are listed as follows:

1. *ETt−1* Time lagged ET_0 for one step
2. *ETt−2* Time lagged ET_0 for two steps
3. *VPt* Vapour pressure
4. *VPt−1* Time lagged VP for one step
5. *RHmint* Minimal relative humidity
6. *RHmint−1* Time lagged RHmin for one step
7. *RHmaxt* Maximal relative humidity
8. *RHmaxt−1* Time lagged RHmax for one step
9. *RHmaxt−2* Time lagged RHmax for two steps
10. *Sn* Sunshine hours
11. *Tmin* Minimal temperature
12. *Tmax* Maximal temperature

Table 1
Input combinations.

		Input parameters											
		1	2	3	4	5	6	7	8	9	10	11	12
Combinations	1	×	×										
	2	×	×	×	×								
	3	×	×			×	×						
	4	×	×	×		×							
	5	×	×					×	×	×			
	6	×	×	×		×		×	×		×		
	7	×	×									×	×
	8	×	×	×							×	×	×

The dataset is divided in three groups, 70% for training, 15% for testing and 15% for validation of the models. Based on the input parameters the reference evapotranspiration (ET_0) is calculated by FAO–56 Penman–Monteith equation (Allen et al., 1998).

There are eight combination from the given data which are created and analyzed by the evolutionary algorithm. The aim is to determine which combination could provide the best prediction of evapotranspiration. These combinations are presented in Table 1.

2.2. Evolutionary algorithm

In this study genetic algorithm was used as evolutionary algorithm. The genetic algorithm applied in this article has space solution of genetic programs. Therefore this methodology is known as genetic programming (GP) Koza, 2007. GP is method which used approach for principle of Darwinian natural selection procedure. As the population in the algorithm fixed-length strings or programs are used. Problem solutions in GP model are expressed with terminals and nodes as tree structure. Nodes represents connection points between the individuals.

The basic flowchart of the GP methodology is represented in Fig. 1. Before starting of the GP methodology there is need to set the terminals which will be used for construction of the function threes. These sets could be found in article (Koza, 2007). Afterward there is need to determine the set of functions, arithmetic operations, Boolean functions, and testing functions. The next step represents fitness measure. The measure express program capability for solving a problem. Therefore the GP search space represents the terminals, functions and fitness functions. There are also control parameters which should be selected like population size and crossover rate. Finally there is need to determine the termination criteria. Tables 2 shows GP parameters which are used in the study.

To estimate the performance of GP models the root-mean-square error (RMSE) and coefficient of determination (R^2):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}}, \tag{1}$$

$$R^2 = \frac{[\sum_{i=1}^{n}(O_i - \overline{O_i})\cdot(P_i - \overline{P_i})]^2}{\sum_{i=1}^{n}(O_i - \overline{O_i})\cdot\sum_{i=1}^{n}(P_i - \overline{P_i})} \tag{2}$$

where n is the total number of data samples and O_i and P_i are the predicted and calculated ET values, respectively.

3. Results

To estimate reference evapotranspiration, two statistical indicators are used for comparison between real and predicted values of reference