Computers and Electronics in Agriculture 143 (2017) 302–313

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Monthly pan-evaporation estimation in Indian central Himalayas using different heuristic approaches and climate based models

Anurag Malik[a], Anil Kumar[a], Ozgur Kisi[b,*]

[a] Department of Soil and Water Conservation Engineering, College of Technology, G.B. Pant University of Agriculture and Technology, Pantnagar 263145, Uttarakhand, India

[b] Faculty of Natural Sciences and Engineering, Ilia State University, Tbilisi, Georgia

ARTICLE INFO

Keywords:
Pan evaporation
Heuristic approaches
Gamma test
Climate-based models

ABSTRACT

Estimation of pan-evaporation has a vital importance in water resources planning and management especially in arid/semi-arid regions. In this study, four heuristic approaches, multi-layer perceptron neural network (MLPNN), co-active neuro-fuzzy inference system (CANFIS), radial basis neural network (RBNN) and self-organizing map neural network (SOMNN) were utilized to estimate monthly pan-evaporation (EP_m) at two locations, Pantnagar and Ranichauri, in the foothills of Indian central Himalayas. The monthly climatic data, minimum and maximum air temperatures, relative humidity in the morning and afternoon, wind speed, sunshine hours and pan-evaporation, were used for model calibration and validation. The combination of appropriate input variables for the applied models was decided using gamma test. The results obtained by MLPNN, CANFIS, RBNN and SOMNN models were compared with climate-based empirical models, such as Stephens-Stewart (SS) and Griffith's (G), on the basis of root mean squared error, coefficient of efficiency and coefficient of correlation. The results indicated that the performance of CANFIS (RMSE = 0.627 mm, COE = 0.936, COC = 0.979) and MLPNN (RMSE = 0.214 mm, COE = 0.989, COC = 0.970) models with six input variables was superior than the others models in estimating monthly pan evaporation at Pantnagar and Ranichauri stations.

1. Introduction

Accurate estimation of pan-evaporation is essential for sustainable water resources planning and management especially in arid and semi-arid regions (Lenters et al., 2005; Dinpashoh, 2006; Sabziparvar et al., 2010; Nourani and Fard, 2012; Malik and Kumar, 2015). Evaporation is a non-linear process which occurs in nature due to temperature differences. Pan evaporation is an index of lake and reservoir, irrigation and potential or reference crop evapotranspiration (Irmak et al., 2002; Shiri et al., 2011). The factors which adversely influence the rate of evaporation are sun-shine hours, air and soil temperature, relative humidity, vapor pressure deficit, atmospheric pressure, and wind speed. Numerous studies have been made by various researchers to estimate the pan-evaporation using empirical or semi-empirical equations from climatic variables (Penman, 1948; Stephens and Stewart, 1963; Griffith's, 1966; Christiansen, 1968; Priestley and Taylor, 1972; Jensen et al., 1990), but the applications of these techniques are often limited. Empirical equations are connected with different climatic factors and affecting rate of evaporation. Singh et al. (1995) developed a relationship between evaporation and meteorological parameters at Hisar (India). They found that the wind velocity, sunshine hours, mean air temperature and solar radiation were positively correlated while relative humidity was negatively correlated with evaporation.

In recent times, the artificial intelligence approaches such as co-active neuro-fuzzy inference system (CANFIS), adaptive neuro-fuzzy inference system (ANFIS) which is hybrid of artificial neural networks (ANN) and fuzzy inference system (FIS), fuzzy-logic (FL), radial basis neural network (RBNN) which is a type of ANN, support vector machines (SVM), generalized regression neural network (GRNN) which is a type of ANN, genetic algorithm (GA), wavelet transformation (WT) and multi-layer perceptron neural network (MLPNN) have been significantly utilized in diverse fields such as modelling daily pan-evaporation (Kisi, 2009, 2013, 2015; Kim et al. 2012; Guven and Kisi, 2013; Kisi and Tombul, 2013; Goyal et al. 2014; Wang et al., 2017a, 2017b). Kisi (2009) applied MLPNN, RBNN, multi linear regression (MLR) and Stephens-Stewart (SS) models to simulate pan-evaporation in California. The results revealed that the MLPNN and RBNN models performed superior to the MLR and SS models. Chang et al. (2010) proposed SOMNN to estimate daily pan evaporation in Taiwan. They compared the results obtained by SOMNN with back propagation

* Corresponding author.
E-mail address: ozgur.kisi@iliauni.edu.ge (O. Kisi).

https://doi.org/10.1016/j.compag.2017.11.008
Received 15 August 2017; Received in revised form 27 September 2017; Accepted 7 November 2017

neural network (BPNN), modified Penman and Penman–Monteith formulas. The results showed that the SOMNN provided better estimates than the BPNN, modified Penman and Penman–Monteith formulas. Kim et al. (2012) investigated the performances of MLPNN, GRNN, SVM, Linacre and multiple linear regression (MLR) models to estimate daily pan-evaporation in Korea and Iran. The results indicated that the MLPNN model performed superior to other models. Tabari et al. (2012) used MLPNN and CANFIS models to predict daily pan-evaporation of semi-arid region in Iran. They reported that the MLPNN models performed superior to the CANFIS models. Guven and Kisi (2013) utilized linear genetic programming (LGP), fuzzy genetic (FG), ANFIS, artificial neural networks (ANN) and SS methods in estimating pan-evaporation of Turkey. They found that the LGP models provided better estimates than the FG, ANFIS, ANN and SS models. Kisi and Tombul (2013) investigated the ability of FG, ANFIS, ANN and SS in modelling pan-evaporation at Turkey, and the results of comparison revealed that the FG models performed better than the others models. Kim et al. (2013) developed MLP, GRNN, ANFIS and MLR models for estimation of daily E_p in south Korea. They concluded that the developed neural network models could be successfully applied. Kisi (2013) used evolutionary neural network (ENN), FG, ANFIS, ANN and SS methods for modelling pan-evaporation from two stations, Antalya and Mersin, in Turkey. He reported that the performace of ENN models was better than the other models. Kisi (2015) utilized LSSVM, multivariate adaptive regression splines (MARS), and M5 Model Tree (M5Tree) in modelling pan-evaporation for Antalya and Mersin stations in Turkey. The results showed that the MARS performed superior to the LSSVM and M5Tree. Malik and Kumar (2015) examined the potential of ANN, CANFIS and MLR models in simulation daily E_p at Pantnagar, India. The results of comparison revealed that the ANN models performed superior to the CANFIS and MLR models. Keshtegar et al. (2016) utilized conjugate gradient, ANFIS and M5Tree for daily pan-evaporation simulation, and reported that the conjugate gradient provided better estimates than the ANFIS and M5Tree models. Wang et al. (2016a) applied the ANFIS with grid partition (ANFIS-GP), ANFIS with subtractive clustering (ANFIS-SC), M5Tree and Angstrom model for estimating daily global solar radiation in China, and the results of comparison revealed the better accuracy of the ANFIS models. Wang et al. (2016b) explored the potential of MLPNN, GRNN, RBNN and improved Bristow-Campbell (IBC) model in predicting the daily global solar radiation from China. The results indicated that the MLP and RBNN models provide better prediction accuracy than the GRNN and IBC models. Wang et al. (2017a) used MLPNN, GRNN, FG, LSSVM, MARS, ANFIS-GP, MLR and SS models to predict pan-evaporation from China. They stated that the soft computing methods outperformed than the MLR and SS models. Wang et al. (2017b) investigated the capability of FG, ANFIS-GP and M5Tree models in estimating EP_m in the Yangtze River Basin, China. The results indicated the better estimation capability of FG model. Wang et al. (2017c) demonstrated the ability of ANFIS-GP, ANFIS-SC and M5Tree techniques in predicting hourly diffuse photosynthetically active radiation in different ecosystems, and results showed that the ANFIS-SC model had more accurate estimates than the other models. Wang et al. (2017d) examined the potential of FG, LSSVR, MARS, M5Tree and MLR models for estimating daily pan-evaporation in the Dongting Lake Basin, China. The results of comparison revealed that the FG and LSSVR models performed better than the other models.

The current study utilized a non-linear modelling and analysis tool i.e. gamma test to identify the appropriate input variables combination for monthly pan-evaporation simulation. Nowadays, several studies reported the utility of gamma test in various fields such as Moghaddamnia et al. (2008) used gamma test for selecting the significant input variables combination for ANN and ANFIS techniques for estimation of daily evaporation in arid and semi-arid regions of Iran. They found that the performances of ANN and ANFIS models was much better than the empirical equations. Piri et al. (2009) utilized ANN and neural network autoregressive exogenous (NNARX) approaches for daily E_p simulation in arid and semi-arid regions of Iran. The appropriate combination of input parameters for ANN and NNARX approaches was decided using gamma test. The results of comparison revealed that the NNARX model performed superior to the ANN and empirical methods. Remesan et al. (2009) applied gamma test to identify the best input variables combination and length of data for rainfall-runoff modelling using local linear regression, ANFIS, NNARX and neuro-wavelet models in Bure catchment, England. They stated that the neuro-wavelet models performed better to the other models. Kakaei Lafdani et al. (2013) used gamma test to identify the most influential input variables for SVM and ANN techniques for estimating suspended sediment load. They have reported that the input configurations selected by utilizing the gamma test for feeding the SVM and ANN models performed better than those identified through the regression technique. Goyal et al. (2014) explored the potential of gamma test to identify the appropriate input variables for ANN, LSSVM, FL, and ANFIS in estimating daily E_p form sub-tropical climates of India. They compared the results obtained by ANN, LSSVM, FL and ANFIS models with the Hargreaves-Samani (HGS) and SS equations. The results indicated that the soft computing methods performed superior to the HGS and SS. Rashidi et al. (2016) applied gamma test-support vector machine (GT-SVM) with the radial basis function (RBF) and GT-SVM with polynomial function in prediction of daily suspended sediment load from Korkorsar basin, Iran. The results of comparison showed that the GT-SVM with RBF kernel model predicted the suspended sediment load better. Malik et al. (2017) examined the potential of CANFIS, MLPNN, MLR, multiple non-linear regression and sediment rating curve techniques in simulation of daily suspended sediment concentration (SSC) from Pranhita river basin, India. The significant combination of the input variables for these techniques were selected using gamma test. The results of the study revealed that the CANFIS model simulated SSC better than the other models.

According to the authors' knowledge, there is not any published work in the literature that compares the accuracy of MLPNN, CANFIS, RBNN and SOMNN in monthly pan-evaporation estimation. This study was conducted with the following objectives: (i) to select appropriate input variables combination for MLPNN, CANFIS, RBNN and SOMNN models using gamma test; (ii) to calibrate and validate the heuristic models with selected input variables; and (iii) to compare the results from MLPNN, CANFIS, RBNN and SOMNN models with those of the climate based models, SS and Griffiths

2. Material and methods

2.1. Study area and data acquisition

This study was conducted at two locations: (i) Pantnagar (79° 38′ 0″E longitude and 29° 0′ 0″N latitude) at an altitude of 243.8 m above mean sea level; and (ii) Ranichauri (78° 24′ 35″E longitude and 30° 18′ 40″N latitude) at an altitude of about 2000 m above mean sea level (Fig. 1). These stations are located in the Indian Central Himalayan region. The mean annual rainfall is about 1400 mm and 1176 mm at Pantnagar and Ranichauri, respectively.

The monthly meteorological data including minimum and maximum air temperatures (T_{min} and T_{max}), relative humidity (RH_1 and RH_2), wind speed (S_w), sun shine hours (H_{ss}), and monthly pan-evaporation (EP_m) were collected from the meteorological observatories located at Crop Research Centre of Pantnagar and Ranichauri stations in Uttarakhand, India. The RH_1 was recorded at 7 AM in the morning and RH_2 was recorded at 2 PM in the afternoon, Indian Standard Time. The measured monthly meteorological data of 27 years (January 1990 to December 2016) for Pantnagar are presented graphically in Fig. 2. Similarly, the measured monthly meteorological data of 13 years (January 2000 to December 2012) for Ranichauri are presented graphically in Fig. 3. The total available data were partitioned into two data sets. Two different splitting rules 50%-50% and 75%-25% were used (a)