ORIGINAL PAPER

Evaluation of multivariate linear regression for reference evapotranspiration modeling in different climates of Iran

Saeed Sharafi[1] · Mehdi Mohammadi Ghaleni[2]

Abstract

The study aimed to evaluate the accuracy of empirical equations (Hargreaves-Samani; HS, Irmak; IR and Dalton; DT) and multivariate linear regression models (MLR1–6) for estimating reference evapotranspiration (ET_{Ref}) in different climates of Iran based on the Köppen method including arid desert (Bw), semiarid (Bs), humid with mild winters (C), and humid with severe winters (D). For this purpose, climatic data of 33 meteorological stations during 30 statistical years 1990–2019 were used with a monthly time step. Based on various meteorological data (minimum and maximum temperature, relative humidity, wind speed, solar radiation, extraterrestrial radiation, and vapor pressure deficit), in addition to 6 multivariate linear regression models and three empirical equations were used as MLR1, MLR2, and HS (temperature-based), MLR3 and IR (radiation-based), MLR4, MLR5 and DT (mass transfer-based), and MLR6 (combination-based) were also used to estimate the reference evapotranspiration. The results of these models were compared using the root mean square error (RMSE), mean absolute error (MAE), scatter index (SI), determination coefficient (R^2), and Nash-Sutcliffe efficiency (NSE) statistical criteria with the evapotranspiration results of the FAO_{56} Penman-Monteith reference as target data. All MLR models gave better results than empirical equations. The results showed that the simplest regression model (MLR1) based on the minimum and maximum temperature data was more accurate than the empirical equations. The lowest and highest accuracy related to the MLR6 model and HS empirical equation with RMSE was 10.8–15.1 mm month^{-1} and 22–28.3 mm month^{-1}, respectively. Also, among all the evaluated equations, radiation-based models such as IR in Bw and Bs climates with MAE = 8.01–11.2 mm month^{-1} had higher accuracy than C and D climates with MAE = 13.44–14.48 mm month^{-1}. In general, the results showed that the ability of regression models was excellent in all climates from Bw to D based on SI < 0.2.

1 Introduction

Accurate estimation of reference evapotranspiration (ET_{Ref}) is one of the priorities for estimating the water requirement of agricultural products. The complexity of the evapotranspiration process and its dependence on meteorological variables, lack of access to all meteorological data, and the lack of generalizability of a model for different climates have made it difficult to accurately estimate this variable. Meanwhile, the development of simple models based on multivariate linear

regression due to the simplicity and the ability to select different models according to the available meteorological variables can help in estimating the appropriate ET_{Ref} in different climates. In general, the estimation of ET_{Ref} can be done through direct and indirect methods and using mathematical models (Jing et al. 2019). Indirect methods for estimating ET_{Ref} began with the development of empirical equations such as Penman-Monteith (PM) (Monteith 1965). Based on the PM method, FAO_{56} published a standard method for estimating ET_{Ref}. The PM method presented in FAO_{56} has been used by many researchers as the base method in calculating ET_{Ref} (Güçlü et al. 2017; Saggi and Jain 2019; Shiri et al. 2019).

In recent years, however, artificial intelligence-based methods such as the neural networks (Kisi et al. 2015; Gavili et al. 2018), the support vector machine (Tabari et al. 2013), the extreme learning machine (Abdullah et al. 2015), decision tree (Huang et al. 2019; Raza et al. 2020), and hybrid methods (Ehteram et al. 2019; Shiri et al. 2020; Tikhamarine et al. 2019; Zhu et al. 2020; Kim et al. 2014; Sanikhani et al.

✉ Saeed Sharafi
s-sharafi@araku.ac.ir

1 Department of Environment Science and Engineering, Arak University, Arak, Iran

2 Department of Water Science and Engineering, Arak University, Arak, Iran

2019; Mehdizadeh et al. 2017) have had many applications in estimating ET_{Ref}, but among them, multivariate linear regression method has been compared with other empirical equations and soft computing, validated by many researchers (Reis et al. 2019; Kisi and Heddam 2019; Mattar and Alazba 2019; Tabari et al. 2012).

The appropriate method for estimating ET_{Ref} in each region depends on climatic conditions, required data, and related costs (Sharafi et al. 2016). Accordingly, Unes et al. (2020) predicted daily ET_{Ref} based on climatic conditions using empirical equations (Hargreaves-Samani, Ritchie, Turc, and Penman FAO 56), multilinear regression (MLR), and different data mining techniques (M5T, ANFIS, SVM). According to their results, Turc empirical formula (radiation-based) is found better than other empirical equations and the highest correlation coefficient is calculated for ANFIS, and the minimum errors are calculated for radial basis function SVM. Also, Chen et al. (2020) estimate daily ET_{Ref} based on limited meteorological data using three deep learning methods, two classical machine learning methods, and seven empirical equations. Their results show that, when temperature-based features were available, the deep learning models performed markedly better than temperature-based empirical models, and when radiation-based or humidity-based features were available, all of the proposed deep and classical learning machine models outperformed radiation-based or humidity-based empirical equations beyond the study areas.

Yirga (2019) estimated ET_{Ref} by multiple linear regression. The results of the MLR model could estimate and predict ET_{Ref} in the Megecha basin and can be used in similar parts of data-sparse regions. Karbasi (2018) investigated the Gaussian process regression (GPR) and Wavelet-GPR models to forecast multi-step ahead daily ET_{Ref} at the synoptic station of Zanjan (Iran). The results of the Wavelet-GPR model showed that the performance of the model during the warmer season is better than its performance throughout the year.

dos Santos Farias et al. (2020) evaluated the performance of machine learning techniques and stepwise multiple linear regression method to estimate daily ET_{Ref} with limited weather data in a Brazilian agricultural frontier. Their results showed that machine learning methods are robust in estimating ET_{Ref}, even in the absence of some variables. On the other hand, the use of artificial intelligence models in estimating ET_{Ref} with high accuracy has become prevalent in recent years, but the complexity of these models makes their application difficult for different regions. For this purpose, the present paper uses multivariate linear regression models with different data input to develop a simple comprehensive model with the minimum data required to estimate the ET_{Ref} in different climates of Iran. Although in most of the previous researches, the accuracy of the results has been investigated based on meteorological stations (points), in this research, the accuracy of multivariate linear regression models has been investigated based on different climates in Iran. This innovation extends the results of this paper to similar climates around the world, and also reduces the uncertainty of results as a result of fluctuations in climate variables on a point-by-point basis.

2 Materials and methods

2.1 Study area

Iran is located in the northern hemisphere between two longitudes of eastern 44° and 64° and two latitudes of northern 40° and 25°. Meteorological data of different synoptic stations, including different climates, are collected and analyzed. Some records of data input were incomplete, or not available for some stations; therefore, only stations with long statistical period length remained. Accordingly, in the present study, meteorological data of 33 synoptic stations in Iran over a statistical period of 30 years (1990–2019) are used. Figure 1 shows the location of the studied stations along with climatic classification based on the Köppen method (Köppen 1931) and according to the values of air temperature and precipitation. According to the study of Sharafi and Karim (2020) (the Köppen climate classification), 7 stations (Bandar Abbas, Abadan, Ahwaz, Bushehr, Bam, and Zabol), 7 stations (Zahedan, Semnan, Sabzevar, Kerman, Isfahan, Shahroud and Torbat-e Heydarieh), 13 stations (Shiraz, Gorgan, Tehran, Khorramabad, Bandar Anzali, Birjand, Rasht, Kermanshah, Mashhad, Arak, Qazvin, Sanandaj, and Tabriz), and 6 stations (Shahrekord, Khoy, Saqez, Urmia, Hamedan, and Zanjan) are located in the arid desert (Bw), semiarid (Bs), humid with very mild winters (C), and humid with severe winters (D) zones, respectively (Fig. 1).

To calculate ET_{Ref} values using empirical equations and regression models, the monthly data, parameters of minimum temperature (T_{min}), maximum temperature (T_{max}), relative humidity (RH), wind speed (u), solar radiation (R_s), extraterrestrial radiation (R_a), and vapor pressure deficit (VPD) were used in different stations. The statistical characteristics of these parameters are presented for different climates separately in Table 1.

According to the results in Table 1, the average monthly ET_{Ref} values decreased from 123.3 to 90.5 mm month^{-1}, respectively, from arid desert (Bw) to humid climate with severe winters (D). Also, the highest and lowest coefficient of variation (CV%) of ET_{Ref} changes in Bw and D climates was reported to be equal to 39.2% and 53.8%, respectively. The coefficients of variation of minimum and maximum
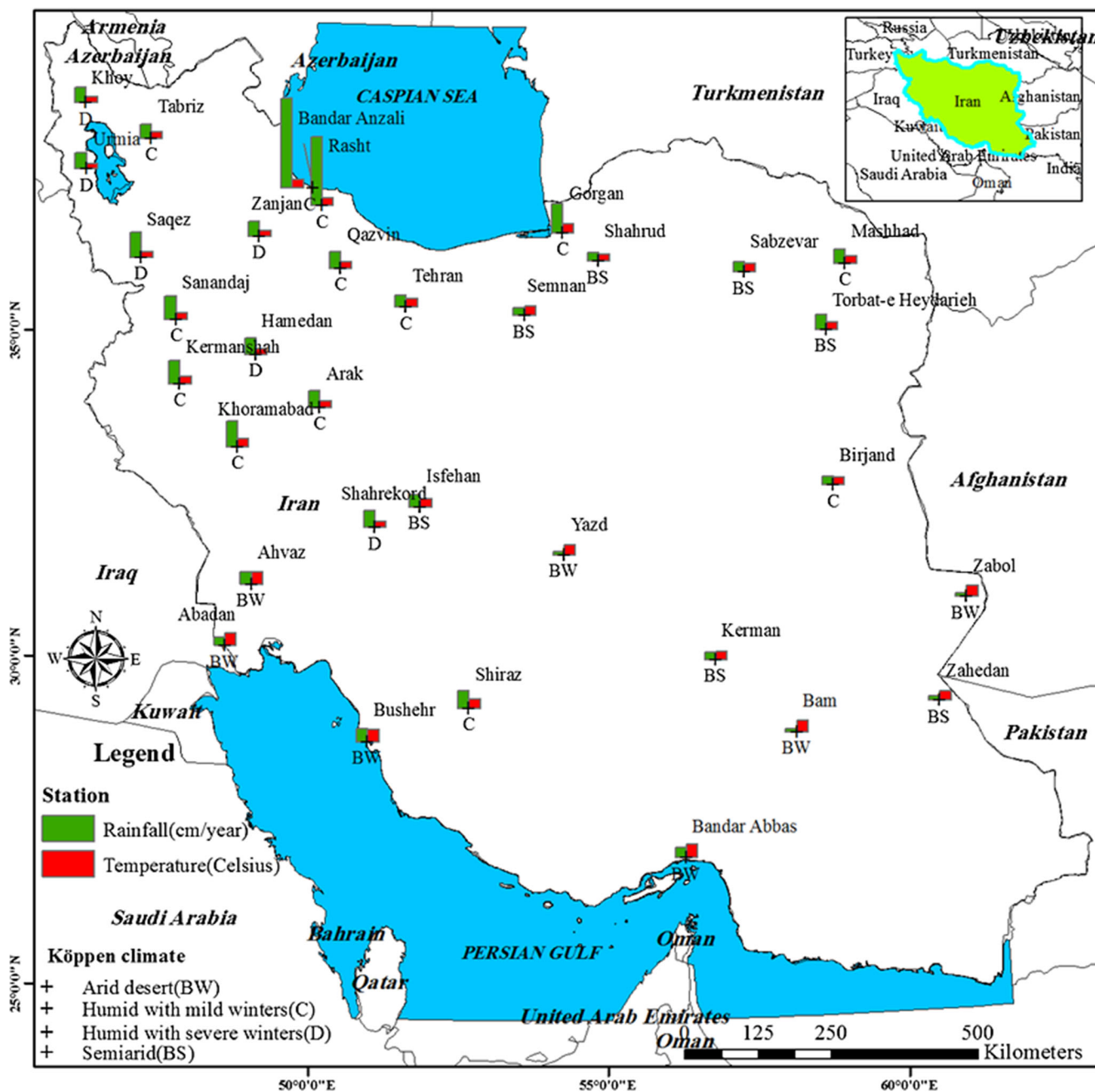
Fig. 1 Location and classification of studied stations based on the Köppen method

temperatures similar to ET_{Ref} increased from Bw to D climates (Table 1).

R, rainfall; *RH*, relative humidity; T_{min}, minimum temperature; T_{max}, maximum temperature; *u*, wind speed; R_s, solar radiation; R_a, extraterrestrial radiation; ET_{Ref}, reference evapotranspiration by Penman-Monteith (Allen et al. 1998); *Min*, minimum; *Max*, maximum; *Ave*, average; *SD*, standard deviation; *CV (%)*, coefficient of variation; *Kur*, kurtosis; *Sk*, skewness

In Fig. 2, the Pearson correlation coefficient values between ET_{Ref} and other meteorological variables (average, maximum and minimum temperature, vapor pressure deficient, wind speed, relative humidity, solar radiation, and extraterrestrial radiation) are shown for 33 study stations based on Köppen method. Pearson correlation coefficient values of $+1$ and -1 show the highest correlation with a direct and inverse relationship between the dependent variable (ET_{Ref}) and independent variable (meteorological variables), respectively. Figure 2 shows that ET_{Ref} were directly related to all meteorological variables. The only correlation coefficient between ET_{Ref} and relative humidity was negative, indicating an inverse relationship between

Table 1 Statistical characteristics of meteorological variables

Köppen climate	Variable	Min	Max	Ave	SD	CV%	Kur	Sk
Arid desert (Bw)	T_{min}	1.3	30.2	17.1	7.8	48.0	−1.4	−0.1
	T_{max}	8.6	43.9	30.2	8.8	29.4	−1.2	−0.3
	R_a	1.9	12.8	8.5	1.8	21.7	0.0	−0.2
	R_s	1.0	7.9	5.0	1.3	25.8	−0.2	−0.2
	u	1.3	15.5	6.0	2.5	39.4	0.2	0.5
	RH	20.4	79.2	45.1	12.8	32.3	−0.5	0.3
	VPD	0.6	7.2	3.7	1.7	43.3	−1.3	0.1
	R	0.0	148.9	13.0	23.8	184.9	7.8	2.6
	ET_{Ref}	42.7	219.0	123.3	48.7	39.2	−1.3	0.1
Semiarid (Bs)	T_{min}	−6.4	20.3	8.0	7.1	79.6	−1.1	0.0
	T_{max}	3.7	33.7	20.5	8.4	36.3	−1.2	−0.2
	R_a	1.2	10.7	7.0	2.0	25.2	−0.5	−0.1
	R_s	0.6	7.2	4.4	1.5	29.3	−0.8	−0.1
	u	0.7	9.6	3.7	1.5	38.5	0.4	0.5
	RH	14.3	66.1	34.6	12.9	33.2	−0.8	0.4
	VPD	0.6	5.2	2.6	1.3	43.8	−1.2	0.2
	R	0.0	99.2	13.4	17.7	120.2	3.5	1.6
	ET_{Ref}	49.2	215.8	120.9	44.3	36.6	−1.3	−0.1
Humid with mild winters (C)	T_{min}	−7.9	21.1	8.2	7.1	88.8	−1.1	0.0
	T_{max}	0.2	34.7	20.2	9.1	42.4	−1.2	−0.1
	R_a	1.8	11.1	7.0	2.2	30.6	−0.8	0.1
	R_s	0.8	7.7	4.2	1.7	37.6	−0.8	0.1
	u	0.6	8.8	3.6	1.4	38.2	0.9	0.7
	RH	23.7	79.1	49.9	13.6	28.9	−0.6	0.0
	VPD	0.3	5.4	2.4	1.4	51.6	−1.0	0.4
	R	0.0	250.2	42.5	43.5	103.5	2.3	1.3
	ET_{Ref}	16.4	199.9	97.3	53.7	50.9	−1.2	0.2
Humid with severe winters (D)	T_{min}	−15.0	15.4	3.2	6.7	197.3	−0.7	−0.2
	T_{max}	−3.5	31.0	15.8	9.3	50.7	−1.1	−0.1
	R_a	1.6	10.5	6.5	2.2	28.6	−0.8	0.0
	R_s	0.8	7.9	4.3	1.8	36.2	−1.0	0.1
	u	0.3	7.1	2.8	1.3	41.4	0.0	0.4
	RH	22.3	74.0	45.9	12.7	24.1	−0.9	0.2
	VPD	0.4	5.1	2.1	1.3	53.4	−1.1	0.3
	R	0.0	160.0	24.9	27.1	92.7	2.5	1.3
	ET_{Ref}	10.9	196.2	90.5	56.7	53.8	−1.2	0.2

ET_{Ref} and RH. Accordingly, in all 33 stations studied, there was the highest correlation coefficient of ET_{Ref} with average, maximum, and minimum temperature and VPD. Also, the correlation coefficient between ET_{Ref} and wind speed and relative humidity had the lowest values in most of the studied stations. ET_{Ref} in C and D climates (Fig. 2 c and d) compared to Bw and Bs climates (Fig. 2 a and b) show more correlation with solar radiation (R_s) and extra-terrestrial radiation (R_a) variable. On the other hand, based on the results of Fig. 2 a and d, there were high and low changes in the correlation coefficient values between ET_{Ref} and meteorological variables in stations with Bw and D climates, respectively.

2.2 ET_{Ref} estimation models

2.2.1 Empirical equations

In the present paper, the FAO₅₆ Penman-Monteith (ET_{Ref} - PMF56) method was used as the target ET_{Ref} values to compare with the results of empirical equations and MLR models, which can be calculated as:
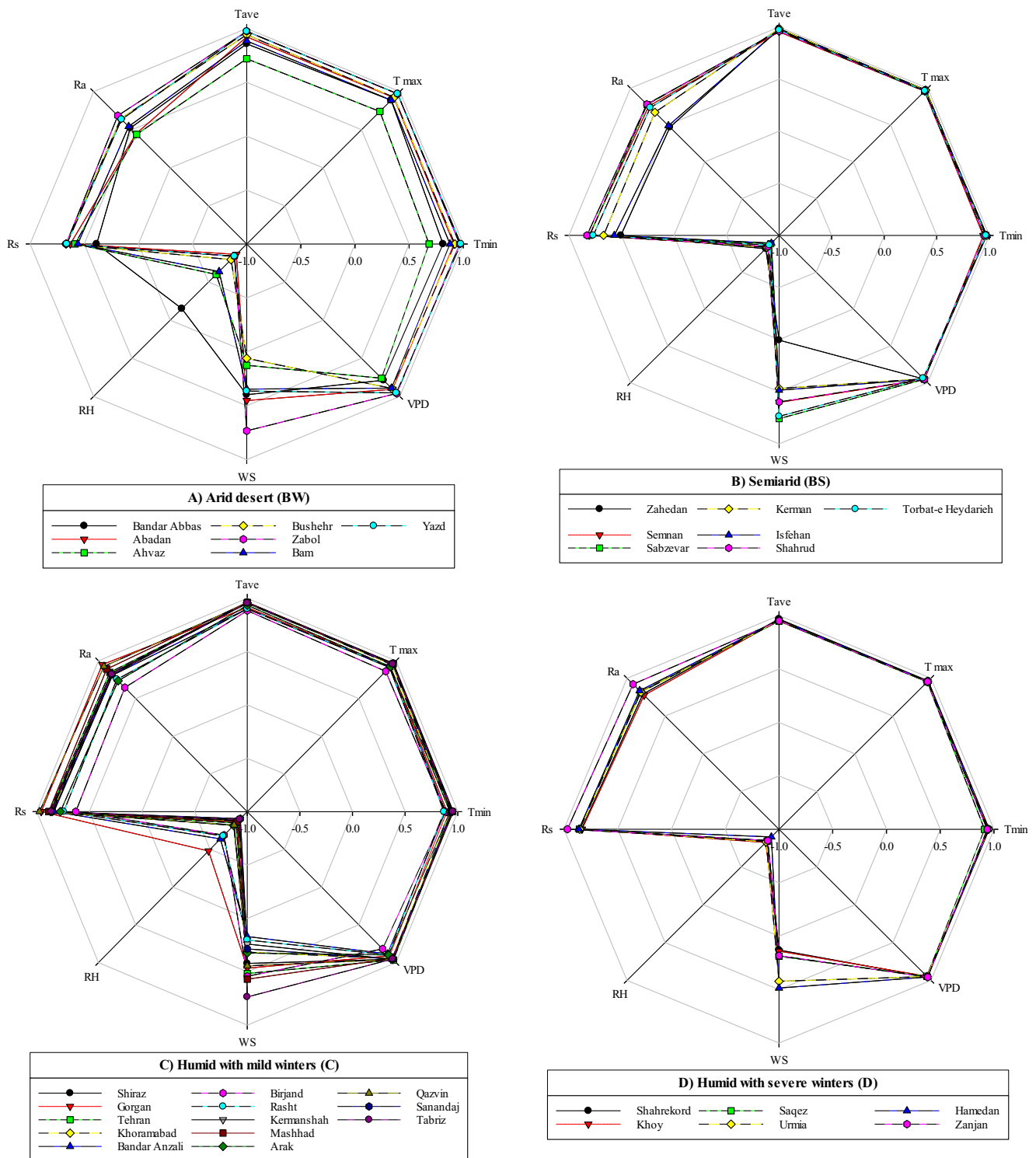
Fig. 2 Radar diagram of the Pearson correlation coefficient between ET_{Ref} and meteorological variable based on the Köppen method

$$ET_{Ref} = \frac{0.408\Delta(R_n-G) + \gamma\dfrac{900}{T_a + 273}u_2(e_s-e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

For more information on this method, you can refer to FAO_{56} (Allen et al. 1998). The developed empirical equations can be categorized into the temperature-based, radiation-based, and mass transfer-based according to the data used (Liu et al. 2017; Zhang et al. 2018). In this research, the empirical equations of Hargreaves-Samani (temperature-based),

Table 2 Characteristics of multivariate linear regression models for ET_{Ref} estimation

Model type	Inputs variables	Models	Train (80%)	Validation (20%)
Temperature-based	Tmin, Tmax	MLR1	1990–2013 (288)	2014–2019 (72 data)
	Tmin, Tmax, Ra	MLR2&HS		
Radiation-based	Tmin, Tmax, Ra, Rs	MLR3&Ir	1990–2013 (288)	2014–2019 (72 data)
Mass transfer-based	Tmin, Tmax, RH, u	MLR4	1990–2013 (288)	2014–2019 (72 data)
	u, VPD	MLR5&DT		
Combination-based	Tmin, Tmax, Ra, Rs, RH, u, VPD	MLR6	1990–2013 (288)	2014–2019 (72 data)

Irmak (radiation-based), and Dalton (mass transfer-based) are used based on Eqs. (2) (Hargreaves-Samani), (3) (Irmak), and (4) (Dalton), respectively, to estimate ET_{Ref} (Hargreaves and Samani 1985; Irmak et al. 2003; Dalton 1802).

$$ET_{Ref} = aR_a(T_a + b)\Delta T^{0.5} \tag{2}$$

$$ET_{Ref} = aR_s(bT_a + c) \tag{3}$$

$$ET_{Ref} = (a + bu_2)VPD \tag{4}$$

In these equations, ET_{Ref}, reference evapotranspiration (mm month^{-1}); Δ, the slope of the saturation vapor pressure function (kPa °C); γ, psychometric constant (kPa °C); R_n, net radiation (MJ m^{-2} day^{-1}); G, soil heat flux density (MJ m^{-2} day^{-1}); u, average wind speed at 2 m height (m s^{-1}); e_s, saturation vapor pressure (kPa); e_a, actual vapor pressure; VPD, vapor pressure deficit; α, 1.26; λ, latent heat of the evaporation (MJ kg^{-1}); R_a, extraterrestrial radiation (mm month^{-1}); R_s, monthly solar radiation (MJ m^{-2} month^{-1}); RH, relative humidity (%); T_a, average air temperature (°C); T_{max}, maximum air temperature (°C); T_{min}, minimum air temperature (°C); the values of a, b, and c are empirical coefficients.

Table 3 Equations derived for multivariate linear regression models in different climates

Model	*Köppen Climate	Equations	Eq.
MLR1	Bw	$ET_{Ref} = -18.82 + 0.09T_{min} + 4.33T_{max}$	(11)
	Bs	$ET_{Ref} = 14.44 + 4.04T_{min} + 3.10T_{max}$	(12)
	C	$ET_{Ref} = -4.92 + 0.99T_{min} + 4.91T_{max}$	(13)
	D	$ET_{Ref} = 12.63 + 1.36T_{min} + 4.80T_{max}$	(14)
MLR2	Bw	$ET_{Ref} = -12.65 + 0.92T_{min} + 4.44T_{max} - 1.20R_a$	(15)
	Bs	$ET_{Ref} = 14.44 + 4.21T_{min} + 2.95T_{max} + 0.35R_a$	(16)
	C	$ET_{Ref} = -15.40 + 1.27T_{min} + 3.85T_{max} + 4.15R_a$	(17)
	D	$ET_{Ref} = -3.03 + 1.76T_{min} + 3.72T_{max} + 4.49R_a$	(18)
MLR3	Bw	$ET_{Ref} = 28.51 + 3.95T_{min} + 1.45T_{max} - 11.82R_a + 17.61R_S$	(19)
	Bs	$ET_{Ref} = 45.70 + 6.75T_{min} + 0.37T_{max} - 8.73R_a + 15.59R_S$	(20)
	C	$ET_{Ref} = 40.33 + 6.11T_{min} - 1.25T_{max} - 12.40R_a + 29.37R_S$	(21)
	D	$ET_{Ref} = 68.78 + 6.87T_{min} - 1.26T_{max} - 16.59R_a + 32.94R_S$	(22)
MLR4	Bw	$ET_{Ref} = -30.63 + 0.49T_{min} + 5.16T_{max} + 0.16RH - 0.49u$	(23)
	Bs	$ET_{Ref} = -12.82 + 3.60T_{min} + 3.40T_{max} + 0.19RH - 3.66u$	(24)
	C	$ET_{Ref} = -26.98 + 0.56T_{min} + 5.41T_{max} + 0.16RH + 3.99u$	(25)
	D	$ET_{Ref} = -8.60 + 1.02T_{min} + 5.12T_{max} + 0.14RH - 2.55u$	(26)
MLR5	Bw	$ET_{Ref} = 66.47 - 0.04u + 23.45VPD$	(27)
	Bs	$ET_{Ref} = -18.93 + 6.55u + 38.38VPD$	(28)
	C	$ET_{Ref} = -4.15 + 4.84u + 37.80VPD$	(29)
	D	$ET_{Ref} = -10.42 + 5.62u + 39.69VPD$	(30)
MLR6	Bw	$ET_{Ref} = 20.45 + 3.29T_{min} - 0.57T_{max} + 0.18RH - 0.12u + 20.08VPD$	(31)
	Bs	$ET_{Ref} = 4.48T_{min} + 3.81u + 19.35VPD$	(32)
	C	$ET_{Ref} = -6.11 + 4.00T_{min} + 3.18R_S + 0.17RH3.14u + 21.50VPD$	(33)
	D	$ET_{Ref} = 29.54 + 5.83T_{min} - 3.38T_{max} + 3.34u + 27.68VPD$	(34)

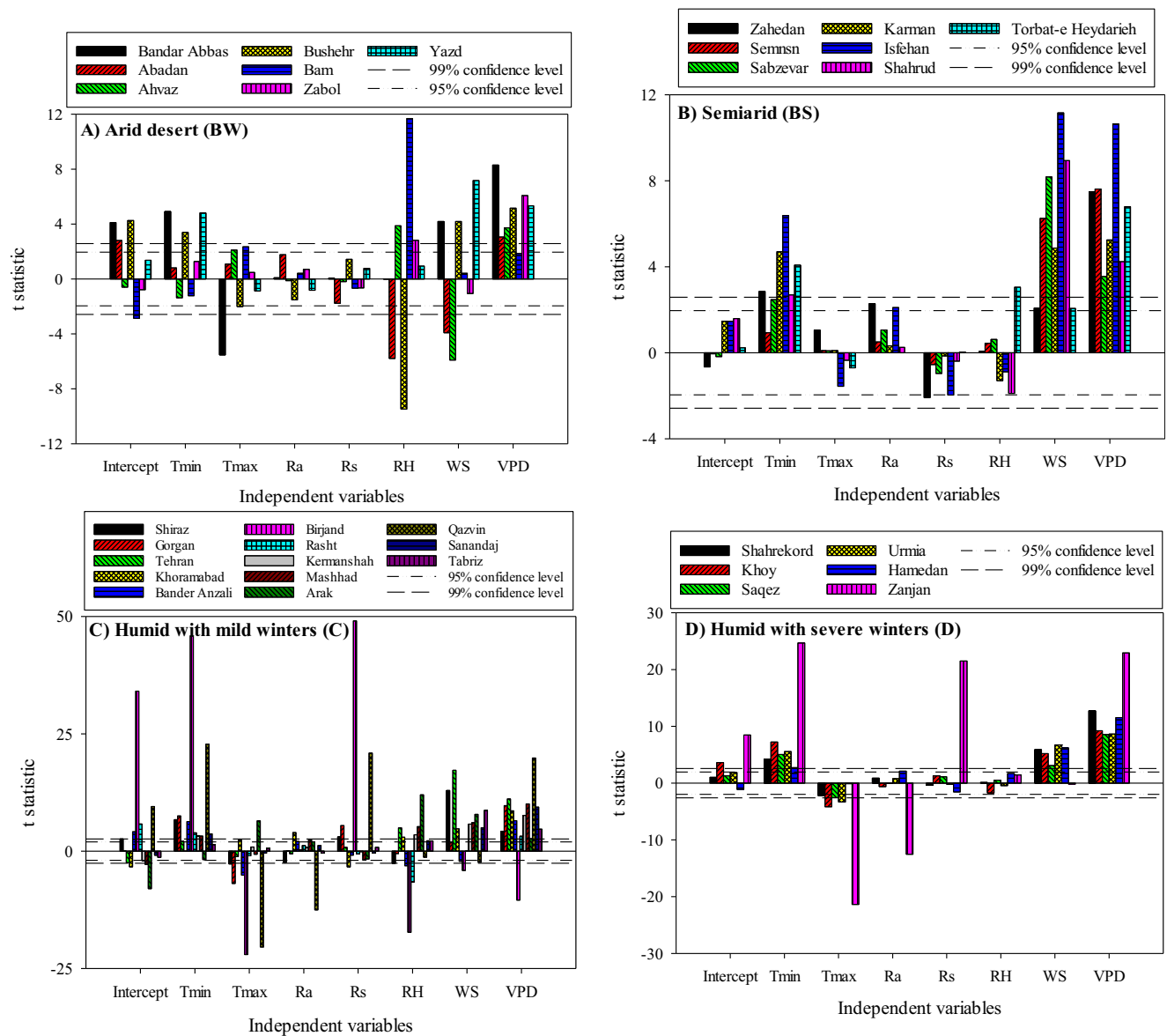* Bw, arid desert; Bs, semiarid; C, humid with mild winters; D, humid with severe winters

Fig. 3 *t* statistic of independent variables in different climates

Table 4 RMSE and R^2 criteria for different models in different climates

*Köppen climate	Criteria	HS	Irmak	Dalton	MLR1	MLR2	MLR3	MLR4	MLR5	MLR6
Bw	RMSE	28.3	19.1	23.2	17.9	17.6	17.4	15.9	21.8	15.1
	R^2	0.66	0.81	0.80	0.83	0.84	0.84	0.87	0.76	0.88
Bs	RMSE	27.7	13.0	15.1	12.6	12.6	12.5	11.7	14.3	10.8
	R^2	0.68	0.82	0.81	0.83	0.83	0.83	0.84	0.81	0.84
C	RMSE	26.1	22.6	19.9	18.1	14.8	14.2	17.1	18.4	12.3
	R^2	0.82	0.86	0.87	0.89	0.91	0.93	0.90	0.88	0.95
D	RMSE	22.0	18.6	19.6	19.0	17.1	15.9	18.6	19.0	13.3
	R^2	0.88	0.92	0.92	0.92	0.93	0.94	0.92	0.92	0.96

* *Bw*, arid desert; *Bs*, semiarid; *C*, humid with mild winters; *D*, humid with severe winters
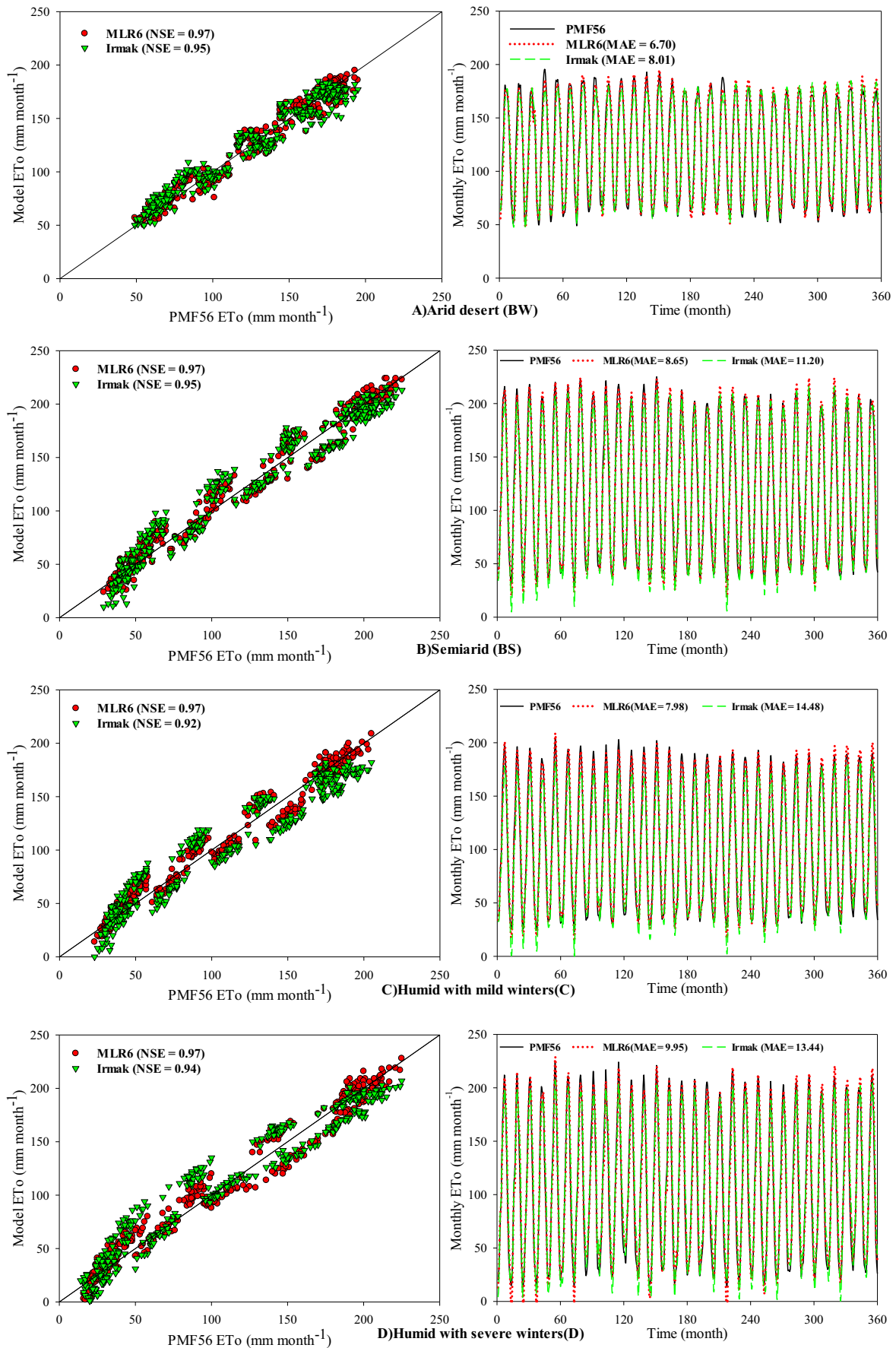
Fig. 4 Comparison of estimated ET_{Ref} values between Irmak and MLR6 models

2.2.2 Multivariate linear regression models

Multivariate linear regression is a method to model the relationship between several independent variables (average, maximum and minimum temperature, vapor pressure deficient, wind speed, relative humidity, solar radiation, and extraterrestrial radiation) with a dependent variable (ET$_{Ref}$). This method, based on the minimum of mean square error, performs the empirical coefficients of the linear relationship between the dependent variable and the independent variables in such a way that the linear model data has the best fit with the target data. In general, the form of multivariate linear regression equations is the following expression 5.

$$\widehat{Y} = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \ldots + b_m X_m \qquad (5)$$

In this expression, $\widehat{Y}$ is a dependent variable, b_0 to b_m are empirical coefficients and X_1 to X_m are independent variables. Also, the accuracy of the results of multivariate linear regression models extremely depends on the number and type of input variables to the model. In the
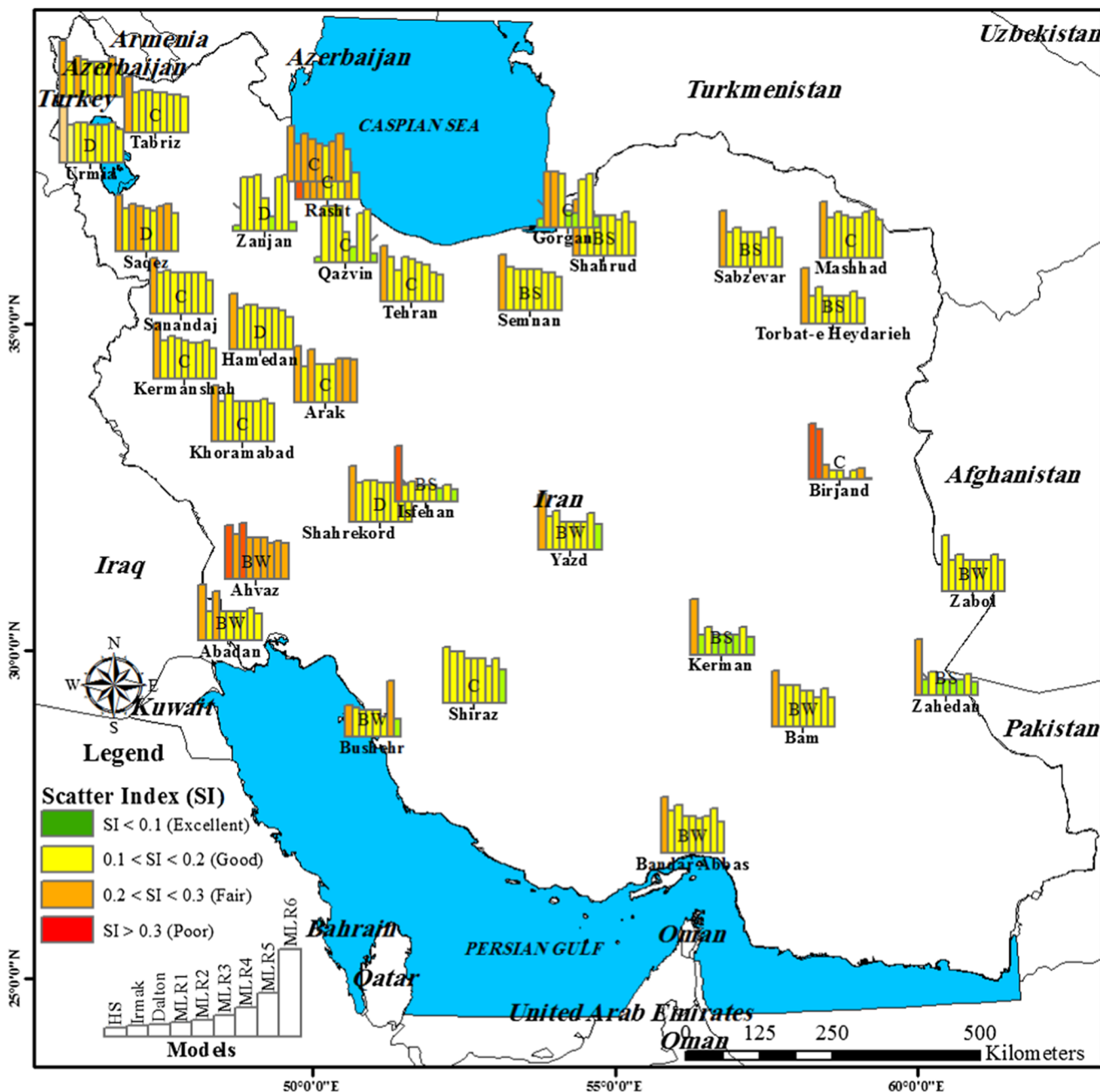

Fig. 5 SI criteria values using different empirical equations and MLR models in the studied stations

present study, 6 multivariate linear regression models presented in Table 2 have been developed. Meteorological variables are referred separately for each empirical equations and multivariate linear regression models. On the other hand, MLR1, MLR2, and HS models are defined as temperature-based, MLR3, and IR as radiation-based, MLR4, MLR5, and DT as mass transfer-based, and MLR6 as combination models, respectively (Table 2).

2.3 Evaluation criteria

In this study, 5 statistical criteria include the following: RMSE, SI, MAE, NSE, and R^2 were used to compare the results of empirical equations and MLR with ET_{Ref}-PMF56 based on Eqs. (6) to (10).

Root Mean Square Error (RMSE) RMSE

$$= \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(ET_{Ref_i}^{model} - ET_{Ref_i}^{PMF56} \right)^2} \tag{6}$$

Scatter Index (SI) $\text{SI} = \dfrac{RMSE}{\overline{ET}_{Ref}^{PMF56}}$ (7)

Mean Absolute Error (MAE)   MAE

$$= \frac{1}{N} \sum_{N}^{i=1} \left| ET_{Ref_i}^{model} - ET_{Ref_i}^{PMF56} \right| \tag{8}$$

Nash−Sutcliffe Efficiency (NSE)   NSE

$$= 1 - \left[\frac{\sum_{N}^{i=1} \left(ET_{Ref_i}^{PMF56} - ET_{Ref_i}^{model} \right)^2}{\sum_{i=1}^{N} \left(ET_{Ref_i}^{PMF56} - \overline{ET}_{Ref}^{PMF56} \right)^2} \right] \tag{9}$$

Coefficient of determination (R2) R^2

$$= \left[\frac{\sum_{N}^{i=1} \left(ET_{Ref_i}^{PMF56} - \overline{ET}_{Ref}^{PMF56} \right) \left(ET_{Ref_i}^{model} - \overline{ET}_{Ref}^{model} \right)}{\sqrt{\left[\sum_{i=1}^{N} \left(ET_{Ref_i}^{PMF56} - \overline{ET}_{Ref}^{PMF56} \right)^2 \right] \left[\sum_{i=1}^{N} \left(ET_{Ref_i}^{model} - \overline{ET}_{Ref}^{model} \right)^2 \right]}} \right]^2 \tag{10}$$

In Eqs. (6) to (10), $ET_{Ref_i}^{PMF56}$ and $ET_{Ref_i}^{model}$ are the ET_{Ref} based on PMF56 and modeled ET_{Ref}, $\overline{ET}_{Ref}^{PMF56}$ and $\overline{ET}_{Ref}^{model}$ are the mean values of ET_{Ref} based on PMF56 and modeled ET_{Ref}, and N is the number of data set (360 months).

The perfect value for MAE, RMSE, SI, and MAE indices, except NSE and R^2, is zero, and for NSE and R^2 is unity. According to Li et al. (2013), the range of SI for the accuracy of the models as:

$$IF \left\{ \begin{array}{l} SI < 0.1 \, \text{Excellent} \\ 0.1 < SI < 0.2 \, \text{Good} \\ 0.2 < SI < 0.3 \, \text{Fair} \\ SI > 0.3 \, \text{Poor} \end{array} \right\}$$

3 Results and discussion

In this paper, a mathematical equation was developed for each of the 6 predefined multivariate linear regression models (MLR1–6) to more accurately estimate ET_{Ref} for each of the four climates of Iran. The developed equations illustrated in Table 3 are based on various independent variables (meteorological parameters) and dependent variables (ET_{Ref}). Thus, MLR1 and MLR5 models were developed with two variables, MLR2 model with three variables, MLR3 and MLR4 models with four variables, and MLR6 model with the combination variables, i.e., 2 to 7 variables (Table 3). The MLR6 model has been developed as the best selection from 7 independent variables to the estimation of minimum error between ET_{Ref}^{PMF56} and ET_{Ref}^{model}. For example, according to Eq. (32), ET_{Ref}^{model} values have a linear relationship with the three variables T_{min}, u, and VPD in Bw climate (Table 3).

Kiafar et al. (2017) conducted the comparison of gene expression programming (GEP) models and empirical models led to the development of new equations for each climatic region. Their research showed that the models developed for very dry (Bw) and humid (C and D) climates had more accurate results. Also, based on their results, models that used the more climatic variables had a more accurate estimate of ET_{Ref}. Other researchers have reported similar results in the same climates (Traore et al. 2010; Ozkan et al. 2011; Huo et al. 2012).

To evaluate the relationship between ET_{Ref} and meteorological variables, a t test with a 95% confidence level ($\alpha = 0.05$) was used. The significance of the multivariate regression linear relationship between ET_{Ref} and each independent variable will be significant at the 95% level if the t stat value is greater than 1.96 or less than -1.96 (Mattar and Alazba 2019). Figure 3 shows the t stat parameter values for all stations studied by climates classified according to the Köppen method (Fig. 3).

According to the results of Fig. 3, in Bw climate, the two variables R_a and R_s had no significant linear relationship with ET_{Ref} in none of the stations of this climate at a 95% confidence level. Also in this climate, changes in the T_{min} variable had a significant relationship with ET_{Ref} only in three stations of Bandar Abbas, Bushehr, and Yazd, and the relationship of this variable with ET_{Ref} was not significant in four stations of Abadan, Ahvaz, Bam, and Zabol at 95% confidence level. Also, in Bs climate (Fig. 3b), there was only a significant linear relationship between the three variables T_{min}, u, and

VPD at 95% confidence level with ET_{Ref} in most stations of this climate. For other variables including intercept width, T_{max}, R_a, R_s, and RH, there was a significant linear relationship with ET_{Ref} at a 95% level in two out of seven stations in this climate (Fig. 3b).

Figure 3c shows that there is a significant linear relationship between ET_{Ref} and most meteorological variables in station C climates. In other words, there is a significant linear relationship between ET_{Ref} and most meteorological parameters in this climate at a 95% level (Fig. 3c). There is no significant linear relationship between ET_{Ref} values and R_a, R_s, and RH variables in the most stations with D climate (Fig. 3d). Based on the results of Fig. 3, there is a significant linear relationship between ET_{Ref} and VPD, u, and T_{min} variables in 32, 28, and 25 stations of the 33 studied stations ($\alpha = 0.05\%$), respectively. Also, the least significant linear relationship between ET_{Ref} and R_a and R_s variables was observed in 8 and 10 stations of all studied stations, respectively (Fig. 3). According to the results of Shiri et al. (2013) and Khanmohammadi et al. (2018), models in which the RH variable was used showed better results in C and D climates, because the effect of RH on the ET_{Ref} is greater in these climates. Also, based on the results of Yassin et al. (2016), the Irmak model (radiation-based) was more accurate in Bw and Bs climates. RMSE values for the Irmak model in arid climates were 25.78% lower than humid climates.

Table 4 shows the average values of RMSE and R^2 criteria in stations based on climates classified according to the Köppen method for empirical equations and MLR models. Also, the highest and lowest accuracy of the models in all climates was related to the MLR6 model and the empirical Hargreaves-Samani (HS) equation, respectively. The Irmak method had the best results in all climates in comparison with the other empirical equations. After the MLR6 model, which was estimated based on the linear relationship between ET_{Ref} and all climatic variables, the MLR4 model (based on the minimum and maximum temperature, relative humidity, and wind speed) was more accurate in Bw and Bs climates. In C and D climates, the MLR3 model (based on minimum, maximum temperatures, and solar and extraterrestrial radiations) showed the best estimates of ET_{Ref} between radiation-based models. This can be justified due to the higher correlation of ET_{Ref} with radiation variables in C and D climates (Fig. 2 c and d).

Based on the results of this study, a significant linear relationship between ET_{Ref} and VPD, u, and T_{min} variables were reported in 32, 28, and 25 stations, but this linear relationship was only significant in three variables T_{min}, u, and VPD with ET_{Ref} in most semiarid climate stations (Bs). However, the linear relationship between ET_{Ref} and most meteorological variables in C climate was significant.

Furthermore, according to the results of Table 4, results were more accurate of the simplest regression model (i.e., MLR1) in estimating the ET_{Ref} compared to all empirical equations. Also, in terms of RMSE criteria, the lowest error values were reported for C climate and with changes in the range of 13.3–22.0 mm month^{-1}. On the other hand, the highest values of R^2 for this climate were around 0.88–0.96 (Table 4).

The results of Shiri et al. (2012) showed that the HS method has less accurate in estimating daily reference evapotranspiration in the north of Spain in comparison with the Priestley-Taylor empirical method, intelligent gene expression programming (GEP), and adaptive neuro-fuzzy inference system (ANFIS) methods. Their results corroborate the results of the present study due to the lower accuracy of the HS method in all four climates of Iran.

Figure 4 shows the mean of the estimated ET_{Ref} values for the best empirical method (Irmak) and the best regression model (MLR6) in different climates with two criteria NSE and MAE (Fig. 4). Accordingly, NSE values in all climates for the MLR6 model were 0.97, which indicates the high accuracy of this model in estimating ET_{Ref} at all stations in different climates of the country.

The results also showed that the highest and lowest accuracy of the models in all climates was related to the MLR6 model and the empirical Hargreaves-Samani (HS) equation, respectively. Of the three empirical equations, the Irmak method showed the best results in all climates. After the MLR6 regression model, which was estimated based on the linear relationship between ET_{Ref} and all climatic variables, in Bw and Bs climates, the MLR4 model in climates C and D, the MLR3 model showed the best estimates of ET_{Ref}. Another noteworthy point was the high accuracy of the simplest regression model (i.e., MLR1) in estimating the ET_{Ref} compared to all empirical equations.

Also, Fig. 4 shows that in terms of NSE and MAE criteria, the Irmak empirical method is more accurate in arid and semiarid climates (Bw and Bs) than in humid climates (C and D).

According to the value of the SI criteria, the performance of a model can be divided into four levels: excellent to poor. Figure 5 shows the SI values at different stations for all empirical equations and the MLR models. Also, the MLR6 model had excellent estimates in 10 stations including Bushehr and Yazd stations (Bw); Zahedan, Kerman, and Isfahan (Bs); Shiraz, Gorgan, Birjand, and Qazvin (C); and Zanjan (D) (SI < 0.1) which in this regard had the highest accuracy in different climates. Then, MLR3 model in two stations in Bs climate (Zahedan and Kerman), in three stations in C climate (Gorgan, Birjand, and Qazvin), and one station in D climate (Zanjan) had the best estimate of ET_{Ref}. In other words, the HS equation in four stations of Ahwaz, Isfahan, Birjand, and Rasht had the lowest SI values and ET_{Ref} estimation using this method in these stations was poor (SI > 0.3). In general, Fig. 5 shows that

except for the HS equation, the accuracy of estimation of other methods in most of the studied stations concerning SI values was good (0.1 < SI < 0.2).

Figure 5 shows that the MLR6 model has the best estimation of ET_{Ref} in almost of the studied station located on different climates. However, in a few stations such as Arak, a model other than MLR6 has had better results in ET_{Ref} estimation based on the scatter index. According to Fig. 5, the Irmak method has been more accurate to estimate ET_{Ref} among the three studied empirical equations in most stations located in different climates. According to Unes et al. (2020) results, radiation-based empirical equations (Turc and Irmak) were found better than other empirical equations, the same with our results about Irmak empirical equation (radiation-based).

4 Conclusion

Based on the results of the study, the MLR6 model was developed as the best multivariate linear regression model with a minimum error value between ET_{Ref} -PMF56, and ET_{Ref}. The results of the model showed a significant linear relationship between ET_{Ref} as a dependent variable and other meteorological variables as independent variables, because of the MLR6 model. Also, the model had excellent estimates in 10 stations including Bushehr and Yazd stations (Bw), Zahedan, Kerman, and Isfahan (Bs); Shiraz, Gorgan, Birjand, and Qazvin (C); and Zanjan (D) (SI < 0.1), which in this regard had the highest accuracy of estimation in different climates. Therefore, the development of multivariate linear regression models provided the necessary preconditions for evaluating the function of the empirical equations of ET_{Ref} in the study.

Generally, the MLR6 shows efficient results under various climatic conditions. However, intensive data is required for this method and in developing countries; numerous climatic data are not available readily for all the weather stations. Also, the data always lack reliable quality. Certainly, there is a need to develop some approaches that can estimate ET_{Ref} precisely with available limited climatic data. The study demonstrated that all MLR models (even MLR1 with the T_{min} and T_{max} as input data) gave reliable results to estimation of ET_{Ref} in different climates.

Appendix

Table 5 The values of *t* stat and *p* value of independent variables in multivariate linear regression analysis for synoptic stations of Iran

*Köppen Climate	Station	Sta	Intercept	T_{Min}	T_{Max}	R_a	R_s	RH	u	VPD
Bw	Bandar Abbas	t	4.11	4.91	− 5.53	0.09	0.06	0.01	4.19	8.30
		p	< 1E-4	< 1E-4	< 1E-4	0.93	0.96	0.99	< 1E-4	< 1E-4
	Abadan	t	2.81	0.81	1.09	1.77	− 1.77	− 5.79	− 3.92	3.06
		p	0.01	0.42	0.28	0.08	0.08	< 1E-4	0.00	0.00
	Ahvaz	t	− 0.59	− 1.37	2.11	− 0.12	− 0.19	3.88	− 5.91	3.73
		p	0.55	0.17	0.04	0.90	0.85	0.00	< 1E-4	0.00
	Bushehr	t	4.26	3.39	− 2.01	− 1.52	1.44	− 9.47	4.19	5.16
		p	< 1E-4	0.00	0.05	0.13	0.15	< 1E-4	< 1E-4	< 1E-4
	Bam	t	− 2.87	− 1.22	2.35	0.42	− 0.67	11.69	0.42	1.86
		p	0.00	0.22	0.02	0.67	0.50	< 1E-4	0.68	0.06
	Zabol	t	−0.78	1.28	0.48	0.70	− 0.65	2.82	−1.05	6.08
		p	0.44	0.20	0.63	0.48	0.52	0.01	0.29	< 1E-4
	Yazd	t	1.37	4.81	− 0.87	− 0.82	0.78	0.94	7.17	5.33
		p	0.17	< 1E-4	0.38	0.41	0.43	0.35	< 1E-4	< 1E-4
Bs	Zahedan	t	− 0.66	2.85	1.05	2.28	− 2.09	0.06	2.08	7.49
		p	0.51	0.00	0.29	0.02	0.04	0.95	0.04	< 1E-4
	Semnan	t	− 0.05	0.93	0.10	0.50	− 0.55	0.44	6.25	7.62
		p	0.96	0.35	0.92	0.62	0.58	0.66	< 1E-4	< 1E-4
	Sabzevar	t	− 0.18	2.47	0.09	1.06	− 0.96	0.62	8.18	3.55
		p	0.86	0.01	0.93	0.29	0.34	0.54	< 1E-4	0.00

Table 5 (continued)

*Köppen Climate	Station	Sta	Intercept	T_{Min}	T_{Max}	R_a	R_s	RH	u	VPD
	Karman	t	1.47	4.70	0.10	0.32	− 0.15	− 1.30	4.87	5.25
		p	0.14	*< 1E-4*	0.92	0.75	0.88	0.19	*< 1E-4*	*< 1E-4*
	Isfahan	t	1.46	6.39	− 1.56	2.10	− 1.97	− 0.89	11.16	10.65
		p	0.14	*< 1E-4*	0.12	*0.04*	*0.05*	0.37	*< 1E-4*	*< 1E-4*
	Shahrud	t	1.59	2.69	− 0.35	0.25	− 0.39	− 1.89	8.94	4.24
		p	0.11	*0.01*	0.73	0.80	0.69	0.06	*< 1E-4*	*< 1E-4*
	Torbat-e Heydarieh	t	0.23	4.08	− 0.70	0.00	0.03	3.05	2.06	6.80
		p	0.82	*< 1E-4*	0.48	1.00	0.98	*0.00*	*0.04*	*< 1E-4*
C	Shiraz	t	2.61	6.71	− 2.72	− 2.39	3.01	− 2.61	12.94	4.21
		p	*0.01*	*< 1E-4*	*0.01*	*0.02*	*0.00*	*0.01*	*< 1E-4*	*< 1E-4*
	Gorgan	t	− 0.04	7.52	− 6.88	0.11	5.45	− 0.51	1.78	9.73
		p	0.97	*< 1E-4*	*< 1E-4*	0.91	*< 1E-4*	0.61	0.08	*< 1E-4*
	Tehran	t	− 2.49	2.12	− 1.17	− 0.57	0.81	4.95	17.25	11.13
		p	*0.01*	*0.04*	0.24	0.57	0.42	*< 1E-4*	*< 1E-4*	*< 1E-4*
	Khoramabad	t	− 3.34	− 0.04	2.38	3.97	− 3.37	2.98	4.75	8.62
		p	*0.00*	0.97	*0.02*	*< 1E-4*	*0.00*	*0.00*	*< 1E-4*	*< 1E-4*
	Bander Anzali	t	4.14	6.29	− 5.12	2.07	− 0.87	− 3.12	− 2.16	6.49
		p	*< 1E-4*	*< 1E-4*	*< 1E-4*	*0.04*	0.38	*0.00*	*0.03*	*< 1E-4*
	Birjand	t	34.12	45.80	− 22.04	0.27	137.06	− 17.30	− 4.12	− 10.44
		p	*< 1E-4*	*< 1E-4*	*< 1E-4*	0.79	*< 1E-4*	*< 1E-4*	*< 1E-4*	*< 1E-4*
	Rasht	t	5.78	3.90	− 0.99	1.19	− 0.61	− 6.62	0.01	3.18
		p	*< 1E-4*	*0.00*	0.32	0.24	0.54	*< 1E-4*	0.99	*0.00*
	Kermanshah	t	− 2.05	3.23	0.87	0.76	− 0.12	3.48	5.77	7.64
		p	*0.04*	*0.00*	0.38	0.45	0.91	*0.00*	*< 1E-4*	*< 1E-4*
	Mashhad	t	− 2.85	3.21	− 0.66	2.37	− 1.93	5.22	6.06	10.09
		p	*0.00*	*0.00*	0.51	*0.02*	0.05	*< 1E-4*	*< 1E-4*	*< 1E-4*
	Arak	t	− 7.99	− 1.80	6.45	1.98	− 1.67	12.02	7.86	2.65
		p	*< 1E-4*	0.07	*< 1E-4*	*0.05*	0.10	*< 1E-4*	*< 1E-4*	*0.01*
	Qazvin	t	9.54	22.79	− 20.42	− 12.52	20.90	− 1.31	− 2.36	19.85
		p	*< 1E-4*	*< 1E-4*	*< 1E-4*	*< 1E-4*	*< 1E-4*	0.19	*0.02*	*< 1E-4*
	Sanandaj	t	− 0.89	3.69	− 0.43	1.20	− 0.44	2.21	4.98	9.41
		p	0.37	*0.00*	0.67	0.23	0.66	*0.03*	*< 1E-4*	*< 1E-4*
	Tabriz	t	− 1.34	1.38	0.64	− 0.44	0.84	2.19	8.70	4.68
		p	0.18	0.17	0.52	0.66	0.40	*0.03*	*< 1E-4*	*< 1E-4*
D	Shahrekord	t	1.03	4.22	− 2.24	0.88	− 0.37	0.12	5.88	12.73
		p	0.30	*< 1E-4*	*0.03*	0.38	0.71	0.90	*< 1E-4*	*< 1E-4*
	Khoy	t	3.62	7.25	− 4.16	− 0.61	1.29	− 1.73	5.20	9.23
		p	*0.00*	*< 1E-4*	*< 1E-4*	0.54	0.20	0.08	*< 1E-4*	*< 1E-4*
	Saqez	t	1.30	5.05	− 2.55	− 0.06	1.12	0.53	3.09	8.56
		p	0.19	*< 1E-4*	*0.01*	0.95	0.26	0.60	*0.00*	*< 1E-4*
	Urmia	t	1.87	5.57	− 3.31	0.79	− 0.21	− 0.44	6.72	8.65
		p	0.06	*< 1E-4*	*0.00*	0.43	0.84	0.66	*< 1E-4*	*< 1E-4*
	Hamedan	t	−1.10	2.74	−0.09	2.12	−1.56	1.70	6.21	11.54
		p	0.27	*0.01*	0.93	*0.03*	0.12	0.09	*< 1E-4*	*< 1E-4*
	Zanjan	t	8.48	24.67	−21.37	−12.54	21.50	1.44	−0.16	22.92
		p	*< 1E-4*	*< 1E-4*	*< 1E-4*	*< 1E-4*	*< 1E-4*	0.15	0.87	*< 1E-4*

Italic values represent significant difference at $p < 0.05$

* Bw, arid desert; Bs, semiarid; C, humid with mild winters; D, humid with severe winters

References

Abdullah SS, Malek MA, Abdullah NS, Kisi O, Yap KS (2015) Extreme learning machines: a new approach for prediction of reference evapotranspiration. Journal of Hydrology 527:184–195

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration-guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56 Fao. Rome 300:D05109

Chen Z, Zhu Z, Jiang H, Sun S (2020) Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. Journal of Hydrology 591:125286

Dalton J (1802) Experimental essays on the constitution of mixed gases Manchester Literary and Philosophical Society Memo 5:535–602

dos Santos Farias DB, Althoff D, Rodrigues LN, Filgueiras R (2020) Performance evaluation of numerical and machine learning methods in estimating reference evapotranspiration in a Brazilian agricultural frontier Theoretical and Applied Climatology:1–12

Ehteram M et al (2019) An improved model based on the support vector machine and cuckoo algorithm for simulating reference evapotranspiration PloS one 14:e0217499

Gavili S, Sanikhani H, Kisi O, Mahmoudi MH (2018) Evaluation of several soft computing methods in monthly evapotranspiration modelling. Meteorological Applications 25:128–138

Güçlü YS, Subyani AM, Şen Z (2017) Regional fuzzy chain model for evapotranspiration estimation Journal of hydrology 544:233–241

Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. Applied engineering in agriculture 1:96–99

Huang G, Wu L, Ma X, Zhang W, Fan J, Yu X, Zeng W, Zhou H (2019) Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions. J Hydrol 574:1029–1041

Huo Z, Feng S, Kang S, Dai X (2012) Artificial neural network models for reference evapotranspiration in an arid area of northwest China Journal of arid environments 82:81–90

Irmak S, Irmak A, Allen R, Jones J (2003) Solar and net radiation-based equations to estimate reference evapotranspiration in humid climates Journal of irrigation and drainage engineering 129:336–347

Jing W et al. (2019) Implementation of evolutionary computing models for reference evapotranspiration modeling: short review, assessment and possible future research directions Engineering applications of computational fluid mechanics 13:811-823

Karbasi M (2018) Forecasting of multi-step ahead reference evapotranspiration using wavelet-Gaussian process regression model. Water Resour Manag 32:1035–1052

Khanmohammadi N, Rezaie H, Montaseri M, Behmanesh J (2018) The application of multiple linear regression method in reference evapotranspiration trend calculation Stochastic environmental research and risk assessment 32:661–673

Kiafar H, Babazadeh H, Marti P, Kisi O, Landeras G, Karimi S, Shiri J (2017) Evaluating the generalizability of GEP models for estimating reference evapotranspiration in distant humid and arid locations. Theoretical and Applied Climatology 130:377–389

Kim S, Singh VP, Seo Y, Kim HS (2014) Modeling nonlinear monthly evapotranspiration using soft computing and data reconstruction techniques Water resources management 28:185–206

Kisi O, Heddam S (2019) Evaporation modelling by heuristic regression approaches using only temperature data. Hydrological Sciences Journal 64:653–672

Kisi O, Sanikhani H, Zounemat-Kermani M, Niazi F (2015) Long-term monthly evapotranspiration modeling by several data-driven methods without climatic data. Computers and Electronics in Agriculture 115:66–77

Köppen W (1931) Grundriss der klimakunde. de Gruyter,

Li M-F, Tang X-P, Wu W, Liu H-B (2013) General models for estimating daily global solar radiation for different solar radiation zones in mainland China Energy conversion and management 70:139–148

Liu X, Xu C, Zhong X, Li Y, Yuan X, Cao J (2017) Comparison of 16 models for reference crop evapotranspiration against weighing lysimeter measurement. Agric Water Manag 184:145–155

Mattar MA, Alazba A (2019) GEP and MLR approaches for the prediction of reference evapotranspiration. Neural Computing and Applications 31:5843–5855

Mehdizadeh S, Behmanesh J, Khalili K (2017) Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. Computers and Electronics in Agriculture 139:103–114

Monteith J The state and movement of water in living organisms. In: 19th Symposia of the Society for Experimental Biology. Cambridge University Press, London, 1965, 1965. pp 205–234

Ozkan C, Kisi O, Akay B (2011) Neural networks with artificial bee colony algorithm for modeling daily reference evapotranspiration. Irrigation Science 29:431–441

Raza A, Shoaib M, Khan A, Baig F, Faiz MA, Khan MM (2020) Application of non-conventional soft computing approaches for estimation of reference evapotranspiration in various climatic regions. Theoretical and Applied Climatology 139:1459–1477

Reis MM, da Silva AJ, Junior JZ, Santos LDT, Azevedo AM, Lopes ÉMG (2019) Empirical and learning machine approaches to estimating reference evapotranspiration based on temperature data Computers and electronics in agriculture 165:104937

Saggi MK, Jain S (2019) Reference evapotranspiration estimation and modeling of the Punjab Northern India using deep learning Computers and electronics in agriculture 156:387–398

Sanikhani H, Kisi O, Maroufpoor E, Yaseen ZM (2019) Temperature-based modeling of reference evapotranspiration using several artificial intelligence models: application of different modeling scenarios. Theoretical and Applied Climatology 135:449–462

Sharafi S, Karim NM (2020) Investigating trend changes of annual mean temperature and precipitation in Iran. Arabian Journal of Geosciences 13:1–11

Sharafi S, Ramroudi M, Nasiri M, Galavi M, Kamali GA (2016) Role of early warning systems for sustainable agriculture in Iran. Arabian Journal of Geosciences 9:734

Shiri J, Kişi Ö, Landeras G, López JJ, Nazemi AH, Stuyt LC (2012) Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain) Journal of Hydrology 414:302–316

Shiri J, Nazemi AH, Sadraddini AA, Landeras G, Kisi O, Fard AF, Marti P (2013) Global cross-station assessment of neuro-fuzzy models for estimating daily reference evapotranspiration Journal of hydrology 480:46–57

Shiri J, Marti P, Karimi S, Landeras G (2019) Data splitting strategies for improving data driven models for reference evapotranspiration

estimation among similar stations Computers and Electronics in Agriculture 162:70–81

Shiri J, Zounemat-Kermani M, Kisi O, Mohsenzadeh Karimi S (2020) Comprehensive assessment of 12 soft computing approaches for modelling reference evapotranspiration in humid locations. Meteorol Appl 27:e1841

Tabari H, Kisi O, Ezani A, Talaee PH (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, Talaee PH (2013) Applicability of support vector machines and adaptive neurofuzzy inference system for modeling potato crop evapotranspiration. Irrigation Science 31:575–588

Tikhamarine Y, Malik A, Kumar A, Souag-Gamane D, Kisi O (2019) Estimation of monthly reference evapotranspiration using novel hybrid machine learning approaches. Hydrological Sciences Journal 64:1824–1842

Traore S, Wang Y-M, Kerh T (2010) Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone Agricultural water management 97:707–714

Unes F, Kaya YZ, Mamak M (2020) Daily reference evapotranspiration prediction based on climatic conditions applying different data mining techniques and empirical equations THEORETICAL AND APPLIED CLIMATOLOGY 141:763–773

Yassin MA, Alazba AA, Mattar MA (2016) Comparison between gene expression programming and traditional models for estimating evapotranspiration under hyper arid conditions. Water Resources 43:412–427

Yirga SA (2019) Modelling reference evapotranspiration for Megecha catchment by multiple linear regression Modeling Earth Systems and Environment 5:471–477

Zhang D, Lin J, Peng Q, Wang D, Yang T, Sorooshian S, Liu X, Zhuang J (2018) Modeling and simulating of reservoir operation using the artificial neural network, support vector regression, deep learning algorithm. J Hydrol 565:720–736

Zhu B, Feng Y, Gong D, Jiang S, Zhao L, Cui N (2020) Hybrid particle swarm optimization with extreme learning machine for daily reference evapotranspiration prediction from limited climatic data. Computers and Electronics in Agriculture 173:105430