



Comparative Assessment of Reference Evapotranspiration Estimation Using Conventional Method and Machine Learning Algorithms in Four Climatic Regions

ALI RAZA,¹ MUHAMMAD SHOAIB,¹ MUHAMMAD ABRAR FAIZ,² FAISAL BAIG,¹ MUDASSER MUNEER KHAN,³
MUHAMMAD KALEEM ULLAH,⁴ and MUHAMMAD ZUBAIR⁵

Abstract—Reference evapotranspiration (ET_o) is considered as an essential component in hydrological and agro meteorological processes. Its accurate estimation becomes an imperative in the planning and management of irrigation practices. ET_o estimation also plays a vital role in improving the irrigation efficiency, water reuse and irrigation scheduling. The conventional physical model of Penmen Montith (PM56) developed by Food and agriculture organization (FAO) has been recommended worldwide for ET_o estimation. This model was firstly used in this study to determine ET_o by using required meteorological data and obtained results used as referenced values. Afterward, five data machine learning algorithms/data driven models, support vector machine (SVM), multilayer perceptron (MLP), group method of data handling (GMDH), general regression neural network (GRNN) and cascade correlation neural network (CCNN), were applied to estimate ET_o values. The climatic data of maximum and minimum temperatures, wind speed, average relative humidity and sunshine hours of six stations from Pakistan was used to train and test data driven model. Data driven models were also applied on other climatic stations without training data which lie in China, New Zealand and USA to further validate and investigate their performance. Comparison results indicated that model efficiency (ME) and correlation coefficient (r) of SVM were obtained (ranges: ME = 95–99%; $r = 0.96$ –1) maximum for all the selected stations. Alternatively, model errors (RMSE = 0.016, MSE = 0.0001 & MAE = 0.08) for SVM were found minimum in comparison to GMDH, MLP, CCNN and GRNN. In addition, all data driven models show enough divergence from hyper arid to high humid climate except SVM which shows almost identical results for all the climatic zones in comparison to standard FAO-PM56 method. Finally, it can be

concluded that SVM could be considered as a reliable alternative method for ET_o estimation among data driven models.

Keywords: Reference evapotranspiration, machine learning algorithms, cross validation, climatic regions.

1. Introduction

The agriculture sector is found to be the huge water consumer in most of the countries. The quantity of water withdrawal from total amount in developing countries is nearly estimated 81% while accounted 71% globally. In addition to use, more than 55% of World's fresh water withdrawals are utilizing for irrigation purposes (Amarasinghe and Smakhtin 2014). It is impossible to yield large amount of food without the irrigation crop field in order to fulfill world's population needs (Lumia et al. 2005). Researches relevant with enhancement of water productivity and improvement of irrigation water management have attained higher importance in irrigation field (Barker et al. 1999). Irrigators and farmers can make real time decision for irrigation on basis of crop water requirements. The amount of water required to irrigate crop field for entire period is called crop water requirements (ET_c) which can be estimated by multiplying crop coefficient (K_c) with reference evapotranspiration (ET_o). Moreover, values of K_c for a particular crop can be determined by dividing crop-ET with the ET_o value (Allen et al. 1998). Therefore, accurate quantification of ET_o is direly needed in determining crop water requirements to overcome less and excessive irrigation problems. As we know, the amount of water given at right time is called irrigation scheduling which seems

¹ Department of Agricultural Engineering, Bahauddin Zakariya University, Multan, Pakistan. E-mail: alir3201@gmail.com; msho127@aucklanduni.ac.nz; faisalbaig@bzu.edu.pk

² College of Water Conservancy and Civil Engineering, Northeast Agricultural University, Harbin, China. E-mail: abrar_faiz@yahoo.com

³ Department of Civil Engineering, Bahauddin Zakariya University, Multan, Pakistan. E-mail: mudasserkhan@bzu.edu.pk

⁴ Department of Civil Engineering, The University of Lahore, Lahore, Pakistan. E-mail: muhammad.kaleem1@ce.uol.edu.pk

⁵ Department of Civil Engineering, NFC Institute of Engineering and Technology, Multan, Pakistan. E-mail: engineerzubair49@gmail.com

impossible without accurate estimation of ET_o . Different techniques and methods have been applied for ET_o estimation which include direct, indirect and data driven methods (Karimaldini et al. 2011). The estimation of ET_o through lysimeter is considered as the best choice but difficult mainly because of cost intensive, labor intensive and requires huge amount for properly setup. Consequently, indirect and data driven methods were developed to estimate ET_o by using meteorological data. The indirect methods developed for estimation of ET_o includes: (1) empirical and semi empirical Eq. (2) Pan-evaporation method (3) energy budget method (4) mass transfer Eq. (5) combination Eq. (6) radiation base methods. These methods are found in Xu and Singh (2002) and McMahon et al. (2013).

Food and Agriculture Organization (FAO) of United Nations has introduced one of the illustrious and well-known indirect standard method for ET_o estimation. This method involves incorporation of Penman–Monteith equation which was modified and reformed by Allen et al. (1998) as a reference equation (FAO-PM56). This equation (FAO-PM56) is influenced and controlled by various climatic, aerodynamic and surface resistance parameters. These comprise of maximum and minimum air temperature, maximum and minimum relative air humidity, wind speed and solar radiation, saturation vapor pressure deficit, slope vapor pressure curve and psychometric constant. There is an anomaly to air temperature available at the most weather stations, the remaining variables are often incomplete and not always found reliable for many locations (Rahimikhoob 2010). This point may be not valid for developed countries, but seems a big challenge in case of developing countries like Pakistan, where quantity and quality of data remain always questionable. Trajkovic and Kolakovic (2009) stated that reliable weather data sets of radiation, relative humidity, and wind speed are usually limited in developing countries.

Concerning and regarding the above contexts, there are myriad factors including physical, social and the environmental conditions in case of physical models influencing ET estimation which is the foundation of indirect methods. In addition, the geographic data (latitude, longitude, altitude) is also

required for the local adjustment of the different weather parameters e.g. atmospheric pressure, extraterrestrial radiations (R_a) and daylight hours (N) in FAO-PM56 equation. Additionally, the experimental approaches or field measurement efforts are usually time consuming, labor-intensive and post processing output process. Subsequently, it is quite difficult to formulate the ET_o equation to overcome these effects that can produce reliable and verified results (Ventura et al. 1999; A. R. Pereira and Pruitt 2004; Lopez-Urrea et al. 2006; Gavilan et al. 2007).

A few software solutions, namely, DailyET (Hess 1996), Cropwat 8.0 (Swennenhuis 2009) and ET_o calculator (Raes 2009) are efficaciously used to calculate ET_o . These software packages are based upon empirical and semi-empirical equations which require numerous climatic and geographic parameters. Moreover, these software packages will not yield results if even one of the parameter is not available. Thus, the data driven models based on different machine learning algorithms may be utilized as substitute models because of several advantages, for example, no required information of inner framework factors, easier answers for multi-variable issues and realistic calculation. These are innovative approaches which based on computational intelligent system to overcome imprecision and vulnerability in producing results (Chaturvedi 2008; Huang et al. 2010).

The main issue with modeling ET process is its non-linear dynamic and high complexity nature. In this regard, data driven models can be considered and contemplated as proper methods for ET_o estimation. Their ability to deal with complex problems by using only available set of data made them superior to conventional methods. Furthermore, there is also no need to give geographical information or local adjustment in these data driven models. The following literature provides a comprehensive comparison among data driven models and conventional methods in hydrological studies especially for ET_o estimation.

Among various data driven models, artificial neural networks (ANNs); adaptive neuro-fuzzy interference system (ANFIS); gene expression programming (GEP) have been widely applied for ET_o estimation and proved best choice over indirect methods. Various studies reported in literature which used ANN for ET_o estimation and found best

performance for ANN data driven model in comparison to indirect methods (Kumar et al. 2002; Sudheer et al. 2003; Trajkovic 2005; Kim and Kim 2008; Landeras et al. 2008; Traore et al. 2010). The potential for the use of ANNs was examined by Rahimikhoob (2010) to estimate ET_o based on air temperature data under humid subtropical conditions on southern coast of Caspian Sea. It was claimed that ANN model found superior in comparison to indirect empirical equations. In addition to ET_o estimation, ANFIS model was also investigated in parallel to other methods (Kişi and Öztürk 2007; Doğan 2009; Cobaner 2011; Aytek 2009; Kisi and Zounemat-Kermani 2014). Kişi and Öztürk 2007 have applied ANFIS model for ET_o estimation in Los Angeles, city of southern California. They concluded that ANFIS can be considered as a good approach in estimating ET_o . Aytek 2009 has investigated the potential of coactive adaptive neuro fuzzy interference system (CANFIS) in three stations for estimating daily ET_o . The results obtained with CANFIS were further compared with some empirical formulas. The study finally affirmed the CANFIS performance is superior as compared to the conventional methods. In similar manner, Karimaldini et al. (2011) have also examined ANFIS approach for estimating ET_o in arid condition by using limited climatic data. The findings confirmed the utmost performance of ANFIS over conventional empirical equations (calibrated FAO-PM56, Hargreaves, Priestley-Taylor, Makkink, and Blaney-Criddle). Cobaner (2011) examined two types of ANFIS based on grid partition (G-ANFIS) and subtractive clustering (S-ANFIS) methods in modeling ET_o by using daily climatic data. The results were compared with different types of ANNs and three empirical models. He claimed that S-ANFIS was performed better than ANNs and empirical formulas in estimation of ET_o . In addition, the study Kisi and Zounemat-Kermani (2014) have also applied G-ANFIS and S-ANFIS models in estimating ET_o with daily climatic data. They found ANFIS models superior to the corresponding conventional methods. Due to good generalization ability and attractive interference, gene expression programming (GEP) attained the higher significance in many hydrological studies e.g. stage–discharge curve, prediction in monthly precipitation (Azamathulla et al. 2011; Kisi

and Sanikhani 2015). In addition to literature, the performance of GEP was also investigated in modeling ET_o by using climatic variables (Shiri et al. 2014b, 2014a, 2014c; Guven et al. 2008; Traore and Guven 2012). The studies claimed that the GEP model could be considered as an alternative tool to the selected conventional methods for ET_o estimation.

Nowadays, support vector machine (SVM) and self-organizing networks (SONs) are also considered as the emerging tools in solving complex problems among the data driven models with an efficient outcomes (Asefa et al. 2006; Carpenter and Grossberg 1991, 1992; Lu and Wang 2005; Jorguseski et al. 2014; Østerbø and Grøndalen 2012; Peng et al. 2013). Group method of data handling (GMDH) and cascade correlation neural networks (CCNN) are found as the robust data driven models because of their tendency to develop direct relationship between input and output layers without the hidden layer as required in case of ANNs (Sherrod 2008, Sherrod 2010). Application of SVM, GMDH and CCNN models found in literature is presented thereafter.

Collobert and Bengio (2000) investigated the SVM model for large scale regression problems. It is claimed that time consumed in performing the analysis (training and testing phases) using SVM model was less than the other selected algorithms. The study Huang et al. (2002) also investigated the performance of SVM model. They found that the classifiers used in SVM model have more tendencies to solve complex non-linear problems as compared to the other classifiers. Kisi (2013) investigated the capability of SVM model for estimating the ET_o by using daily climatic data in two stations of California. The comparison results showed that SVM model performed better than the selected empirical equations. Kisi and Cimen (2009) investigated the potential of SVM model in ET_o estimation using daily meteorological data and compared the results with empirical models (CIMIS Penmenne, Hargreaves, Ritchie and Turc methods). They claimed that SVM performed better than empirical models and could be employed as an alternative tool in estimating ET_o . Shiri Nazemi et al. (2014a, b, c) applied heuristic approaches (including SVM) in comparison to the empirical and semi-empirical equations for estimation of ET_o . The

comparison results confirmed the superiority of heuristic approaches to the existing conventional methods. Tabari et al. (2013) investigated the capability of ANFIS and SVM in potato ET_c estimation. The results of ANFIS and SVM were also compared with the empirical equations. They found that SVM model performed better than ANFIS and corresponding applied empirical equations. Diamantopoulou et al. (2011) examined the potential of back propagation algorithm (BPANN) and CCNN in estimating the daily ET_o using weather data obtained from automatic weather stations (AWS) of Northern Greece. The results indicated that CCNN performed better than BPANN and selected empirical equations. Therefore, it could be considered as an appropriate tool in modeling ET_o process. In another study, Awari (2018) also investigated the capability of CCNN in the estimation of ET_o . The comparison results indicated that CCNN performed best than the other selected methods. Najafzadeh et al. (2014) investigated the potential of GMDH and SVM in estimating the scour depth below the pipelines and compared the results with empirical equations. The comparison results showed that the GMDH performed better than SVM and could be successfully employed as an alternative tool to the existing conventional methods. To the best knowledge of the authors, there is not any published article in the literature which investigate the performance of GMDH model in ET_o estimation.

Thus, it can be perceived from the above citations that data driven models are often found best for ET_o estimation. But most of the studies do not apply developed data driven models on other locations, which arise a question regarding models's performance. Nowadays, this practice is not justified for data driven models in determining their performance by incorporating similar data in training and testing phases. Therefore, a new practice must be followed to investigate and authenticate model's performance. According to the knowledge of authors, there is no study conducted in Pakistan that applies data driven models for ET_o estimation. The main aims of this study were: (1) to determine the ET_o using globally accepted method (FAO-PM56) across a range of climatic regions (2) to investigate the potential of different data driven models (i.e. SVM, GMDH,

CCNN, MLP and GRNN) for ET_o estimation (3) to compare ET_o estimation of PM 56 method with that of data driven models.

2. Materials and Methods

2.1. Location and Data Collection of Studied Regions

Pakistan expands north-east to south-west from 23.5° to 37° N and 60° to 75° E in south-west Asia over world map. It is bordered with Afghanistan in north-west, Iran in west, China in far north-east and India in east. The total area of Pakistan is nearly about 80 mha out of which only 37 mha is available for raising crops. A variety of climatic conditions exists in Pakistan. The weather of northern and north-western parts is quite cool because of the presence of snow-packed mountains. Plains of Punjab province are located in the center and are quite hot and semi-arid in nature. Highlands of Balochistan province have cold climate but arid in nature and the coastal areas located in south are warm and humid. Temperature of colder regions located in north falls below freezing point during winter months while a high temperature of about 40 °C prevails during summer months in the semi-arid and arid desert regions of Pakistan with a peak approaching 50 °C frequently. The northern and western highlands of Pakistan receive an annual rainfall of 1500 mm and 500 mm respectively while desert area receives very less rainfall. Thus, climatic distribution of Pakistan is arid to semi-arid, except a short stretch of temperate in north with wide variations.

The climatic data of six selected cities located in various climatic regions of Pakistan, was used in current study for a duration of 30 years (1987–2016). The data includes long term monthly average values of minimum temperature (T_{min}), maximum temperature (T_{max}), average relative humidity ($R.H_{mean}$), sunlight hours (N), and average 24 h. wind speed (U). This dataset was divided into sections. First section contains 70% of the data (1987–2007) and used to train each selected model while remaining 30% of the data (2008–2016) employed to test the developed models. The performance of one machine learning

Table 1

The geographic description of selected climatic regions

Sr. no	Study Site		Coordinates		Elevation m	Duration Years
	Country	City	Lat (DD)	Long (DD)		
1	Pakistan	Multan	30.2	71.45	122	1987–2016
		Faisalabad	31.41	73.11	184	
		Islamabad	33.72	73.06	540	
		Jacobabad	28.28	68.43	61	
		Peshawar	34.01	71.54	331	
		Skardu	35.3	75.68	2228	
2	New Zealand	Zealand				Auckland 2015–2018
		36.84	174.76	196		
3	China	Harbin	45.74	126.63	140	2011–2014
4	USA	Belleville	89.98	38.52	133	2015–2018
		Big Bend	87.87	42.05	182.4	
		Dixon Springs	89.39	40.63	165	

based model could be hardly justified if applied to only one location/country. To further test the validity of the developed models, the climatic data of five more cities located in New Zealand, China and USA was also used. It is pertinent to mention here that the data of six cities located in Pakistan was used to train and test the developed models while the data of five cities located in three countries of New Zealand, China and USA was only used to test the developed models. The details of the data of different cities located in Pakistan, New Zealand, China and USA is given in Table 1.

The location of six selected climatic stations of Pakistan is presented in Fig. 1a while Fig. 1b shows the climatic stations of other selected countries. All selected climatic stations of Fig. 1a and b are classified into six climatic zones which can be seen in Table 2. In addition, classifications of these selected climatic stations are based on global aridity index (Todorovic et al. 2013). This index can be calculated by dividing the mean yearly precipitation (P) to the mean yearly ET_o . Additionally, the humid region is further categorized into warm, mild and high humid climatic zones. The division of these zones is based on average annual rainfall and mean temperature data. This is due to the fact that climatic conditions that lie in humid regions can easily be fathomable.

2.2. Adopted Methodology

2.2.1 FAO-PM56 Method

This method is universally accepted as a standard method for the estimation of ET_o . It is a physical method which requires numerous climatic parameters. The method is based on combine equations which estimate ET_o values. The mathematical relation to estimate ET_o by using climatic parameters is given in Eq. 1.

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma \times \frac{900}{T_{\text{mean}} + 273} \times U_2 \times (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

wherein R_n , and G are net radiation ($MJ/m^2/day$) on crop surface and soil warmness flux density ($MJ/m^2/day$), respectively. Additionally, T_{mean} , U_2 , e_s , e_a , e_{min} , e_{max} , Δ and γ are means of maximum and minimum air temperatures ($^{\circ}C$), air speed (m/s), saturation vapor pressure (kPa), actual vapor pressure (kPa), saturation vapor pressure at minimum air temperature, saturation vapor pressure at maximum air temperature, slope vapor pressure curve ($kPa/^{\circ}C$) and psychrometric constant ($kPa/^{\circ}C$), respectively. As the value of G is smaller than R_n , it may be taken as zero. The value of Δ can be determined by utilizing the procedure which is enlightened by Gocić et al. (2015). The complete derivation of PM equation either in case of complete or missing data can be found in Allen et al. (1998).

2.2.2 Multilayer perceptron (MLP)

MLP is frequently used neural network in solving complex classification problems. In this network, input is given to network along with target output and weights are adjusted in such a way that network can generate target output with several attempts. The MLP structure consists of an input, a hidden and an output layers. Each layer contains a number of neurons that are linked in same pattern with neurons inside the subsequent layer. Each neuron carries a number of predictor variables from previous layer and outputs to the following layer. Each layer plays a significant role in the overall performance of structure. The neurons of one layer are linked to the

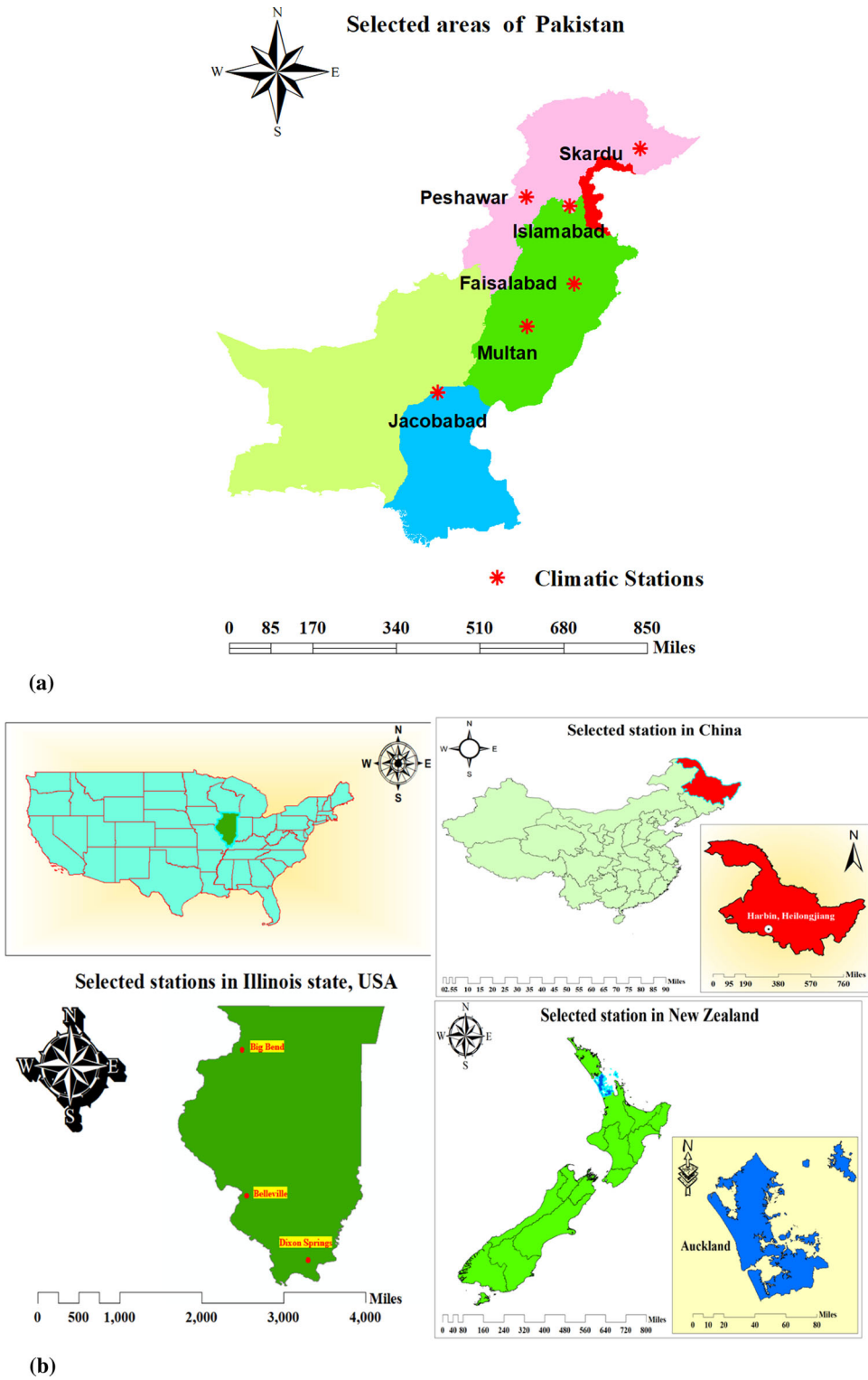


Figure 1 Location of the six selected climatic stations in Pakistan. **a** Location of the selected climatic stations in other countries

Table 2

Division of climate zones, respective precipitation to ET_o ratios

Climate zone	Ratio P/ ET_o	Climate zone	Ratio P/ ET_o
Hyper-arid (H.A)	< 0.05	Dry sub humid (D.H)	0.51–0.65
Arid (A)	0.05–0.20	Moist sub humid (M.H)	0.66–1.0
Semi-arid (S.A)	0.21–0.50	Humid (H)	> 1
Stations	Climatic zone	Stations	Climatic zone
Multan	H.A	Harbin	High Humid
Jacobabad	H.A	Belleville	H
Faisalabad	S.A	Big Bend	Warm humid
Islamabad	D.H	Dixon Springs	Mild humid
Peshawar	S.A	Auckland	H
Skardu	H		

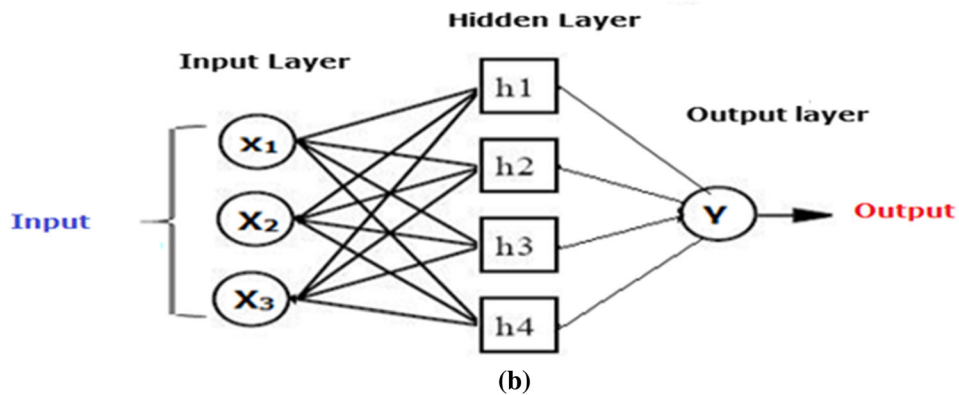
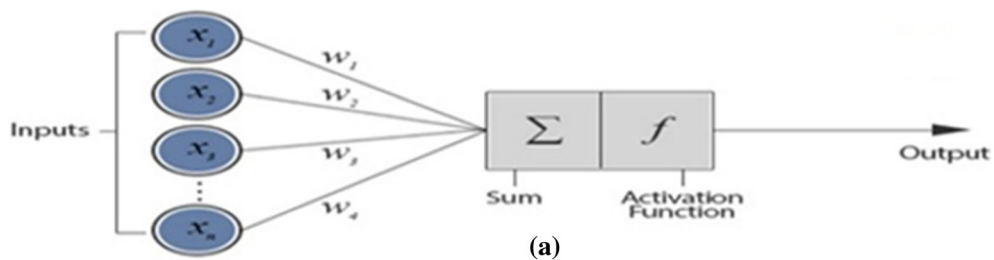


Figure 2
Development of MLP model for ET_o estimation

neurons existed in other layer through a specific connection called weights. These weights are made accountable to carry results from one to another layer. Output of each neuron within input layer will become an input for the hidden layer. Furthermore, output of each neuron from hidden layer converts into an input for every neuron inside output layer. A structure of MLP neural network with one hidden layer can be seen in Fig. 2a. A function is utilized to process input

data among the neurons of each layer known as activation function. Moreover, the sigmoid and linear functions are utilized to process input data from input to hidden layer and hidden to output layer, respectively. The network can be trained by adjusting weights through an error back propagation process explained by Hassanpour Kashani et al. (2008). The present study utilizes scaled conjugate gradient (SCG) method which can be swiftly performed twice

Table 3
Parametric values of MLP model after tuning

Parameters	Values	Parameters	Values
Number of layers	3	Number of iterations	5,000
Min to max neurons	2–20	Convergence tolerance	(1.000e ⁻⁰⁰⁵)
Neurons in hidden layer	6	Minimum improvement delta	(1.000e ⁻⁰⁰⁶)
Hidden layer function	Sigmoid	Minimum gradient	(1.000e ⁻⁰⁰⁷)
Output layer function	Linear	Maximum execution time	0

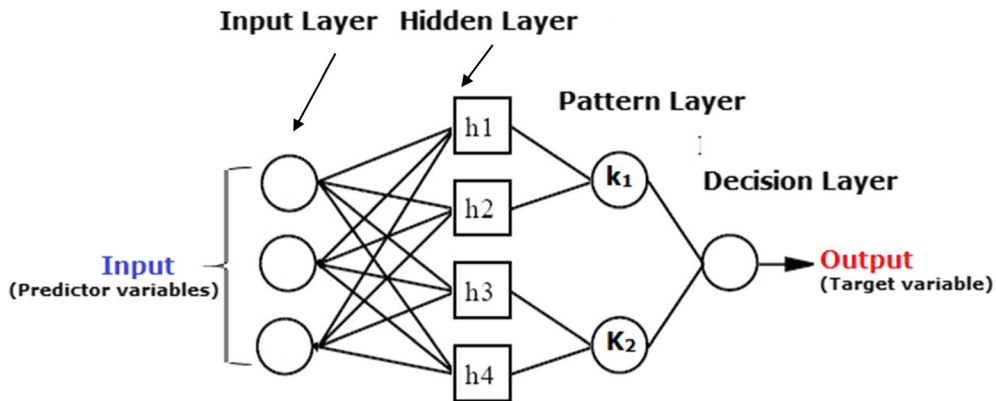


Figure 3
Configuration of GRNN model

as compared to Conjugate gradient and converging up to 20 times than scaled conjugate gradient (Sherrod 2008). Møller (1993) has explained SCG algorithm in detail. development of MLP model is described in the next lines. five meteorological variables (T_{\max} , T_{\min} , $R.H_{\text{mean}}$, N , U) are fed in an input layer as predictor variables. calculated ET_o is used as a target variable in an output layer. The mapping of input–output relationship with MLP model is presented in Fig. 2b. Six neurons are selected in hidden layer after tuning the neurons. The V-fold cross validation is presently applied to check and validate input data. essential parameters for the development of MLP model are given in Table 3.

2.2.3 General Regression Neural Networks (GRNN)

Specht (1990) has proposed a method that allows to contrive and formulate the weighted-neighbor method. A general regression neural network

(GRNN) can be utilized for regression problems while probabilistic neural network (PNN) is implemented to solve classification problems. This study employs GRNN network in order to estimate ET_o by using meteorological variables. Architecture of GRNN model is shown in Fig. 3.

Development of GRNN model to predict ET_o is as followed: Five meteorological variables, (T_{\max} , T_{\min} , $R.H_{\text{mean}}$, N , U) are used as predictor variable while ET_o is obtained as a target variable. Gaussian function is chosen as a kernel function which is considered one of the best yet among invented (Sherrod 2008). Eliminating neurons frequently minimize the error of developed model. The procedure to eliminate redundant neurons is considered slow. This is an iterative process due to the fact that model needs to be assessed with each persisting neuron to determine the best one as a way to eliminate. Thus, leave on out validation (LOO) with minimize error method is employed to find optimal number of

Table 4
Tuning parameters of GRNN model

1. Network parameters			
Minimum sigma value	0.001	Search step	20
Maximum sigma value	10	Kernel function	Gaussian
Unnecessary removing Neurons method	Minimize error	Cross validation	LOO method
2. Additional parameters			
Number of Iterations	10,000 Convergence tolerance 0.006737		
Minimum improvement delta	0.002478 Minimum gradient 0.00091		
Maximum execution time	0		

neurons. The essential parametric values for GRNN model are given in Table 4.

2.2.4 Support Vector Machine (SVM)

Eventually, the concept and idea of learning machine is suggested by Turing (1950). Support vector machine is scrutinized and regarded as a supervised learning model which is built on statistical learning theory. The statistical learning theory is familiarized and made conversant by Vapnik (1998). Support vector regression (SVR) is frequently used to narrate SVM approach. SVR has capability to apply structural risk minimization principle in a reliable way. It is an inductive principle for model selection which is used for finite training data sets. It describes a general model of capacity control and provides a relationship between space complexity and quality of fitting the training data. Moreover, it emphasizes on an upper bound instead of training error in the data set. Minimum value of upper bound produces a best model. Based on this principle, SVR achieves an optimum network structure. More details regarding this principle can be found in Wu et al. (2008). The main objective of SVM modeling is to find optimal hyper plane which isolates the clusters of vector. This can be accomplished by separating variables using linear and non-linear dividing. The division of attributes in non-linear case is more complex than linear case. Furthermore, the separation in high dimensional case is easier than low dimensional. Figure a–d illustrate the pictorial view of linear, complex, low and high dimensional cases, respectively (Han and Kamber 2001; Markowetz 2003). The vectors adjoining to hyper plane are termed as support vectors.

This study employs two kinds of SVR, namely, epsilon (ϵ) and neuron (Nu). There is no general method to determine model of the largest capacity in order to estimate ET_o , except iterative method. After a large number of iterations, it was found that ϵ -SVR performed good than Nu-SVR. Therefore, ϵ -SVR is used in SVM model for present study. The accuracy of SVR model depends upon following parameters: penalty factor(C), gamma (γ), P, Nu etc. In addition, there is a dire need to determine proper optimal values for these parameters to achieve satisfactory, better and acceptable results. For example, large value of C creates over fitting problem while a small value of C causes under-fitting the model. Therefore, it is mandatory to select optimal values for all parameters to obtain reliable and better results. Grid and pattern (GAP) search are concurrently applied in this study in order to find optimal parametric values. A grid search tries to attempt values of each parameter across the specified search range. This can be done by using geometric steps. A pattern search starts at the center of search range and makes trial steps in each direction for each parameter. More comprehensive details about these searching methods can be found in Sherrod (2008). The attributes (predictor variable) in SVM model are meteorological variables while ET_o is selected as predicands (target variable). Four kinds of kernel functions (linear, RBF, polynomial and sigmoid) are concerned and applied with V-fold cross validation method. It is used to construct an efficient model. In addition, number of parameters in SVM model largely depends upon the applied kernel function. Thus, Table 5 provides epigrammatic details about the development of SVM model. It also shows optimal values for all required parameters which are needed corresponding

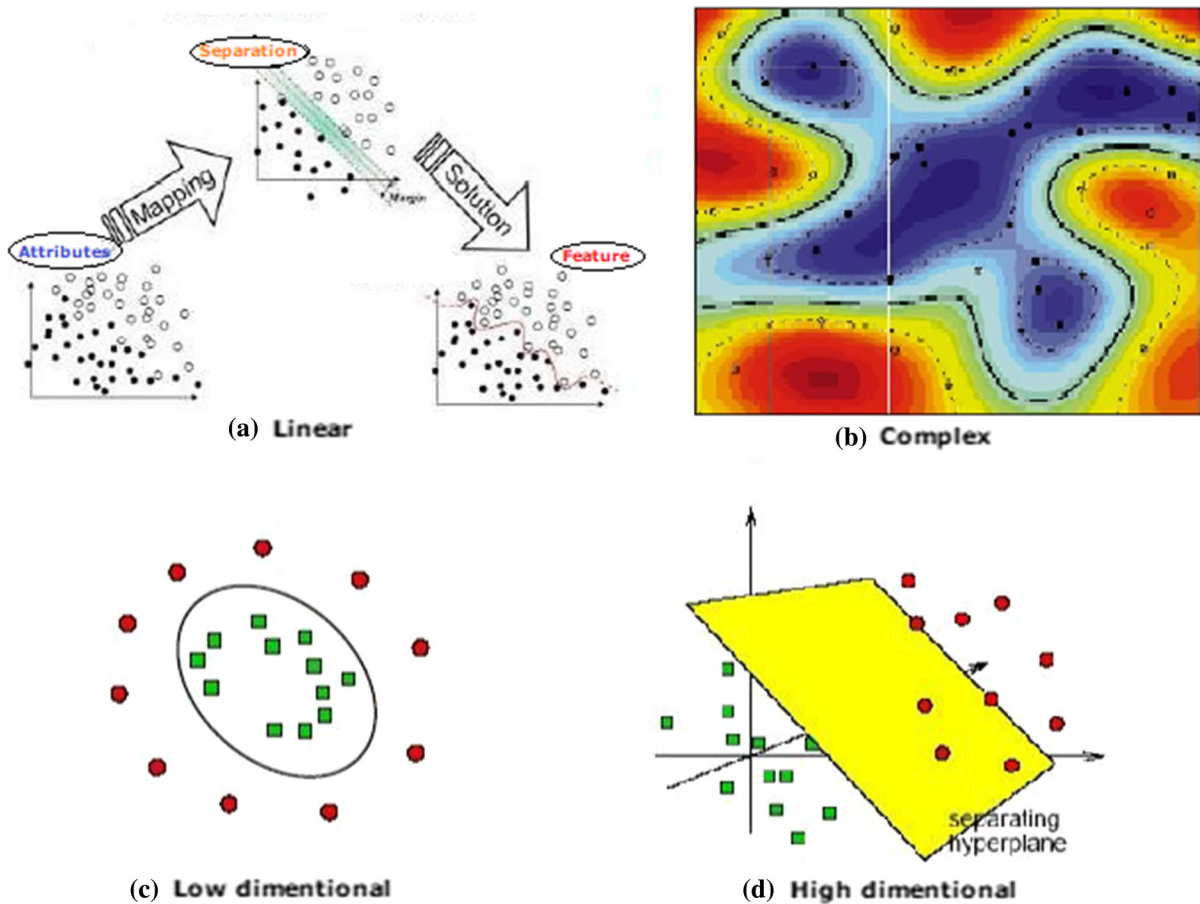


Figure 4
Schematic diagram of SVR model

to applied function. Moreover, number of parameters which do not show any contribution in the development of SVM model regarding kernel function is represented by symbol “X” and can be seen in Table 5.

2.2.5 Group Method of Data Handling (GMDH)

GMDH model was developed by Ivakhnenko (1968). This algorithm presents a unique approach to the problem of modeling and even a new philosophy to scientific research. It is used in various fields such as data mining, prediction and complex system modeling etc. Li et al. (2017) found GMDH better than other classical forecasting algorithms. A schematic diagram of GMDH is shown in Fig. 5. The network starts with just input neurons. The main layer (at the

top) presents one contribution for every indicator variable. Each neuron within second layer draws its inputs from two of the input variables. In similar way, the neurons within third layer draw their inputs from two of the neurons inside preceding layer. This procedure proceeds until second final layers fetch their inputs to the ultimate layer. The final layer (at the lowest) draws its two inputs from past layer and produces a single value that’s the output of network. Additionally, the links between neurons inside this network are not fixed. It can be observed in Fig. 5 that neuron at the left end of third layer is linked with an input variable rather than output of preceding layer. Thus, current study applies GMDH model in order to draw a relationship between climatic variables and ET_0 . Five meteorological variables, symbolically, T_{max} , T_{min} , RH_{avg} , U , N are used as

Table 5
Optimal ranges of parameters for SVR model

Models	Kernel Function	Optimal ranges of model parameters					
		C	P	γ	Nu	Coef0	Degree
ϵ -SVR	Linear	0.001–78	0.001–58	X	X	X	X
	RBF	0.001–78	0.001–58	0.001–58	X	X	X
	Polynomial	0.001–78	0.001–58	0.001–58	X	0–100	1
	Sigmoid	0.001–78	0.001–58	0.001–58	X	0–100	
Nu-SVR	Linear	0.001–78	X	X	0.001–0.6	X	
	RBF	0.001–78	X	0.001–58	0.001–0.6	X	
	Polynomial	0.001–78	X	0.001–58	0.001–0.6	0–100	1
	Sigmoid	0.001–78	X	0.001–58	0.001–0.6	0–100	

predictor variables. The output layer generates ET_o as a target value. Various types of functions such as linear, quadratic, cubic, asymptotic etc. (Table 6) are used in GMDH model to improve model's accuracy. combination of two quadratic functions (1 variable and 2 variable) is applied in the training and testing process for estimation of ET_o . results generated with this combination of functions were found the best among other functions. Table 6 indicates essential parametric values for development of GMDH model.

2.2.6 Cascade Correlation Neural Network (CCNN)

Primarily, a CCNN model includes the most effective input and output layers without a hidden layer. Every

input is attached to each output neuron by way of a connection with an adjustable weight. Each output neuron receives its input as weighted sum from all of input neurons alongside the bias value. At that point, yield neuron sends this weighted input sum through its transfer function to withdraw final output as shown in Fig. 6a. A simple CCNN also shows good performance in case of small data. A hidden layer is introduced into CCNN model for solving large data set and complex problems. Neurons are delivered to hidden layer one after another. Every new hidden neuron gets a connection from each of the inputs and every pre-present hidden neuron. The hidden neuron's input weights train and then fixed. Simplest output connection weights train over and over. Each new neuron, consequently, provides a new one hidden unit to the network. This prompts the formation of an effective powerful high-order function which includes identifier and deep network with a big variety of inputs to output neurons. Now, the neurons inside hidden layer do not change, once they have been brought. It is known as a cascade because output from all neurons already inside network feeds into the new neurons. A complete structure of CCNN model which is used in the current study has illustrated in the Fig. 6(b). It can be observed that the ET_o values are obtained as target outputs in CCNN model after giving meteorological parameters as predictor variables. Mixture of Sigmoid and Gaussian functions is employed as a kernel function in the hidden layer. In the CCNN model, advance parameters are also used with training parameters which enhance the predictive accuracy

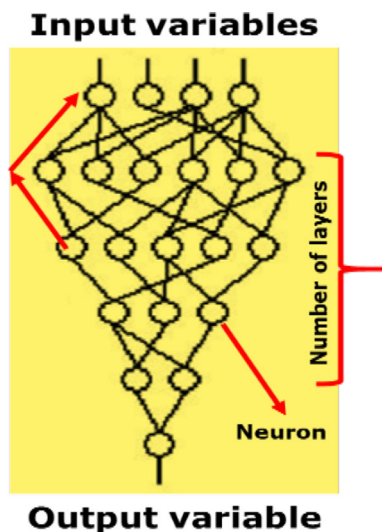


Figure 5
Structure of GMDH model

Table 6
Essential parameters and list of functions for GMDH model

Parameters		Values	
Minimum number of network layer		04	
Maximum polynomial order		16	
Convergence tolerance		1.000e ⁻⁰⁰⁴	
Functions to use GMDH model			
Functions	Equations	Functions	Equations
Linear 1 variable	$y = p_1 + p_2x_1$	Ratio 2 variables	$y = p_1 + p_2 \frac{x_1}{x_2}$
Linear 2 variables	$y = p_1 + p_2x_1 + p_3x_2$	Asymptotic 1 variable	$y = \frac{p_1+p_2}{p_1+p_2}$
Linear 3 variables	$y = p_1 + p_2x_1 + p_3x_2 + p_4x_3$	Gaussian 1 variable	$y = \frac{x_1^2+p_3}{p_1+p_2 \exp(-(x_1-p_3)^2)}$
Quadratic 1 variable	$y = p_1 + p_2x_1^2$	Logistic 1 variable	$y = \frac{p_4}{p_1+p_2}$
Quadratic 2 variables	$y = p_1 + p_2x_1 + p_3x_1^2 + p_4x_2 + p_5x_2^2 + p_6x_1x_2$	Exponential 1 variable	$y = p_1 + p_2 \exp(p_3(x_1 + p_4))$
Cubic	$y = p_1 + p_2x_1 + p_3x_1^2 + p_4x_1^3$	Product 2 variables	$y = p_1 + p_2x_1x_2$
1 variable			

y = output variable; x_1, x_2, x_3, \dots = input variables; p_1, p_2, p_3, \dots = constant factors; maximum iterations = 10,000; execution time = zero

and minimize the error. Table 7 provides essential parametric values for CCNN model to generate reliable results.

2.2.7 Weaknesses of FAO-PM56 Method vs Strengths of Numerical Methods

The concept of ET_o has been widely discussed and used for decades (Doorenbos and Pruitt 1977; Pereira et al. 1999). Allen et al. (1998) introduced a clear definition of FAO-PM56 method for estimation of ET_o. One of the most important reasons for advocating a simpler method than FAO-PM56 is the substantial likelihood for inaccuracy in weather data measurement and collection, especially for the developing countries and meteorological stations managed by non-experts. In these situations, the accuracy of data and especially for more advanced parameters such as radiation and humidity, can be very low (Droogers and Allen 2002). Table 8 shows data requirements for FAO-PM56 and data driven models. It can be seen in Table 8 that FAO-PM56

depends upon several parameters which is not easily accessible especially in the developing countries. Alternatively, data driven models only depend upon climatic parameters and require less parameter in comparison to FAO-PM56 method.

3. Model Performance Evaluation

The goodness of fit is assessed through a set of indicators that are used to compare all the pairs of observed and model-predicted values of selected variables. Nine performing indicators are selected in this study to evaluate the performance of MLP, GRNN, GMDH and CCNN models. These include square of Pearson correlation coefficient (r²), linear regression coefficient (R²) model efficiency (M.E), mean absolute error (MAE) and root mean square error (RMSE), scatter index (SI) which are defined as follows (Legates and McCabe 1999; Alazba et al. 2011; Mattar and Alamoud 2015; Mattar et al. 2015)

$$r^2 = \frac{n \times [\{ \sum_{i=1}^n (ET_{obs} \times \overline{ET}_{est}) \}] - [\{ (\sum_{i=1}^n ET_{obs}) \times (\sum_{i=1}^n ET_{est}) \}]}{\sqrt{[n \times \{ \sum_{i=1}^n (ET_{obs})^2 \} - \{ (\sum_{i=1}^n ET_{obs}) \}^2] \times [n \times \{ \sum_{i=1}^n (ET_{est})^2 \} - \{ (\sum_{i=1}^n ET_{est}) \}^2]}} \quad (2)$$

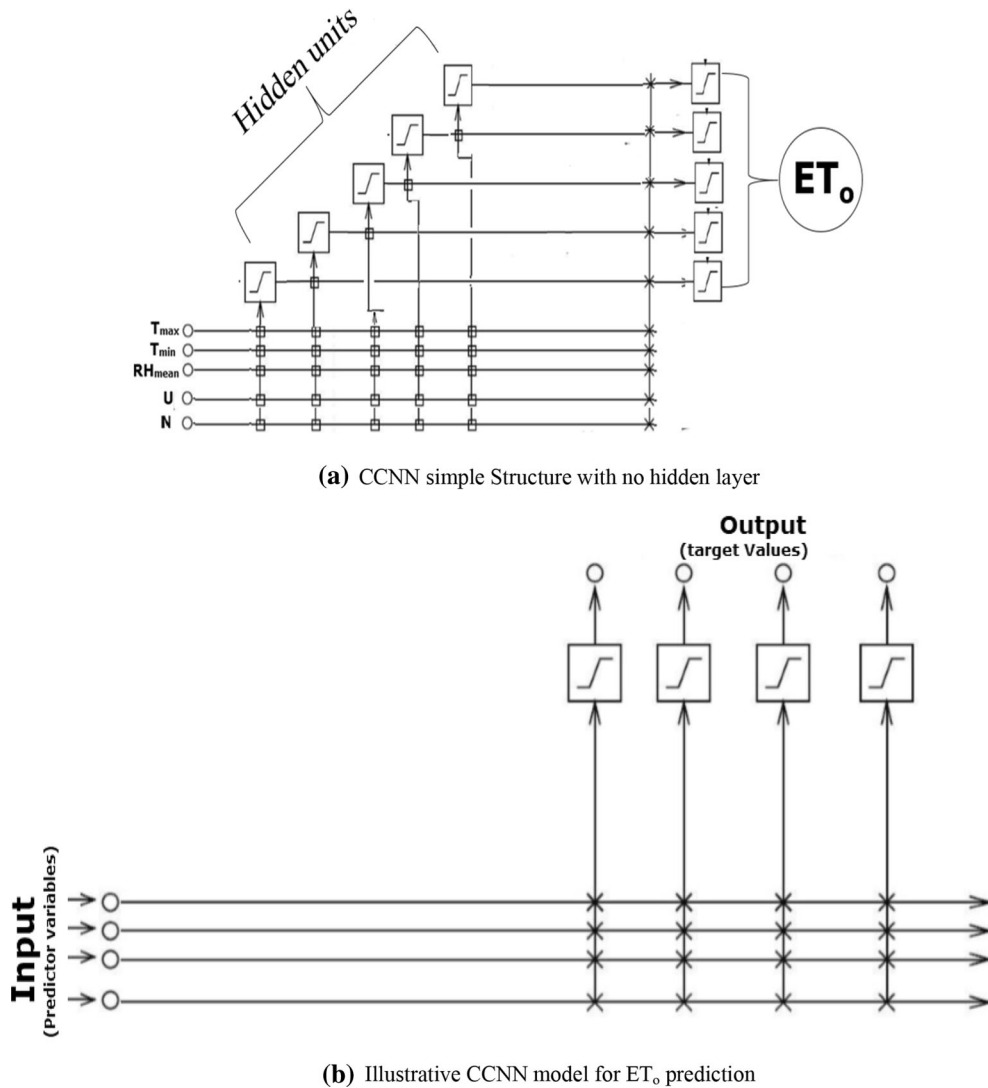


Figure 6
a CCNN simple Structure with no hidden layer. **b** Illustrative CCNN model for ET₀ prediction

$$R^2 = \frac{[(\sum_{i=1}^n (ET_{obs} - \overline{ET_{obs}}) \times (ET_{est} - \overline{ET_{est}})]^2}{[(\sum_{i=1}^n (ET_{obs} - \overline{ET_{obs}}) \times (ET_{est} - \overline{ET_{est}})]} \quad (3)$$

$$M.E = 1 - \frac{\sum_{i=1}^N (ET_{obs} - ET_{sim})}{\sum_{i=1}^N (ET_{obs} - \overline{ET_{obs}})} \quad (4)$$

$$MAE = \sum_{i=1}^N \frac{|ET_{obs} - ET_{sim}|}{N} \quad (5)$$

$$MAPE = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{ET_{obs} - ET_{est}}{ET_{obs}} \right| \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (ET_{obs} - ET_{est})^2}{N}} \quad (7)$$

$$SI = \frac{RMSE}{ET_{mean}} \quad (8)$$

wherein, ET_{obs}, ET_{est}, $\overline{ET_{obs}}$, $\overline{ET_{est}}$, ET_{mean} are observed value, estimated value, average of observed value, average of estimated value, mean of observed

Table 7

Training and advance parameters of CCNN model

1. Training parameters				
Minimum neurons	0	Candidate neurons	21	
Maximum neurons	26	Candidate epochs	200	
Cross fold	04	Cross validation	V fold method	
2. Advance parameters				
Regularization parameters	Weight decay	Epsilon	Mu	
Candidate	0–0.001	1000–0.01	1–2	
Output	0–0.001	1000–0.01	1–2	

Table 8

Data requirements for the applied ET_o methods (Standard (FAO-PM56) method vs. Data driven models)

Required parameters	T _{min}	T _{max}	R.H	U	N	Rn	es	ea	e _{min}	e _{max}	Other factors (Δ, Z and γ)
FAO-PM56 Method	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Data driven models	✓	✓	✓	✓	✓						

value of ET_o, respectively and N represents total number of observations. Further two coefficients, namely, skewness (C_s) and variation (CV) are also used in present study to distinguish the data.

$$C_s = \frac{\sum_{i=1}^n (O_i - \mu)^3}{(n-1) \times \sigma^3} \quad (9)$$

$$C.V = \frac{\sigma}{\mu} \quad (10)$$

Here, O_i, n, σ, and μ represent observed data, number of data points, standard deviation and mean value of data set, respectively. C_s determine the symmetry of distribution and σ measures amount of variation in the given data set (Bland and Altman 1996).

4. Results and Discussions

4.1. Experimental Data

The values of meteorological data on monthly basis are used to develop data driven models. Data sets used for data driven models based on climatic data are divided into two subsets. These include training and testing data sets. Training data set of each climatic station includes 70% of observed data while testing data set contains 30% of observed data collected in the period of years. The climatic data

duration corresponding to each climatic station is given in Table 1. Moreover, Table 9 shows the value of statistical parameters of climatic data such as minimum, maximum, mean, standard deviation and skewness coefficient values of the selected countries. Additionally, averages of aforementioned parameters for six selected cities of Pakistan are estimated and presented in Table 9. Likewise, averages of statistical parameters of meteorological parameters for other locations used in the current study are also given in Table 9.

4.2. Performance Analysis

At the beginning, data driven models are trained with extracted data from the FAO-PM56 equation. As a baseline for comparison of the performance of SVM with other data driven models, the same training and testing data set was considered for stations of Pakistan. To estimate FAO-PM56 ET_o values, r², RMSE, MAE and ME are used in comparison between the FAO-PM56 ET_o and predicted values of SVM, GMDH, GRNN, CCNN and MLP models. The performing indices, R² and RMSE of testing data set for each selected city of Pakistan are presented in Fig. 7a and b, respectively. To present results in elegant way, abbreviations of the selected cities of Pakistan are used for graphical representations. These are followed as: Faisalabad (Fsd), Multan (Mul),

Table 9
 Statistical parameters of selected countries used in current study

Countries	Data Source	Variables	Mean	S.D	Min	Max	Cs
Pakistan	https://www.pmd.gov.pk/	T_{max}	26.81	1.1	19.4	34.22	- 0.42
		T_{min}	12.48	1.87	4.65	20.32	- 0.07
		$R.H_{avg}$	45.17	1.98	38	56	0.775
		U	139.64	2.6	78.37	250.9	- 0.31
		N	7.33	2.88	6	8	- 0.12
New Zealand	https://www.niwa.co.nz	T_{max}	17.39	4.82	8.10	27.00	0.02
		T_{min}	6.82	3.84	- 0.80	15.40	- 0.06
		$R.H_{avg}$	89.37	6.00	73.90	99.70	- 0.42
		U	125.40	2.49	69.12	172.80	- 0.12
		N	4.85	1.77	1.54	9.39	0.25
China	https://www.cma.gov.cn/en2014/	T_{max}	10.57	3.63	- 16.20	31.40	- 0.30
		T_{min}	0.33	4.49	- 26.10	20.70	- 0.22
		$R.H_{avg}$	53.11	5.92	25.90	73.00	- 0.36
		U	199.15	4.36	111.50	335.20	0.69
		N	5.94	1.72	0.90	9.90	- 0.37
USA	https://www.isws.illinois.edu	T_{max}	14.22	3.29	- 4.57	30.01	- 0.11
		T_{min}	10.40	4.08	- 11.63	25.68	- 0.15
		$R.H_{avg}$	71.57	4.80	59.87	89.33	0.59
		U	6.09	1.56	3.23	9.23	- 0.03
		N	4.98	2.1	1.25	9.70	0.12

T_{max} maximum air temperature ($^{\circ}C$), T_{min} minimum air temperature ($^{\circ}C$), N sunlight hours (h), U air speed at 2 m height (km/day), RH_{avg} average 24 h. relative humidity (%)

Islamabad (ISD), Jacobabad (Jac), Peshawar (Pes) and skardu (Skr). It can be observed in Fig. 7a that SVM model looks superior to other numerical computing models in all the selected cities of Pakistan. However, GMDH and GRNN models show almost equal outcomes in comparison to MLP model. In addition, CCNN model generates low r^2 values except in Faisalabad city. It is evident from Fig. 7a that all models have acceptable square of Pearson correlation coefficient (r^2), which is higher than 90%. It means that there is no over fitting of training and testing data of the numerical computing models. In other words, the data driven models capture all nonlinearity of ET_o , which is the main advantage of data driven models. Likewise, it can be seen in Fig. 7b that SVM model generates low RMSE value for the selected cities of Pakistan which varies from 0.13 to 0.50 mm/month. On contrary, GMDH, GRNN, MLP and CCNN models show high value of RMSE. Although GRNN and MLP models generate high value of RMSE as compared to SVM model but have performed well in comparison to GMDH and CCNN models for all the selected cities

of Pakistan. The RMSE values for all the chosen cities of Pakistan by employing GRNN and MLP models varies from 0.7 to 0.70 and 0.18–0.81 mm/month, respectively. Thus, it can be concluded that SVM model has more capability among all other numerical computing models to predict accurate ET_o in various climatic regions. However, GRNN and MLP models also show good performance in comparison to GMDH and CCNN models for all selected cities of Pakistan. The summary performance indices, M.E MAE, MAPE, SI, of various approaches towards estimating ET_o for various climatic zones of Pakistan can be seen in Table 10. It can be observed in Table 10 that model efficiency of all the data driven models in various climatic zones of Pakistan was found greater than 90%. As it can be seen model efficiency is higher for SVM model as compared to the use of other models in each climatic zone of Pakistan. However, MLP model looks superior in comparison to GRNN, CCNN and GMDH models in Multan, Peshawar and Skardu climatic stations. In contrast, GMDH model shows good performance as compared to MLP,

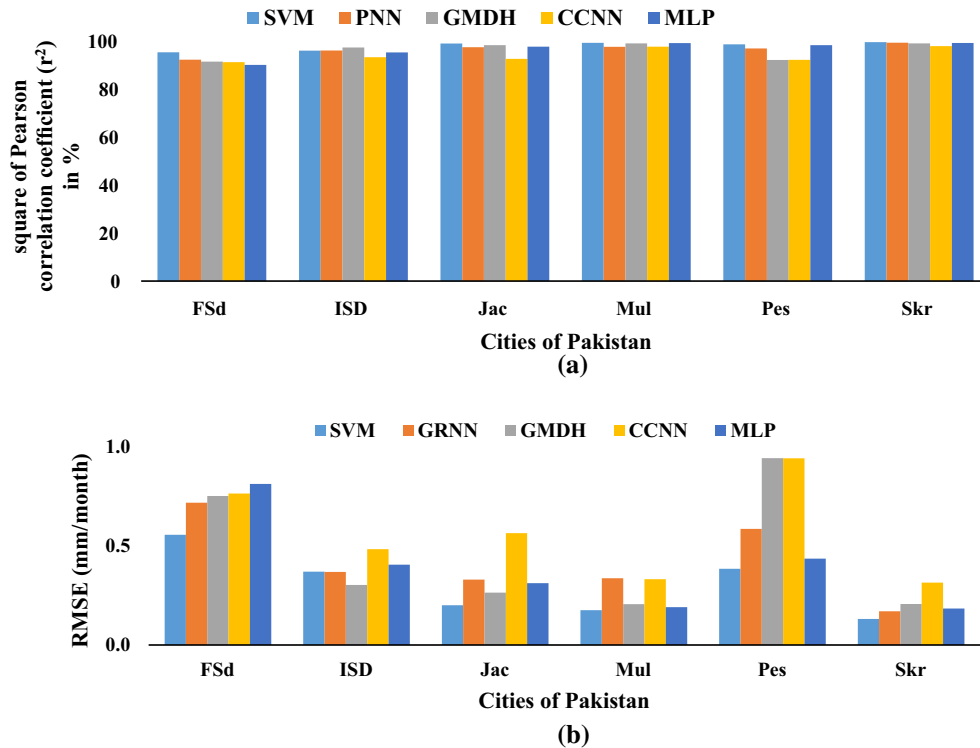


Figure 7
 r^2 and RMSE values for the selected cities of Pakistan

CCNN and GRNN models in Islamabad and Jacobabad cities. Furthermore, CCNN model has greater value for M.E in Multan city as compared to MLP and GMDH models while it shows slightly better performance than GRNN model. Table 10 further supports our above discussion and expresses the performing indices of all numerical computing methods for each climatic station of Pakistan. Additionally, SVM model has attained 99 percent value for M.E in order to estimate ET_o . Moreover, Table 10 also demonstrates M.E value for SVM approach in selected climatic stations of Pakistan which lies between 95 and 100 percent. In general, the results indicate that SVM approach has the best estimating capability for ET_o estimation. similar results were obtained by Shrestha and Shukla (2015) for tropical environment. Two indicators, namely, MAE and SI are further calculated to endorse the selected numerical computing methods for each climatic station. These two indicators have values range from 0 to 1. If

value approaches to 0, it shows the perfect fit of model. Alternatively, model shows less predictive accuracy when its outcome reaches to 1. It can also be observed in Table 10 that MAE and SI indices have least values for SVM model in each climatic station of Pakistan. In addition, SVM model has 0 values for SI while it shows minimum value of MAE as compared to the other models in each selected cities of Pakistan. According to the results depicted in Table 10 that MLP has smaller values for SI and MAE in comparison of GRNN, CCNN and GMDH models in Multan, Peshawar and Skardu climatic stations. On the other hand, MLP, CCNN and GRNN models generate higher values for MAE and SI as compared to GMDH model in Islamabad and Jacobabad cities. Additionally, CCNN has lower values of SI and MAE than MLP, GMDH and GRNN models for Multan city. above whole discussion could be summarized as all selected data driven models except SVM do not maintain consistency of their results in

Table 10
Performance parameters for ET_o assessment in Pakistan using selected approaches

Climatic stations	Numerical methods	Training Algorithms (for optimal model structure)	Performing indices			
			M.E	MAE	MAPE	SI
FSD	SVM	Grid and pattern search	95.4	0.42	8.57	0.11
	GRNN	Random/mixture of transfer function	93.92	0.50	10.18	0.14
	GMDH	Equation combination/function type	91.53	0.52	9.85	0.15
	CCNN	Conjugate gradient	94.54	0.49	12.77	0.15
	MLP	Nguyen widrow/scaled conjugate gradient	90.11	0.58	10.35	0.16
ISD	SVM	Grid and pattern search	95.1	0.21	7.63	0.07
	GRNN	Random/mixture of transfer function	96.14	0.21	7.81	0.07
	GMDH	Equation combination/function type	97.39	0.21	8.46	0.06
	CCNN	Conjugate gradient	97.11	0.23	8.99	0.10
	MLP	Nguyen widrow/scaled conjugate gradient	95.33	0.28	10.56	0.08
JAC	SVM	Grid and Pattern Search	99.06	0.14	4.04	0.04
	GRNN	Random/mixture of transfer function	97.48	0.24	6.48	0.07
	GMDH	Equation combination/function type	98.38	0.19	6.14	0.05
	CCNN	Conjugate gradient	98.2	0.21	6.56	0.11
	MLP	Nguyen widrow/scaled conjugate gradient	97.74	0.24	6.26	0.06
MUL	SVM	Grid and pattern search	99.36	0.13	4.33	0.03
	GRNN	Random/mixture of transfer function	97.66	0.24	6.58	0.07
	GMDH	Equation combination/function type	98.15	0.24	7.58	0.04
	CCNN	Conjugate gradient	99.12	0.16	5.74	0.07
	MLP	Nguyen widrow/scaled conjugate gradient	99.24	0.15	4.29	0.04
PES	SVM	Grid and pattern search	98.7	0.20	4.36	0.08
	GRNN	Random/mixture of transfer function	96.99	0.39	6.69	0.12
	GMDH	Equation combination/function type	96.58	0.46	8.76	0.19
	CCNN	Conjugate gradient	96.58	0.46	8.76	0.19
	MLP	Nguyen widrow/scaled conjugate gradient	98.34	0.29	5.92	0.09
SKR	SVM	Grid and pattern search	99.64	0.08	4.27	0.03
	GRNN	Random/mixture of transfer function	99.4	0.11	4.69	0.03
	GMDH	Equation combination/function type	98.74	0.17	5.76	0.04
	CCNN	Conjugate gradient	95.75	0.17	6.25	0.06
	MLP	Nguyen widrow/scaled conjugate gradient	99.3	0.12	4.86	0.04

The maximum iterations for optimal model structure was chosen 10,000 and execution time was zero

each climatic zone which can be clearly observed in Table 10. This may be due to the fact that the climatic parameters in training process for all selected models excluding SVM generate higher residual values which become responsible to reduce their efficiencies in order to estimate ET_o . Furthermore, GMDH, MLP, CCNN and GRNN show acceptable values for all performing indices (M.E, MAE, MAPE and SI) in each selected climatic stations of Pakistan. An inclusive detail about each data driven model regarding performing indices for selected climatic stations of Pakistan is given in Table 10. In addition to summarized discussion, the model efficiencies (M.Es.) of SVM model for FSd, Jac, Mul, Pes and Skr are 95.4%, 99.06%, 99.36%, 98.7% and 99.64%,

respectively that stipulate the supremacy of SVM among other applied data driven models. Surprisingly, the efficiency of SVM for ISD station is lower as compared to other models but the least value of error (i.e. MAPE = 7.63%) further assist the aforementioned remarks regarding consistency and strength. The performance of each data driven model depends upon its training algorithm. It is difficult to decide which data driven model has the best performance except applied in different conditions especially in case of climatic studies. Although data driven models always performed better than local calibrated physical model or conventional methods (Karimaldini et al. 2011; Aytek 2009; Tabari and Talae 2013). But mostly, it was found that the model

considered as the best for one climatic region does not produce efficient result in another climatic condition. This is due to fact that the model optimal structure determined by using data driven model based on machine learning algorithms are of critical importance. Determination of the optimal structure of data driven model seems impossible without iteration process. Hence, various types of training algorithm with a number of iterations have been used to determine the best fitted optimal structure. In addition to use of training algorithms, the number of iterations, convergence value and execution times are considered as the prime parameters. Thus, data driven models based on some advanced machine learning algorithms e.g. pattern and search training algorithm in SVM model and/or equation combinations (linear, polynomial, quadratic) in GMDH model were applied in this study in order to investigate the performance.

4.3. Results of Developed Machine Learning Models in Other Climatic Regions

In section, developed data driven models were tested on five other stations located in China, New Zealand and USA (Illinois State) by using 4-year climatic data. The climatic data of these stations is also comprised of same meteorological input parameters. The statistical average values of all meteorological variables for selected countries calculated and presented in Table 8. Geographical description and data duration of each climatic station is mentioned in Table 2. climatic characteristics of these stations can be seen in Table 3. ET_o values of these stations were estimated by FAO-PM 56 standard equation. Selected data driven models, SVM, GMDH, MLP, CCNN and GRNN, directly applied on these climatic stations to estimate ET_o values without using their data in training phase. The performance indices, namely, CV, RMSE, r , and MSE are calculated for each station to investigate their performance. results of each data driven model are summarized in Table 11. As the results indicated (Table 11), the SVM model had the best performance in Auckland station ($r = 0.99$, $RMSE = 0.09$ mm/day, $CV = 0.17$ and $MSE = 0.12$ mm/day) among other models. However, the models GMDH and MLP had performed equally well with $r = 0.98$

and $RMSE = 0.13$ mm/day in the test phase. results generated with GRNN and CCNN models are comparatively lower than SVM, GMDH and MLP models. The GRNN model had yielded results as $r = 0.97$, $RMSE = 0.12$ mm/day, $CV = 0.19$, $MSE = 0.15$ mm/day while CCNN model had generated $r = 0.96$, $RMSE = 0.15$ mm/day, $CV = 0.23$ and $MSE = 0.22$ mm/day, respectively. It can be evidently observed that $RMSE$ values of GRNN and MLP models are remained identical i.e. 0.12 mm/day. In addition, the correlation between observed and predicted values for all data driven models had persisted above 0.95, although SVM model attains value equivalent to 1. Likewise, SVM model shows good performance for the Harbin station which is located in china country and have extremely cold climatic conditions. The values of $r = 1$ and $RMSE = 0.10$ mm/day makes SVM an ideal model for this climatic station while the values for CV and

Table 11

Results of each applied data driven model in testing phase for other five selected climatic stations

Climatic Stations	Numerical methods	Performing indices			
		CV	RMSE	r	MSE
Auckland	SVM	0.17	0.11	0.99	0.09
	GRNN	0.19	0.12	0.97	0.15
	GMDH	0.17	0.11	0.98	0.13
	CCNN	0.23	0.15	0.96	0.22
	MLP	0.17	0.12	0.98	0.13
Harbin	SVM	0.03	0.10	1.00	0.03
	GRNN	0.17	0.43	0.96	0.18
	GMDH	0.12	0.30	0.97	0.09
	CCNN	0.18	0.46	0.95	0.21
	MLP	0.06	0.15	0.98	0.06
Belleville	SVM	0.13	0.016	0.99	0.0001
	GRNN	0.17	0.021	0.96	0.0005
	GMDH	0.12	0.015	0.98	0.0002
	CCNN	0.13	0.026	0.95	0.0003
	MLP	0.17	0.022	0.94	0.0005
Big Bend	SVM	0.16	0.019	0.97	0.0002
	GRNN	0.24	0.025	0.92	0.0007
	GMDH	0.18	0.021	0.97	0.0004
	CCNN	0.23	0.026	0.91	0.0007
	MLP	0.34	0.039	0.92	0.0015
Dixon Spring	SVM	0.15	0.025	0.96	0.0003
	GRNN	0.22	0.022	0.90	0.0005
	GMDH	0.19	0.037	0.93	0.0014
	CCNN	0.25	0.029	0.89	0.0008
	MLP	0.21	0.023	0.92	0.0005

The maximum iterations for optimal model structure was chosen 10,000 and execution time considered as zero

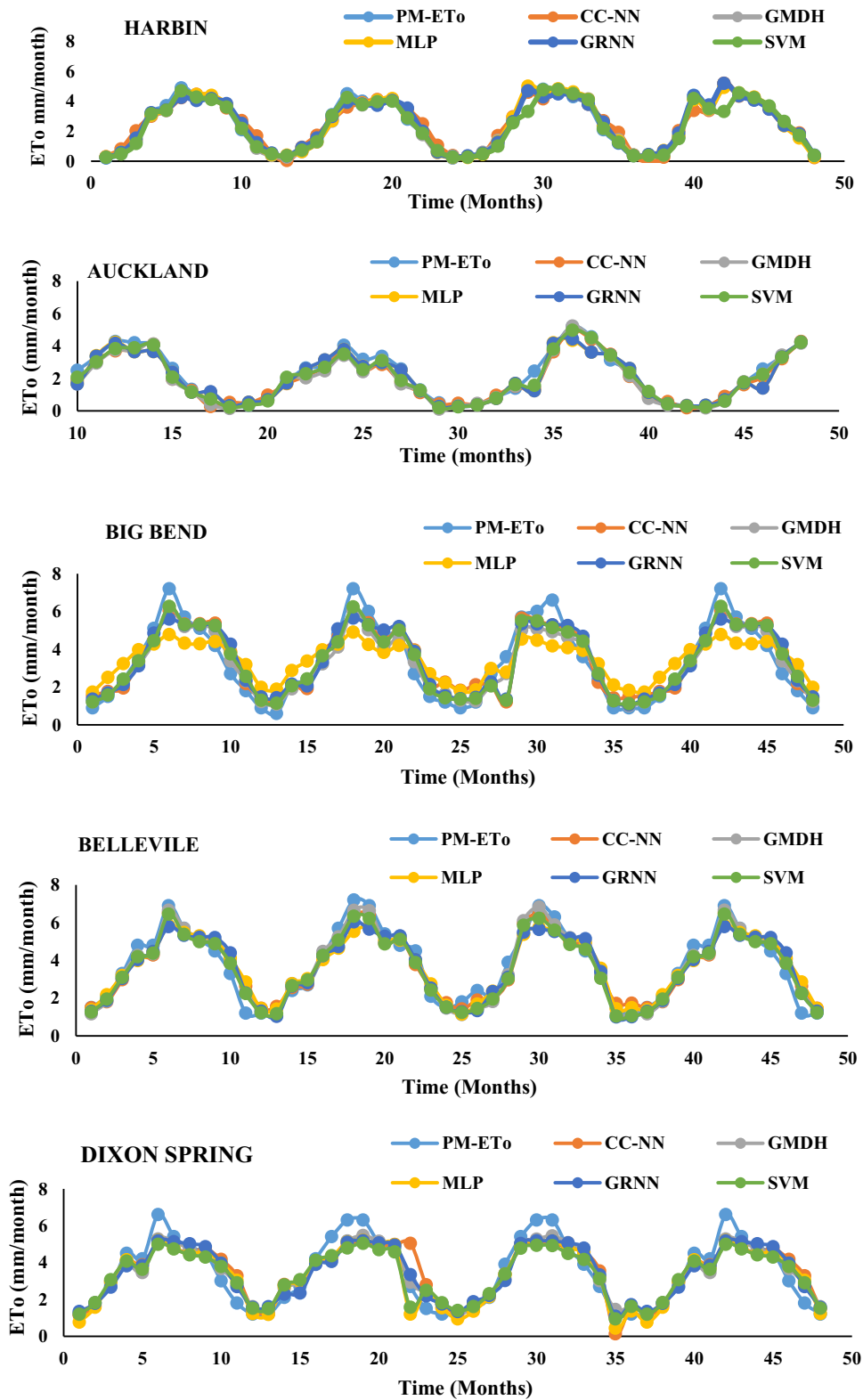


Figure 8
The ET₀ values estimated by the FAO-PM56 and other models in the test phase

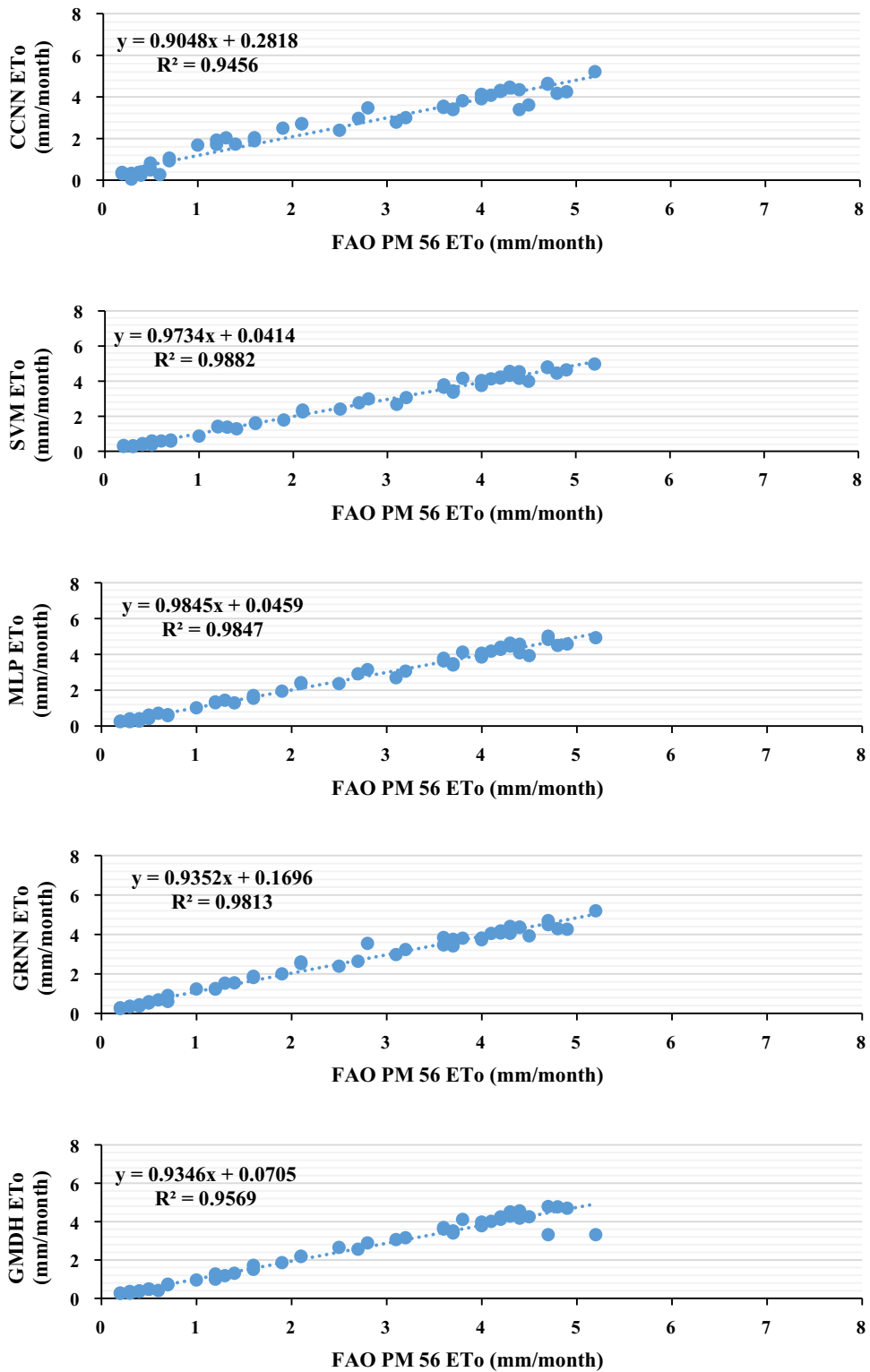


Figure 9
Scatter Plot for Harbin city (ET₀ of FAO-PM56 vs data driven models)

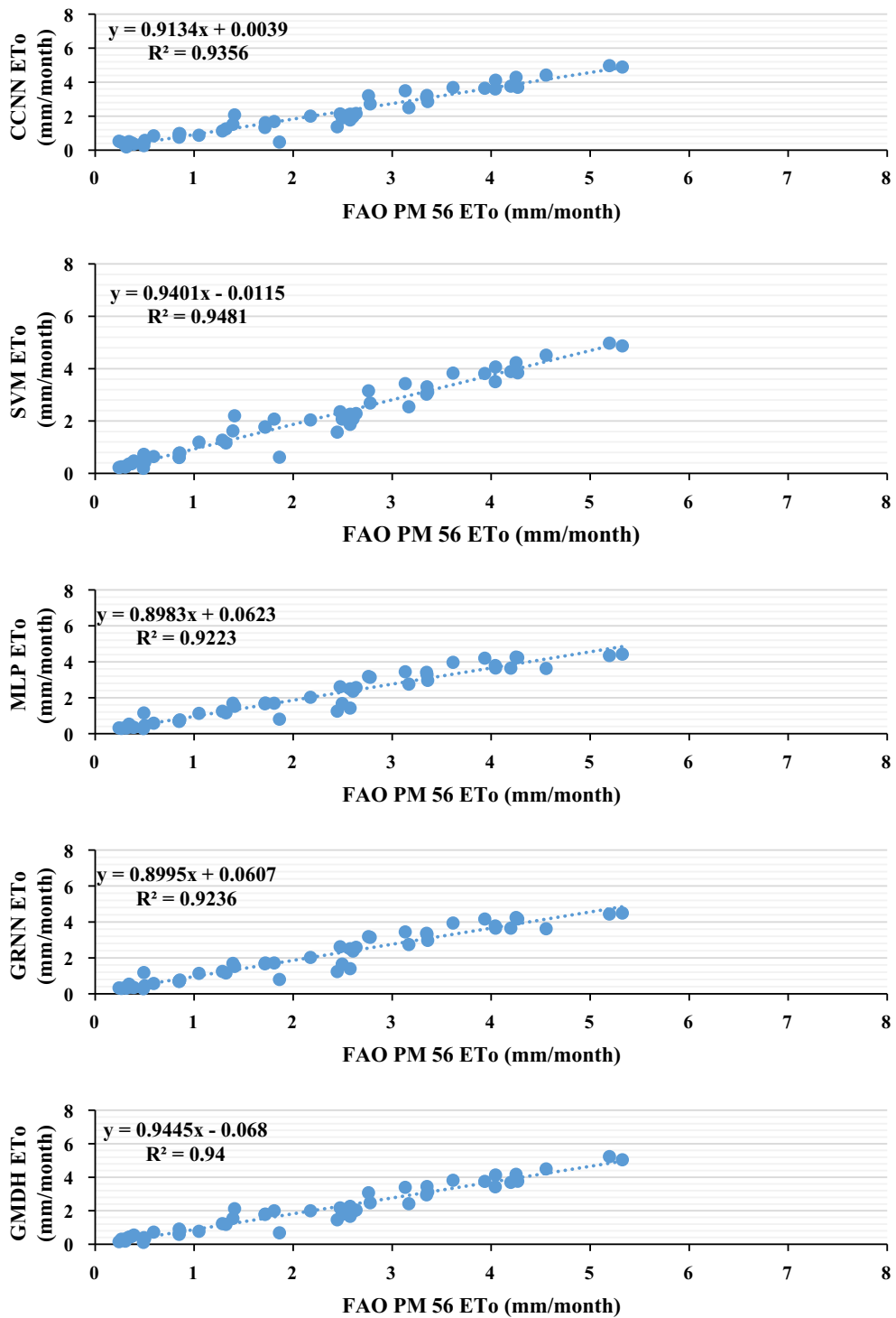


Figure 10
Scatter Plot for Auckland city (ET₀ of FAO-PM56 vs data driven models)

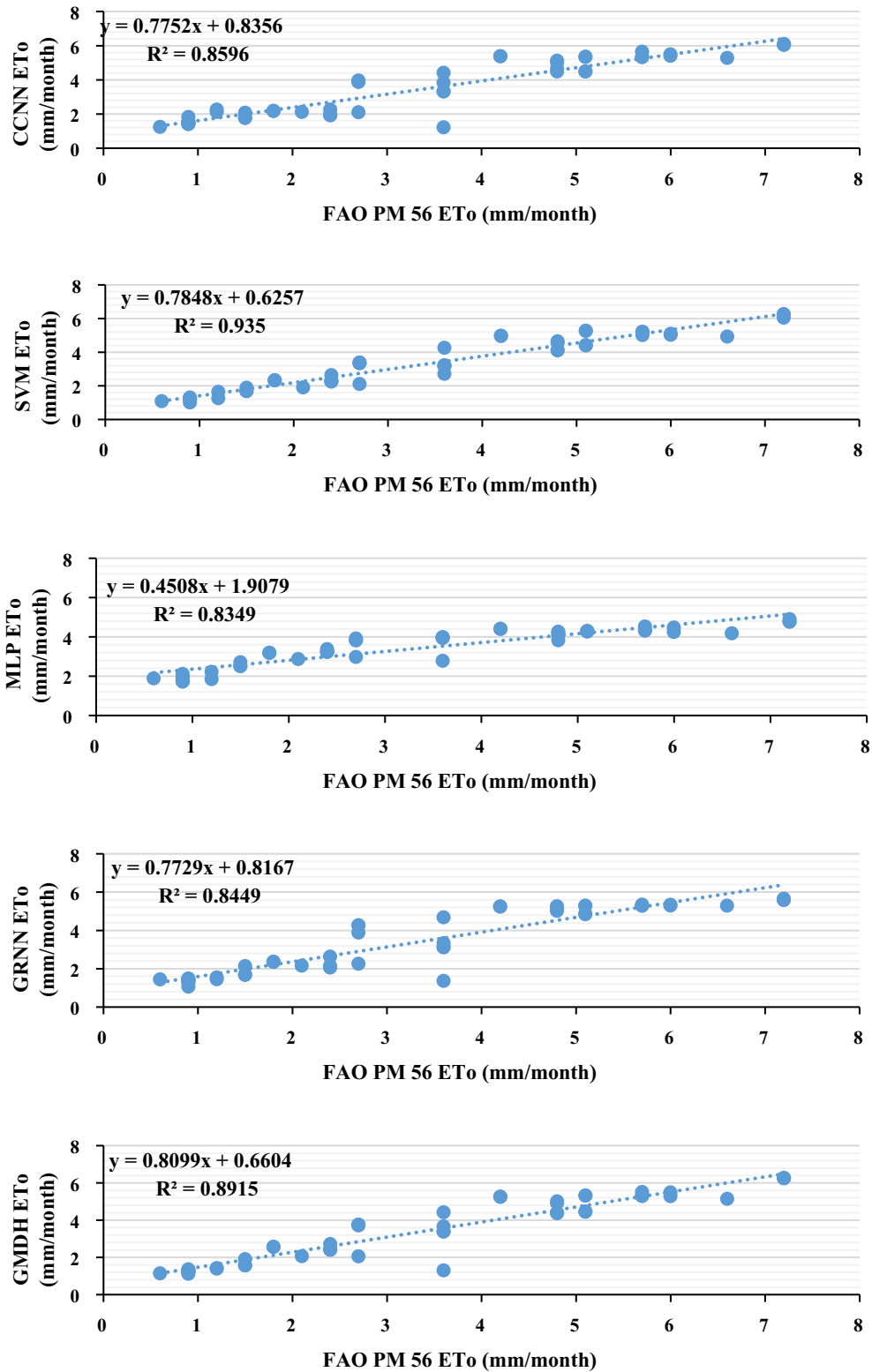


Figure 11
Scatter Plot for Big Bend city (ET₀ of FAO-PM56 vs data driven models)

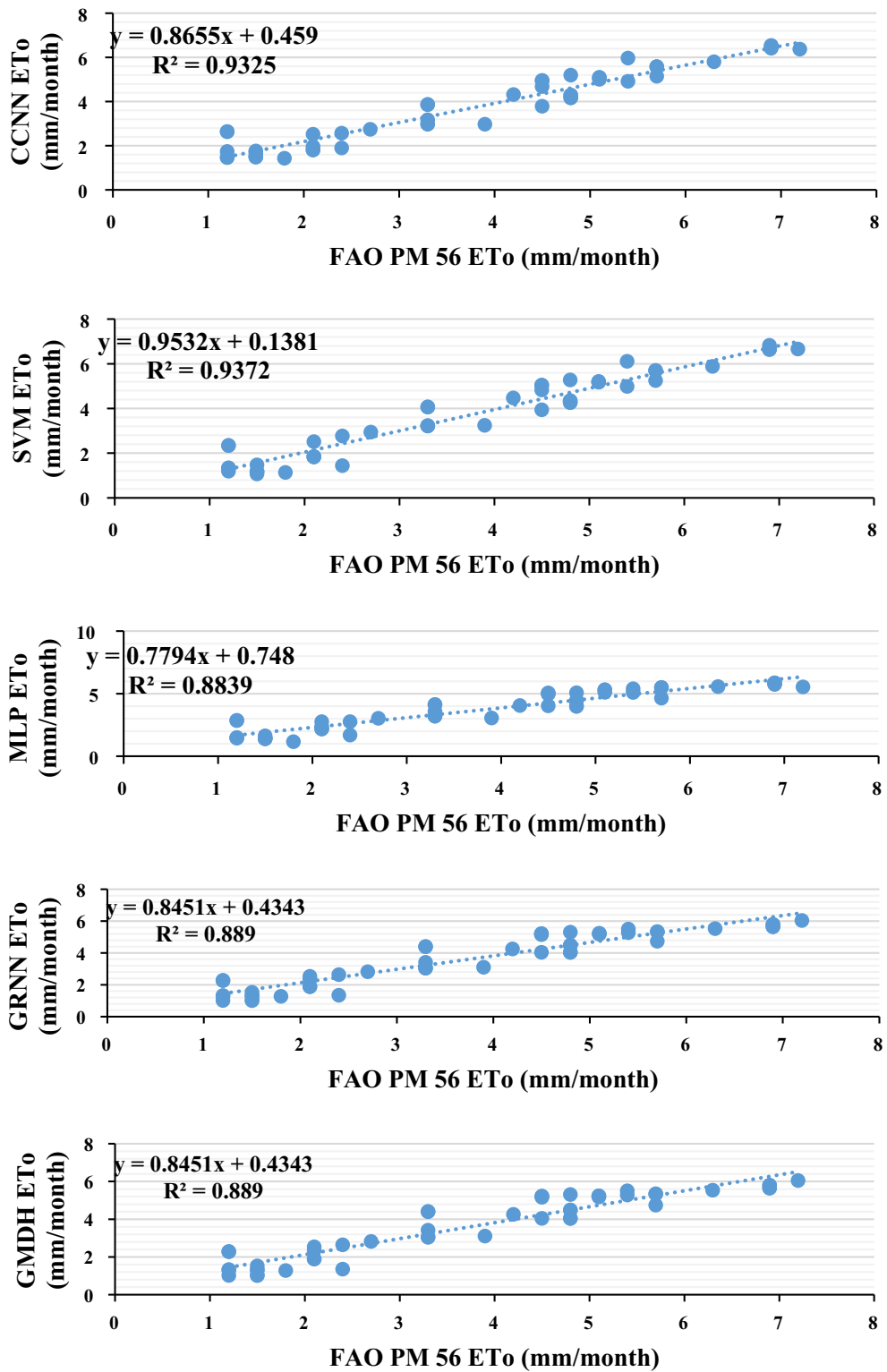


Figure 12
Scatter Plot for Belleville city (ET₀ of FAO-PM56 vs data driven models)

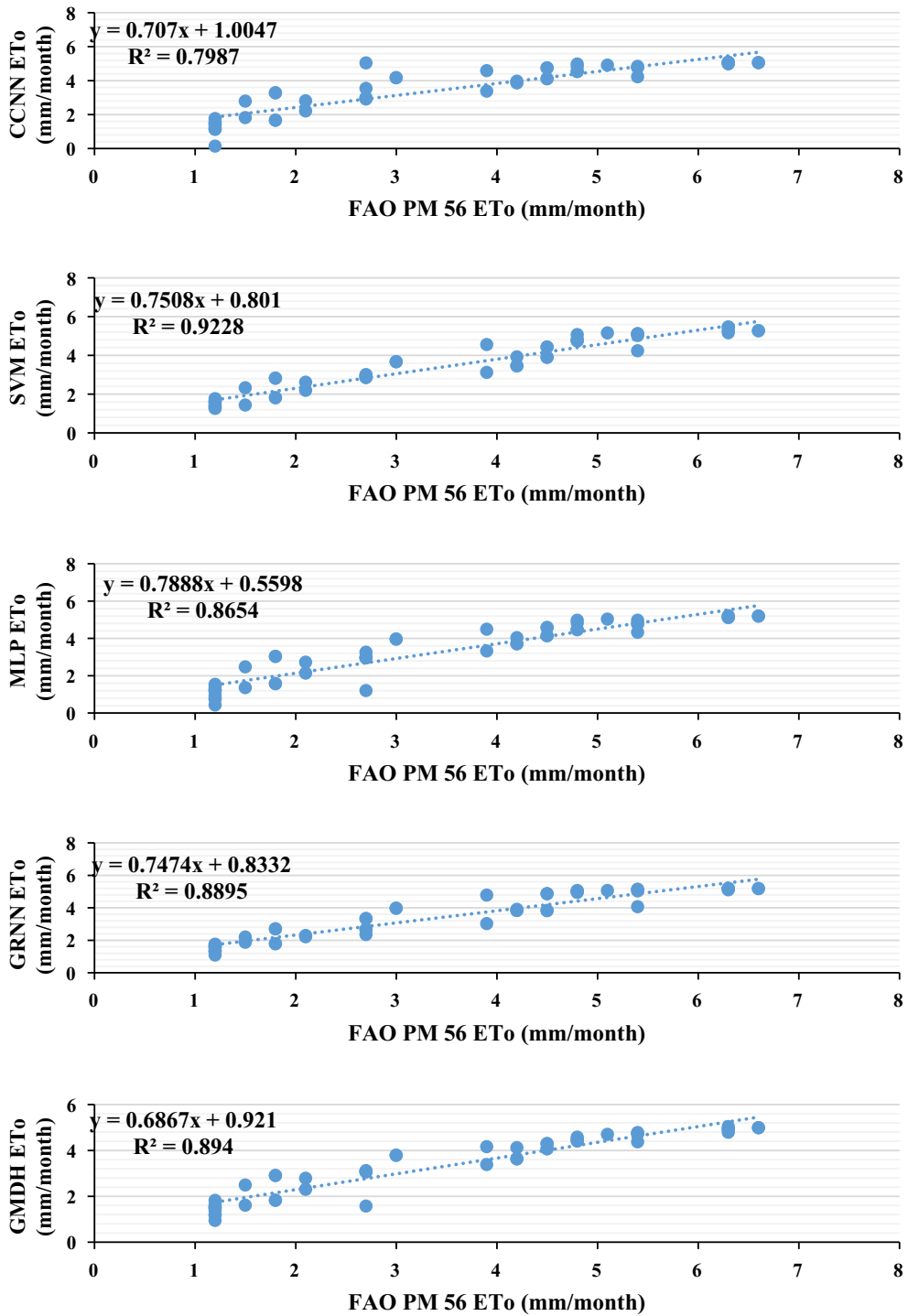


Figure 13
Scatter Plot for Dixon Spring city (ET₀ of FAO-PM56 vs data driven models)

MSE equal to 0 further legitimates it. It should be noted that r term provides information for the linear dependence between observations and corresponding estimates. Furthermore, MLP and GMDH models provide closer ET_o estimates to the benchmark ET_o than GRNN and CCNN models. The values of r and RMSE for both MLP and GMDH models are greater than GRNN and CNN models. However, MLP model had performed better as compared to GMDH model. As shown in Table 9, GMDH model generates exactly twice values for CV and RMSE which are followed as 0.12 and 0.30 mm/day, respectively. Although, the performance of GRNN and CCNN models are shown lower in comparison to SVM, MLP and GMDH models but yield good results in comparison to Auckland station. In addition, MSE value for each model is categorically zero while the value of correlation coefficient had attained 0.95 or above. Generally, it can be concluded from the performance evaluation indicators that SVM had the best performance and superior to other data driven models. The MLP model with a RMSE of 0.15 mm/day and MSE of 0.06 mm/day has provided the second best ET_o estimates. In contrast, the CCNN model has the highest error rates (RMSE = 0.46 mm/day and MSE = 0.21 mm/day). The outcomes of all data driven models are found in Table 11 for Belleville, Big Bend and Dixon Spring stations. According to the performance statistics, SVM model with extreme less values of CV, r and RMSE (CV = 0.13–0.16, r = 0.96–0.99 RMSE = 0.016–0.025 mm/day) is considered as the best among other data driven models for the Belleville, Big Bend and Dixon Spring stations. In addition, for all aforementioned climatic stations, SVM model has zero value for MSE which further supports its supremacy. It can also be inferred from results in Table 11 that the second best model for estimation of ET_o can be considered as a GMDH model. The outcomes of GMDH model are more precise as compared to MLP, GRNN and CCNN models. In addition, this model yields more accurate ET_o with increasing correlation coefficient from 0.96 to 0.99 and decreasing RMSE and MSE values from 0.025 and 0.0003 to 0.016 and 0.0001, respectively. From above analysis, it can be generally concluded that the performance of SVM model is superior to MLP, GMDH, GRNN and CCNN models for all five

selected climatic regions situated in USA, China and New Zealand. As mentioned earlier and can be seen from Table 4, radial basis function (RBF) was the best kernel for SVM model. Tabari et al. (2012) had found RBF as a best kernel function among others for all developed SVM models. A precise selection of kernel function is responsible for the given data set to provide feature space for its separation. This unique feature creates a boundary between SVM and other previously developed algorithms (Bray and Han 2004). Furthermore, Table 11 also revealed that GMDH is comparatively superior to MLP in Auckland while produces lower results as compared with MLP model in Harbin stations. In addition, MLP model had the best performance in Harbin station but generates poor results for Belleville and Big Bend stations with increasing values of CV and RMSE from 0.17 and 0.022 to 0.34 and 0.039, respectively. However, GMDH model could be endured the second best model in order to estimate ET_o . Consequently, it may be evident that GMDH model has utmost performance as compared to MLP, GRNN and CCNN models but inferior to SVM model. Moreover, GRNN model generates more reliable results than CCNN model by decreasing values of RMSE from 0.021 to 0.43 mm/day. Additionally, the values of correlation coefficient for CCNN model are decreasing from 0.96 to 0.89. Finally, it can be concluded that MLP is the third best model while GRNN and CCNN models are ranked as fourth and fifth best model for estimation of ET_o .

The mean monthly ET_o of five selected stations estimated by data driven model and FAO-PM56 method to authenticate model's performance. The hydrographical form of ET_o values for five selected stations can perceptibly be seen in Fig. 8. Initially, ET_o values were observed as low and then increased gradually when number of months increased. But after certain interval, ET_o goes in downward direction from peak value. Similarly, this could be happened for all intervals and observed lower and higher peaks of ET_o . Surely, this trend could be indicated winter and summer seasons. In winter season, temperature recorded minimum and relative humidity was maximum, the value of ET_o estimated minimum. Alternatively, for summer season, temperature recorded maximum and relative humidity was

Table 12

Comparison of FAO-PM56 and data driven models between monthly averages value of ET_o for selected stations

Countries	Stations	PM-method	CCNN	GMDH	MLP	GRNN	SVM
Pakistan	Skardu	3.395	3.368	3.389	3.370	3.355	3.390
USA	Belleville	3.825	3.570	3.780	3.630	3.499	3.820
	Big bend	3.391	3.487	3.378	3.655	3.332	3.389
	Dixon Spring	3.491	3.124	3.370	3.289	3.026	3.490
New Zealand	Auckland	2.176	1.879	2.017	1.980	1.799	2.164
China	Harbin	2.510	2.384	2.498	2.392	2.310	2.506

The value of ET_o calculated is in mm/month using FAO-PM56 and all the applied data driven models

minimum, ET_o value estimated maximum. To further check the underestimation and overestimation of ET_o , data driven models against FAO-PM56 method was compared. It can be perceived from Fig. 8 that MLP and CCNN models indicate overestimation while values of GRNN and GMDH models remain under peak of FAO-PM56 ET_o values for the Harbin station. For SVM data driven model, estimated ET_o is trailing well with FAO-PM56 ET_o values. In contrast, overestimation for MLP, GRNN and GMDH models in comparison to FAO-PM56 ET_o values was observed in Auckland station. However, the values of SVM model closely overlap with the standard FAO-PM56 ET_o values. Likewise, ET_o estimated by SVM model agrees with the FAO-PM56 ET_o and follows same trend in both Big bend and Belleville stations as observed in Fig. 8. Conversely, results of MLP model look extremely below from the FAO-PM56 ET_o values in Big bend station while remain equal with other models for Belleville station. Additionally, GMDH model has performed better than GRNN and CCNN models in both cities. This also authenticates our results for GMDH model indicated in Table 11. For Dixon Spring station, all data driven models generate smooth peaks for ET_o values as compared to FAO-PM56 method. This also indicates that these data driven models only deal with available data set and do not affect with other myriad factors. The deviation of ET_o values for SVM among data driven models was very minute and more close that estimated by FAO-PM56 method. Thus, it can be concluded from earlier discussion that SVM model performed the best in all climatic conditions for ET_o estimation. However, GMDH model generates reliable results as compared to MLP, GRNN and CNN

models, therefore ranked as the 2nd best model. While MLP, CCNN and GRNN models ranked 3rd place to 5th, respectively in the estimation of ET_o . The linear regression test facilitates the extension of statistical analysis. Figures 9, 10, 11, 12, 13 indicates ET_o estimates of all data driven models in form of scatter plots for each station. In these plots, ET_o estimated by the FAO-PM56 method are plotted against with numerical estimated ET_o values, in the form of scatter plots. It can be observed that there is a linear pattern of points for all applied data driven models. There were some misplaced estimated points which correspond to incorrectly estimated ET_o data, with some high values for all methods except SVM model. similar results were observed in Tabari et al. (2012). This may be due to intermittent character of ET_o distribution and deviate from trend line. It can be observed in Fig. 9 that data driven models generate acceptable R^2 values for the Harbin station (China) ranges from 94 to 98%. For SVM, MLP and GRNN data driven models, R^2 value found in equal (i.e. 98%) while CCNN and SVM models have 94% and 95% respectively. Likewise, ET_o values estimated by data driven models in comparison to FAO-PM56 methods for Auckland (New Zealand) is shown in Fig. 10. As indicated, SVM models performed the best with the highest value of R^2 (i.e. 95%) as compared to other data driven models. MLP and GRNN models performed equal while CCNN model generates lowest (i.e. 92%) value for Auckland station. In similar manner, comparison of applied numerical models against the standard FAO-PM56 method for Big Bend, Belleville, Dixon Spring (Illinois state, USA) are presented in Fig. 11, Fig. 12 and Fig. 13, respectively. SVM and GMDH

performed the best as compared to CCNN, MLP and GRNN model as observed in Fig. 11. The highest value of R^2 (93%) in Big Bend station was found for SVM model. Similarly, for Bellville (Fig. 12) and Dixon Spring stations (Fig. 13), SVM model outperformed in comparison to GRNN, CCNN, MLP. The R^2 value for SVM in Bellville and Dixon Springs climatic stations was nearly found 94% and 92% respectively. It is obvious from given fitted lines in all scatter plots (Figs. 9, 10, 11, 12, 13) that SVM model have closer to PM- ET_o values in comparison to other data driven models. Plotted points of SVM model correlate well towards linear trend line. However, plotted points of other data driven model are of major underestimation or overestimation. It can be deduced from Figs. 9, 10, 11, 12, 13 that GMDH model has the higher value of R^2 after SVM model for Auckland station (i.e. $R^2 = 94\%$). The superiority of GMDH model over MLP, GMDH, CCNN can also be noticeably comprehended in, Big Bend, Dixon Spring climatic stations. While MLP and CCNN performed better than GMDH model in the Harbin and Belleville climatic stations, respectively. Similar results of GMDH model can be found in Table 11 which demonstrates the supremacy of GMDH model over data driven models except a SVM model. It can be concluded that SVM model has the highest linear regression coefficient (R^2) in all types of climatic conditions and could be considered as the best data driven model for the ET_o estimation. From above discussions, GMDH and MLP data driven models are ranked as the 2nd and 3rd best models for estimation of ET_o . In contrast, GRNN and CCNN model have lower value of R^2 but can be acceptable for all climatic regions. Therefore, CCNN is ranked as 4th while GRNN model is considered as the 5th best model. It is evident from scatter plots that all methods have acceptable linear regression coefficient values.

For all data driven models, regression equation of each selected climatic station is found in Figs. 9, 10, 11, 12, 13. It can be perceived that SVM outperformed than other data driven models and shows higher correlation coefficient value in each selected station. The value of slope (a) and y-intercept (b) in regression equation ($Y = aX + b$) for SVM model was also found good, for "a" achieved higher and "b" look smaller, in comparison to other data driven

models (Figs. 9, 10, 11, 12, 13). In addition, a comparison of ET_o estimated by the standard conventional FAO-PM56 method and data driven models for selected countries is presented in Table 12 to conclude above mentioned results. Skardu station from Pakistan was included to make reasonable comparison among data driven models because it belongs to humid region. It can be concluded that all the data driven models shown good results when applied in different climatic conditions. Overall results indicate that SVM model outperformed for all selected climatic stations and yielded reliable results for ET_o estimation. This may be due to the fact that SVM model always seeks a global optimized solution, avoids over-fitting and no physical bias which eventually leads to the best performance over other data driven models.

5. Conclusions

In this study, the potential of five machine learning algorithms/data driven models (SVM, GMDH, CCNN, MLP and GRNN) were investigated to estimate long term monthly reference evapotranspiration (ET_o). For this purpose, meteorological variables of maximum and minimum temperatures, average relative humidity, wind speed and sunshine hours were employed as input in data driven models for six stations of Pakistan. The results obtained from all data driven models were found acceptable for selected climatic stations. Most of studies existed in literature used 70% of data as training/input data. And rest of 30% data for testing/validation. In this situation, R^2 mostly comes above 95%. This practice is not justified and always arose question regarding model's performance. Therefore, the data set from other countries have been used for validation and justification of each data driven model. The following conclusions can be drawn from the study:

- The long term monthly reference evapotranspiration (ET_o) of four climatic regions were successfully estimated by conventional standard FAO-PM56 method
- Machine learning algorithms/data driven models of SVM, GMDH, CCNN, MLP and GRNN were

found capable for ET_0 estimation under different climatic zones of Pakistan as well as in other parts of the world

- By investigating the performing indices, SVM and GMDH model ranked as 1st and 2nd while 3rd, 4th and 5th followed by MLP, CCCN and GRNN data driven models
- The comparison results indicate that by using the similar meteorological inputs, SVM can be employed as an alternative ET_0 model to the existing conventional method

Acknowledgements

The authors are thankful to the Directors/Assistant Directors of Pakistan meteorological department (PMD) for the provision of climatic data. Cordial thanks extend to the Editor and anonymous reviewers to improve this manuscript.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Alazba, A., Mattar, M., ElNesr, M., & Amin, M. (2011). Field assessment of friction head loss and friction correction factor equations. *Journal of Irrigation and Drainage Engineering*, 138(2), 166–176.
- Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao*, 300(9), D05109.
- Amarasinghe, U. A., & Smakhtin, V. (2014). *Global water demand projections: past, present and future*. Colombo: IWMI.
- Asefa, T., Kemblowski, M., McKee, M., & Khalil, A. (2006). Multi-time scale stream flow predictions: The support vector machines approach. *Journal of Hydrology*, 318(1–4), 7–16.
- Awari, H. (2018). Estimation of reference evapotranspiration and crop water requirement using artificial neural network for semi-arid. Ph.D, Vasantrao Naik Marathwada Krishi Vidyapeeth, Parbhani, Vasantrao Naik Marathwada Krishi Vidyapeeth, Parbhani.
- Aytek, A. (2009). Co-active neurofuzzy inference system for evapotranspiration modeling. *Soft Computing*, 13(7), 691.
- Azamatulla, H. M., Ghani, A. A., Leow, C. S., Chang, C. K., & Zakaria, N. A. (2011). Gene-expression programming for the development of a stage-discharge curve of the Pahang River. *Water Resources Management*, 25(11), 2901–2916.
- Barker, R., Dawe, D., Tuong, T., Bhuiyan, S., & Guerra, L. (1999). The outlook for water resources in the year 2020: challenges for research on water management in rice production. *Southeast Asia*, 1, 1–5.
- Bland, J. M., & Altman, D. G. (1996). Measurement error. *BMJ. British Medical Journal*, 312(7047), 1654.
- Bray, M., & Han, D. (2004). Identification of support vector machines for runoff modelling. *Journal of Hydroinformatics*, 6(4), 265–280.
- Chaturvedi, D. K. (2008). *Soft computing: Techniques and its applications in electrical engineering*. Berlin, Heidelberg: Springer. <https://doi.org/10.1007/978-3-540-77481-5>.
- Cobaner, M. (2011). Evapotranspiration estimation by two different neuro-fuzzy inference systems. *Journal of Hydrology*, 398(3–4), 292–302.
- Collobert, R., & Bengio, S. (2000). *Support vector machines for large-scale regression problems*. Technical Report IDIAP-RR00-17 IDIAP. Switzerland.
- Diamantopoulou, M., Georgiou, P., & Papamichail, D. (2011). Performance evaluation of artificial neural networks in estimating reference evapotranspiration with minimal meteorological data. *Global Nest Journal*, 13(1), 18–27.
- Doğan, E. (2009). Reference evapotranspiration estimation using adaptive neuro-fuzzy inference systems. *Irrigation and Drainage*, 58(5), 617–628.
- Doorenbos, J., & Pruitt, W. (1977). *Crop water requirements. FAO irrigation and drainage paper 24. Land and Water Development Division*. Rome: FAO.
- Droogers, P., & Allen, R. G. (2002). Estimating reference evapotranspiration under inaccurate data conditions. *Irrigation and Drainage Systems*, 16(1), 33–45.
- Gavilan, P., Berengena, J., & Allen, R. G. (2007). Measuring versus estimating net radiation and soil heat flux: impact on Penman-Monteith reference ET estimates in semiarid regions. *Agricultural water management*, 89(3), 275–286.
- Gocić, M., Motamedi, S., Shamshirband, S., Petković, D., Ch, S., Hashim, R., et al. (2015). Soft computing approaches for forecasting reference evapotranspiration. *Computers and Electronics in Agriculture*, 113, 164–173.
- Güven, A., Aytekin, A., Yuce, M. I., & Aksoy, H. (2008). Genetic Programming-based empirical model for daily reference evapotranspiration estimation. *Clean-Soil Air Water*, 36(10–11), 905–912.
- Han, J., Kamber, M., & Pei, J. (2011). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann Publishers.
- Hassanpour Kashani, M., Montaseri, M., & Lotfollahi Yaghin, M. A. (2008). Flood estimation at ungauged sites using a new hybrid model. *Journal of Applied Sciences*, 8, 1744–1749.
- Hess, T. (1996). *Potential Evapotranspiration [DAILYET]*. UK: Silsoe College.
- Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, 23(4), 725–749. <https://doi.org/10.1080/01431160110040323>.
- Huang, Y., Lan, Y., Thomson, S. J., Fang, A., Hoffmann, W. C., & Lacey, R. E. (2010). Development of soft computing and applications in agricultural and biological engineering. *Computers and Electronics in Agriculture*, 71(2), 107–127.
- Ivakhnenko, A. G. (1968). The group method of data of handling: a rival of the method of stochastic approximation. *Soviet Automatic Control*, 13, 43–55.

- Jorguseski, L., Pais, A., Gunnarsson, F., Centonza, A., & Willcock, C. (2014). Self-organizing networks in 3GPP: Standardization and future trends. *IEEE Communications Magazine*, 52(12), 28–34.
- Karimaldini, F., Teang Shui, L., Ahmed Mohamed, T., Abdollahi, M., & Khalili, N. (2011). Daily evapotranspiration modeling from limited weather data by using neuro-fuzzy computing technique. *Journal of Irrigation and Drainage Engineering*, 138(1), 21–34.
- Kim, S., & Kim, H. S. (2008). Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling. *Journal of Hydrology*, 351(3), 299–317.
- Kisi, O. (2013). Least squares support vector machine for modeling daily reference evapotranspiration. *Irrigation Science*, 31(4), 611–619.
- Kisi, O., & Cimen, M. (2009). Evapotranspiration modelling using support vector machines/Modélisation de l'évapotranspiration à l'aide de 'support vector machines'. *Hydrological Sciences Journal*, 54(5), 918–928.
- Kişİ, Ö., & Öztürk, Ö. (2007). Adaptive neurofuzzy computing technique for evapotranspiration estimation. *Journal of Irrigation and Drainage Engineering*, 133(4), 368–379.
- Kisi, O., & Sanikhani, H. (2015). Prediction of long-term monthly precipitation using several soft computing methods without climatic data. *International Journal of Climatology*, 35(14), 4139–4150.
- Kisi, O., & Zounemat-Kermani, M. (2014). Comparison of two different adaptive neuro-fuzzy inference systems in modelling daily reference evapotranspiration. *Water Resources Management*, 28(9), 2655–2675.
- Kumar, M., Raghuvanshi, N., Singh, R., Wallender, W., & Pruitt, W. (2002). Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering*, 128(4), 224–233.
- Landeras, G., Ortiz-Barredo, A., & López, J. J. (2008). Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). *Agricultural Water Management*, 95(5), 553–565.
- Legates, D. R., & McCabe, G. J., Jr. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35(1), 233–241.
- Li, R. Y. M., Fong, S., & Chong, K. W. S. (2017). Forecasting the REITs and stock indices: Group method of data handling neural network approach. *Pacific Rim Property Research Journal*, 23(2), 123–160.
- Lopez-Urrea, R., de Santa Olalla, F. M., Fabeiro, C., & Moratalla, A. (2006). An evaluation of two hourly reference evapotranspiration equations for semiarid conditions. *Agricultural Water Management*, 86(3), 277–282.
- Lu, W.-Z., & Wang, W.-J. (2005). Potential assessment of the "support vector machine" method in forecasting ambient air pollutant trends. *Chemosphere*, 59(5), 693–701.
- Lumia, D. S., Linsey, K. S., & Barber, N. L. (2005). *Estimated use of water in the United States in 2000: US Department of the Interior US Geological Survey*. Reston: US Geological Survey.
- Markowetz, F. (2003). Classification by support vector machines. https://www.academia.edu/download/3241129/Kapitel_16.pdf.
- Mattar, M. A., & Alamoud, A. I. (2015). Artificial neural networks for estimating the hydraulic performance of labyrinth-channel emitters. *Computers and Electronics in Agriculture*, 114, 189–201.
- Mattar, M., Alazba, A., & El-Abedin, T. Z. (2015). Forecasting furrow irrigation infiltration using artificial neural networks. *Agricultural Water Management*, 148, 63–71.
- McMahon, T., Peel, M., Lowe, L., Srikanthan, R., & McVicar, T. (2013). Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: A pragmatic synthesis. *Hydrology and Earth System Sciences*, 17(4), 1331.
- Møller, M. F. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6(4), 525–533.
- Najafzadeh, M., Barani, G.-A., & Azamathulla, H. M. (2014). Prediction of pipeline scour depth in clear-water and live-bed conditions using group method of data handling. *Neural Computing and Applications*, 24(3–4), 629–635.
- Østerbø, O., & Grøndalen, O. (2012). Benefits of self-organizing networks (SON) for mobile operators. *Journal of Computer Networks and Communications*. <https://doi.org/10.1155/2012/862527>.
- Peng, M., Liang, D., Wei, Y., Li, J., & Chen, H.-H. (2013). Self-configuration and self-optimization in LTE-advanced heterogeneous networks. *IEEE Communications Magazine*, 51(5), 36–45.
- Pereira, A. R., & Pruitt, W. O. (2004). Adaptation of the Thornthwaite scheme for estimating daily reference evapotranspiration. *Agricultural Water Management*, 66(3), 251–257.
- Pereira, L. S., Perrier, A., Allen, R. G., & Alves, I. (1999). Evapotranspiration: concepts and future trends. *Journal of Irrigation and Drainage Engineering*, 125(2), 45–51.
- Raes, D. (2009). ETo Calculator: a software program to calculate evapotranspiration from a reference surface. FAO Land Water Division: Digital Media Service
- Rahimikhoob, A. (2010). Estimation of evapotranspiration based on only air temperature data using artificial neural networks for a subtropical climate in Iran. *Theoretical and Applied Climatology*, 101(1–2), 83–91.
- Sherrod, P. H. (2008). DTREG predictive modeling software. Users Manual. Disponível online no url: (<https://www.dtreg.com/DTREG.pdf>) (Acedido 28 Março 2014).
- Shiri, J., Marti, P., & Singh, V. P. (2014a). Evaluation of gene expression programming approaches for estimating daily evaporation through spatial and temporal data scanning. *Hydrological Processes*, 28(3), 1215–1225.
- Shiri, J., Nazemi, A. H., Sadraddini, A. A., Landeras, G., Kisi, O., Fard, A. F., et al. (2014b). Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. *Computers and Electronics in Agriculture*, 108, 230–241.
- Shiri, J., Sadraddini, A. A., Nazemi, A. H., Kisi, O., Landeras, G., Fard, A. F., et al. (2014c). Generalizability of Gene Expression Programming-based approaches for estimating daily reference evapotranspiration in coastal stations of Iran. *Journal of Hydrology*, 508, 1–11.
- Shrestha, N., & Shukla, S. (2015). Support vector machine based modeling of evapotranspiration using hydro-climatic variables in a sub-tropical environment. *Agricultural and Forest Meteorology*, 200, 172–184.
- Specht, D. F. (1990). Probabilistic neural networks. *Neural Networks*, 3(1), 109–118.
- Sudheer, K., Gosain, A., & Ramasastri, K. (2003). Estimating actual evapotranspiration from limited climatic data using neural

- computing technique. *Journal of Irrigation and Drainage Engineering*, 129(3), 214–218.
- Swennenhuis, J. (2009). *CROPWAT 8.0*. Rome: Water Resources Development and Management Service of FAO.
- Tabari, H., Kisi, O., Ezani, A., & Talaei, P. H. (2012). SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. *Journal of Hydrology*, 444, 78–89.
- Tabari, H., Martinez, C., Ezani, A., & Talaei, P. H. (2013). Applicability of support vector machines and adaptive neuro-fuzzy inference system for modeling potato crop evapotranspiration. *Irrigation Science*, 31(4), 575–588.
- Tabari, H., & Talaei, P. H. (2013). Multilayer perceptron for reference evapotranspiration estimation in a semiarid region. *Neural Computing and Applications*, 23(2), 341–348.
- Todorovic, M., Karic, B., & Pereira, L. S. (2013). Reference evapotranspiration estimate with limited weather data across a range of Mediterranean climates. *Journal of Hydrology*, 481, 166–176.
- Trajkovic, S. (2005). Temperature-based approaches for estimating reference evapotranspiration. *Journal of Irrigation and Drainage Engineering*, 131(4), 316–323.
- Trajkovic, S., & Kolakovic, S. (2009). Estimating reference evapotranspiration using limited weather data. *Journal of Irrigation and Drainage Engineering*, 135(4), 443–449.
- Traore, S., & Guven, A. (2012). Regional-specific data driven models of evapotranspiration using gene-expression programming interface in Sahel. *Water Resources Management*, 26(15), 4367–4380.
- Traore, S., Wang, Y.-M., & Kerh, T. (2010). Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone. *Agricultural Water Management*, 97(5), 707–714.
- Turing, A. (1950). *Mind*. *Mind*, 59(236), 433–460.
- Vapnik, V. (1998). *Statistical learning theory*. New York: Wiley.
- Ventura, F., Spano, D., Duce, P., & Snyder, R. (1999). An evaluation of common evapotranspiration equations. *Irrigation Science*, 18(4), 163–170.
- Wu, C., Chau, K., & Li, Y. (2008). River stage prediction based on a distributed support vector regression. *Journal of Hydrology*, 358(1–2), 96–111.
- Xu, C.-Y., & Singh, V. (2002). Cross comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland. *Water Resources Management*, 16(3), 197–219.

(Received June 20, 2019, revised February 2, 2020, accepted March 20, 2020, Published online April 2, 2020)