CrossMark

ORIGINAL ARTICLE

Estimation of reference evapotranspiration using multivariate fractional polynomial, Bayesian regression, and robust regression models in three arid environments

Mojtaba Khoshravesh[1] · Mohammad Ali Gholami Sefidkouhi[1] · Mohammad Valipour[2]

Abstract The proper evaluation of evapotranspiration is essential in food security investigation, farm management, pollution detection, irrigation scheduling, nutrient flows, carbon balance as well as hydrologic modeling, especially in arid environments. To achieve sustainable development and to ensure water supply, especially in arid environments, irrigation experts need tools to estimate reference evapotranspiration on a large scale. In this study, the monthly reference evapotranspiration was estimated by three different regression models including the multivariate fractional polynomial (MFP), robust regression, and Bayesian regression in Ardestan, Esfahan, and Kashan. The results were compared with Food and Agriculture Organization (FAO)-Penman–Monteith (FAO-PM) to select the best model. The results show that at a monthly scale, all models provided a closer agreement with the calculated values for FAO-PM ($R^2 > 0.95$ and RMSE < 12.07 mm month^{-1}). However, the MFP model gives better estimates than the other two models for estimating reference evapotranspiration at all stations.

Keywords Evapotranspiration · Iran · Multivariate fractional polynomial · Bayesian regression · Robust regression

✉ Mohammad Ali Gholami Sefidkouhi
mohammadvalipour@gmail.com

[1] Department of Water Engineering, Sari Agricultural Sciences and Natural Resources University, Sari, Iran

[2] Young Researchers and Elite Club, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran

Introduction

Evapotranspiration (ET) is one of the major elements of the hydrologic cycle, and its accurate predictions is of paramount importance for many investigations such as irrigation system design and management, hydrologic water balance, crop yield simulation, irrigation scheduling, drainage studies, agricultural and forest meteorology, and water resources planning and management (Banihabib et al. 2012; Kumar et al. 2002; Valipour 2014c, d, e, f, g, h).

The most important parameters that affect on variations of evapotranspiration are air temperature, humidity, wind speed, and sun radiation (Pejic et al. 2015; Valipour 2014a, i, j, k; Wrachien and Mambretti 2015; Yannopoulos et al. 2015). At first, there is a distinction between reference and actual evapotranspiration. Reference evapotranspiration corresponds to the amount of water that would return to the atmosphere in case of moisture redundancy, for a reference crop [e.g., alfalfa (ET_r) or grass (ET_0)], from a standard surface and to then apply an appropriate empirical crop coefficient, which accounts for the difference between the standard surface and crop ET (Tzimopoulos et al. 2007; Valipour 2015a, b, c; Valipour et al. 2012).

ET_0 refers to the water emitted from a unit ground area covered with a reference plant, unstressed and healthy as well as with ample water supply (Walter et al. 2002). The ET_0 is used to quantify evapotranspirative demand within a zone and to forecast crop evapotranspiration when the ET_0 is multiplied by a crop coefficient (K_c) to determine for differences between the grass and plant evapotranspiration (Allen et al. 1998; Schuch and Burger 1997).

Measurements of pan evaporation are spatiotemporally limited, especially in developing countries. In locations where pan evaporation measurements are sparse, theoretical forecasts can be used to estimate it from other existing

مدينة الملك عبدالعزيز
KACST للعلوم والتقنية

🕿 Springer

data, which depend strongly on the local conditions. Evapotranspiration is a function of wind speed, temperature, and humidity. Solar radiation is also a significant parameter for evapotranspiration. All of these factors do not act independently. Wind varies the humidity, and if the humidity is low, the evapotranspiration rate is higher. At the same time, air temperature affects surface temperature and humidity. These parameters all contribute to evaporation with temperature likely being the dominant factor influencing evapotranspiration (Rahimikhoob 2009).

Where pan evaporation data are not available, evapotranspiration can be characterized by evapotranspiration models using weather variables. Among them, the Penman model is most frequently used for evapotranspiration prediction. The Penman model requires five climatic parameters: temperature, relative humidity, wind, saturation vapor pressure, and net radiation. It also uses complicated unit conversions and lengthy computations (Wu 1997).

Many researchers have evaluated the reliability of the Penman–Monteith (PM) model for predicting ET_0 (McNaughton and Jarvis 1984; Allen 1986; Allen et al. 1989; De Souza and Yoder 1994; Chiew et al. 1995). Jensen et al. (1990) analyzed the performance of 20 different models against lysimeter measured ET for 11 stations located in various climatic regions around the world. The PM model ranked as the best method for all climatic situations. However, the ranking of other models varied, depending on their local calibration and conditions.

Abtew and Obeysekera (1995) and Abtew (1996) found that the Penman–Monteith model was well suited to predict evapotranspiration from cattails (Typha domingensis), mixed marsh vegetation, and an open water/algae system; however, that calibrated simpler radiation-based methods also provided reasonable forecasts.

Potential evapotranspiration (PET), rather than actual evapotranspiration (AET), is a common input for empirical methods because it offers an upper limit to evapotranspirative water losses. PET is a function of vapor pressure gradient, available energy, and vegetation type (Douglas et al. 2009). In predicting PET, a clear definition of the "best" method for calculation is not evident and the model choice is often subjective. Verstraeten et al. (2008) presented a comprehensive review of the scientific articles on models for evaluating PET and mentioned that the selection of one model from the many is primarily dependent on the goals of the investigation and the type of available data. For example, Weib and Menzel (2008) compared the Priestley–Taylor (PT) method, two methods based on the Penman–Monteith (PM) equation and the Hargreaves method, a temperature-based method for predicting PET in a global-scale hydrologic method.

However, more study is needed on irrigation management to establish the appropriate model to be applied for evaluating crop evapotranspiration (ET), to avoid the excess or deficit water application, waterlogging, and soil salinity (Eslamian et al. 2009). Operational software tools in the domains specified above require the evaluation of ET_0 and the procedures to do that have been repeatedly implemented in software applications. This is one of the reasons that why there is an increasing demand for modular methods in model development (Jones et al. 2001). Reference evapotranspiration can be measured by lysimeters; however, this approach is very expensive and not easy to use (Valipour 2014l, m, 2015e). Many authors have attempted to estimate the evaporation through the indirect models using the weather parameters; however, some of these methods require the data, which cannot easily be measured (Rosenberry et al. 2007).

Today, applications of modeling and regression methods have been widely spread in different fields of the science. These methods were applied for increasing the quality of forecasts using the sample data sets from experiments and different investigations. Appropriate use of these methods is very useful for resolve of uncertainties and results of predictions, when obtain of information is difficult and sometimes impossible.

The most frequently used statistical methods are known as frequentist (or classical) methods. These methods assume that unknown parameters are the fixed constants, and they define probability using limiting relative frequencies. Bayesian methods offer an alternative approach; they treat the parameters as random variables and define probability as "degrees of belief" (i.e., the probability of an event is the degree to believe the event is true). The term "Bayesian" comes from the prevalent usage of Bayes' theorem, which was named after the Reverend Thomas Bayes, an eighteenth-century Presbyterian minister (Rathnayake 2010).

Bayesian analysis is a mode of inductive reasoning that has been applied in a lot of scientific disciplines. A distinctive feature of the Bayesian approach is that it permits the researcher to use sample (data) and prior (expert judgment) information in a logically consistent manner in making inferences. This is done by using Bayes' theorem to produce a "post-data" or posterior distribution for the model variables. Using Bayes' theorem, prior (or initial) values are transformed to post-data views. This transformation can be viewed as a learning process. The posterior distribution is calculated by the variances of the prior and sample information. If the variance of the prior information is smaller than the variance of the sampling information, then a higher weight is assigned to the prior information. On the other hand, if the variance of the sample information is smaller than the variance of the prior information, then a higher weight is assigned to the sample information causing the posterior forecast to be closer to the sample information (Chulani et al. 1998; Singh 2014).

Since their introduction by Royston and Altman (1994), the regression methods based on fractional polynomial (FP) transformation of continuous predictor(s) have found gradual acceptance as a useful technique of analysis which retains such predictors as continuous in the model. The robust regression using functions related to the least squares has been the subject of intense research (Royston and Sauerbrei 2007). Demoster et al. (1980) discussed statistical properties of the predictions under the assumption that the observation errors are independent normal. Coleman et al. (1980) developed a high-quality set of routines to calculate robust predictions for eight weight functions.

The superiority of regression models to artificial intelligence techniques such as ANFIS, ANN, and SVM is that regression models have the less complexity. In addition, these models give us the formula or relationship and have an acceptable accuracy. Although application of nonlinear regression and Bayesian approach has been surveyed in the various fields of hydrology and water management such as irrigation canal demands (Ticlavilca et al. 2013), low flow indices (Joshi et al. 2013), land surface temperature (Ghosh and Joshi 2014), catchment modeling (Marshall et al. 2007), nutrient loading (Vigiak and Bende-Mich 2013; Wellen et al. 2012), heat flux (Ershadi et al. 2013; Yao et al. 2014), solar radiation (Iizumi et al. 2012), and streamflow modeling (Block and Rajagopalan 2009; Liang et al. 2013), there are few studies about application of them in estimation of ET (Bachour et al. 2014; Izadifar and Elshorbagy 2010; Tabari et al. 2012; Zhu et al. 2013). However, these studies focus on some limited methods and/or use low recorded data. Furthermore, comparison of the nonlinear regression methods and Bayesian approach has not been investigated in arid regions (Valipour 2013a, b). In the present paper, the attempts are made to apply MFP, Bayesian regression, and robust regression models to estimate reference evapotranspiration values, in three arid regions of Iran in which water crises are considerable (Mahdizadeh Khasraghi et al. 2015; Valipour et al. 2015), using temperature data.

Materials and methods

Fractional polynomials

Simple power transformations of a covariate have long been used informally in data analysis in cases when nonlinearity was suspected. General power models were proposed by Box and Tidwell (1962). Royston and Altman (1994) formalized the simple power models and called them fractional polynomials of degree 1 (FP1), and extended them to FPs of higher degree. Ignoring the

intercept term (β_0), a polynomial of degree m in a single covariate X may be written:

$$\beta_1 X^1 + \beta_2 X^2 + \cdots + \beta_m X^m \tag{1}$$

In the generalization to an FP function of degree m, written FP $m(X)$, the indices 1,…, m are replaced with the powers p_1, …, p_m, giving:

$$\text{FP}\, m(X) = \beta_1 X^{P1} + \beta_2 X^{P2} + \cdots + \beta_m X^{Pm} \tag{2}$$

As proposed by Royston and Altman (1994), the powers p_1, …, p_m are chosen from a restricted set, $S = \{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$, where X^0 denotes log$X$. No subsequent changes to S have been adopted. The set includes no transformation ($p = 1$) and the reciprocal, logarithmic, square root, and square transformations. FP2 functions take the form:

$$\text{FP2} = \beta_1 X^{P1} + \beta_2 X^{P2} \tag{3}$$

Or

$$\text{FP2} = \beta_1 X^{P1} + \beta_2 X^{P1} \log X \tag{4}$$

The latter being the so-called repeated-powers function obtained in the mathematical limit $p_2 \to p_1$ (Royston and Sauerbrei 2005).

Consider a linear regression model:

$$Y = X\beta_2 + e \tag{5}$$

where Y is an n-vector of observed responses, X is an $n \times k$ matrix representing k covariates, β is a k-vector of parameters and e is an n-vector of residuals with $E(e) = 0$, var(e) $= \sigma^2 I$. The ordinary least squares (OLS) estimator of β is $\hat{\beta} = (X'X)^{-1}X'Y$, and the vector of fitted values is $\hat{Y} = X\hat{\beta} = HY$. The matrix $H = X(X'X)^{-1}X'$ is usually called the "hat" matrix. Its diagonal elements $h_i = H_{ij}x_i(X'X)^{-1}x_i'$ are termed leverages. In the special case of a single covariate, x, the leverages are as follows given (Belsley et al. 1980; Royston and Sauerbrei 2007):

$$h_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{j=1}^n (x_j - \bar{x})^2} \tag{6}$$

The leverage is then a linear function of the squared distance of x_i from the covariate mean, $\bar{x}$.

Multivariable fractional polynomial modeling

The central problem of modeling is the joint effect of several covariates, at least two of which are continuous and some may be categorical or binary. At the same time, the model must be simplified, both/either dropping nonsignificant variables and/or by reducing the complexity of FP functions fitted to continuous covariates. A solution to finding a model in this situation was proposed by Royston

and Altman ([1994](#)) and further refined by Sauerbrei and Royston ([1999](#)). In essence, it is impractical to try all combinations of powers for finding the best-fitting set of FP functions for all of the candidate continuous variables, and an iterative algorithm is needed. This was termed the MFP procedure by Sauerbrei and Royston ([1999](#)). It combines backward elimination of variables with a search for the best FP functions of continuous covariates. First, a significance level such as 0.05 is chosen. Next, the fitting order is determined by fitting a model linear in all variables and eliminating each variable singly, which is the first step of a conventional backward elimination procedure. The p values from the tests are used to order the variables, from the most to the least significant. Having imposed an order, each X is considered in turn. If an X is categorical or binary, a standard likelihood ratio test is performed to determine whether it should be dropped or not. If an X is continuous, the FP model selection procedure is applied to determine the best FP (or to eliminate X). The procedure cycles through the variables in the same order until the selected variables and FP functions do not change. Typically, two or three cycles are required for convergence (Royston and Sauerbrei [2005](#)).

Robust regression

The models described in linear regression models are based on certain assumptions, such as a normal distribution of errors in the observed responses. If the distribution of errors is asymmetric or prone to outliers, model assumptions are invalidated, and parameter estimates, confidence intervals, and other computed statistics become unreliable. The Statistics Toolbox function robust fit is useful in these cases. The function implements a robust fitting method that is less sensitive than OLS to large changes in small parts of the data (Matlab Statistics Toolbox [2010](#)).

Robust regression works by assigning a weight to each data point. Weighting is done automatically and iteratively using a process called *iteratively reweighted least squares*. In the first iteration, each point is assigned equal weight and model coefficients are estimated using ordinary least squares. At subsequent iterations, the weights are recomputed so that the points farther from model predictions in the previous iteration are given a lower weight. The model coefficients are then recomputed using weighted least squares (WLS). The process continues until the values of the coefficient estimates converge within a specified tolerance (Matlab Statistics Toolbox [2010](#)).

Robust demo demonstrates the difference between ordinary (least squares) regression and robust regression. It displays a scatter plot of X and Y values, where Y is roughly a linear unction of X, but one point is an outlier (it falls far from the line). The bottom of the figure shows the fitted

equations using least squares and robust fitting, plus an estimate of the root mean squared error from both. Leverage is a measure of how much influence that point has on the least squares fit. Weight is the weight that point was given in the robust fit. Robust demo (X, Y) uses X and Y data that supply, in place of the sample data supplied with the function (Matlab Statistics Toolbox [2010](#)).

Bayesian regression

Bayesian regression, the theory specialized adaptation is including the development of multivariate regression models, which clearly consider two sources of previous and experimental information.

Bayesian regression analysis is including development of model or prediction of the relations among variables. For example, a prediction equation can be the following linear form (Eq. [7](#)).

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_p x_p \tag{7}$$

where Y is the dependent variable, independent variables are predicted by X_i, and the aim is numerical determine of b_j coefficients.

The Bayesian model makes possible the decision in short time with development of data and judgments and continue of modeling.

Suppose that we are interested in estimating θ from data $y = \{y_1, \ldots, y_n\}$ using a statistical model described by a density $p(y|\theta)$. Bayesian philosophy states that θ cannot be determined exactly, and uncertainty about the parameter is expressed through probability statements and distributions. Also, we can say that θ follows a normal distribution with mean 0 and variance 1, if it is believed that this distribution best describes the uncertainty associated with the parameter. The following steps describe the essential elements of Bayesian inference:

1. A probability distribution for θ is formulated as $\pi(\theta)$, which is known as the *prior* distribution, or just the prior. The prior distribution expresses beliefs (e.g., on the mean, the spread, and the skewness) about the parameter before examining the data.
2. Given the observed data y, choose a statistical model $p(y|\theta)$ to describe the distribution of y given θ.
3. Update of beliefs about θ by combining information from the prior distribution and the data through the calculation of the *posterior* distribution, $p(y|\theta)$.

The third step is carried out using Bayes' theorem, which enables to combine the prior distribution and the model in the following way:

$$p(y|\theta) = \frac{p(\theta.y)}{p(y)} = \frac{p(y|\theta)\pi(\theta)}{p(y)} = \frac{p(y|\theta)\pi(\theta)}{\int p(y|\theta)\pi(\theta)\mathrm{d}\theta} \tag{8}$$

The quantity

$$p(y) = \int p(y|\theta)\pi(\theta)\mathrm{d}\theta \qquad (9)$$

is the normalizing constant of the posterior distribution. This quantity $p(y)$ is also the marginal distribution of y, and it is sometimes called the marginal distribution of the data. The likelihood function of θ is any function proportional to $p(y|\theta)$; that is $L(\theta) \propto p(y|\theta)$.

In this study, daily grass reference ET (ET_0) is determined using the FAO Penman–Monteith (FAO-PM) approach (Allen et al. 1998). There are many studies that applied FAO-PM model as a base model for calculation of reference ET in different regions of Iran (Rahimi et al. 2014; Valipour 2012c, d, e, f, 2015d; Valipour and Eslamian 2014). The FAO-PM equation for a grass reference crop is defined as:

$$\mathrm{ET}_0 = \frac{0.408\varDelta(R_n - G) + \gamma(900/(T + 273))u_2(e_s - e_a)}{\varDelta + \gamma(1 + 0.34u_2)} \qquad (10)$$

where ET_0 is the evapotranspiration (mm day^{-1}), R_n is the net radiation (MJ m^2 day^{-1}), G is the soil heat flux (MJ m^2 day^{-1}), T is the mean daily air temperature (°C), u_2 is the mean daily wind speed at 2 m height (m s^{-1}), $e_s - e_a$ is the saturation vapor pressure deficit (kPa °C^{-1}), and γ is the psychometric constant (kPa °C^{-1}).

In this study, three models were performed including MFP, Bayesian regression, and robust regression models. All of the models were designed with three parameters as input data, including solar radiation (Ra), mean temperature (T_{mean}), and $T_{max} - T_{min}$.

Study area

In this study, three synoptic stations have been used that have at least 14-year data to avoid the errors due to the problem of short record length. These stations are located on central part of Iran with an arid environment. Figure 1 and Table 1 show the position and geographic characteristics of Esfahan, Kashan, and Ardestan stations, respectively. All of the information is in monthly scale and has been collected from Islamic Republic of Iran Meteorological Organization (IRIMO) at this link: http://irimo.ir/eng/index.php.

To evaluate the performance of these models in monthly ET_0 estimates, between the calculated and predicted ET, the root mean square error (RMSE) is applied as given:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\left(\frac{Y_{obs} - Y_{calc}}{Y_{obs}}\right)^2\right)} \qquad (11)$$

where Y_{calc} is the calculated value and Y_{est} is estimation of the models.

Results and discussions

Multivariable fractional polynomial modeling

The results of statistical parameters of MFP model are given in Table 2. In addition, Figs. 2, 3 and 4 show the predicted values of ET versus the calculated values of ET.

To evaluate the MFP model, the amount of ET estimated by this model was plotted versus the calculated values. These results are shown for Esfahan, Kashan, and Ardestan stations in Figs. 2, 3 and 4, respectively. The best performance of the MFP model was obtained for Kashan than Esfahan and Ardestan stations.

Considering the obtained results, in this study, it could be stated that use of MFP model can predict ($R^2 > 0.95$) the amount of ET in different stations, and this application can be added to the other applications of MFP model.

The formula obtained from MFP model for Esfahan, Kashan, and Ardestan stations are shown in Eqs. 12, 13, and 14, respectively.

$$\mathrm{ET} = -323.91 + 54.6\left(\frac{\mathrm{Ra}}{1000}\right)^{-1} + 293.42\left(\frac{\mathrm{Ra}}{1000}\right)^{0.5}$$
$$+ 23.71\left(\frac{T_{mean} + 3.4}{10}\right) + 11.74\left(\frac{T_{max} - T_{min}}{10}\right)^3$$
$$- 16.62\left(\frac{T_{max} - T_{min}}{10}\right)^3 \mathrm{Ln}\left(\frac{T_{max} - T_{min}}{10}\right) \qquad (12)$$

$$\mathrm{ET} = -13.304 + 63.466\left(\frac{\mathrm{Ra}}{1000}\right)^2 + 4.172\left(\frac{T_{mean}}{10}\right)^2$$
$$+ 14.324\left(\frac{T_{max} - T_{min}}{10}\right) \qquad (13)$$

$$\mathrm{ET} = -39.147 + 62.267\left(\frac{\mathrm{Ra}}{1000}\right)^2 + 9.524\left(\frac{T_{mean}}{10}\right)^2$$
$$+ 56.422\left(\frac{T_{max} - T_{min}}{10}\right) \qquad (14)$$

Robust regression

The results of statistical parameters of robust regression model are given in Table 3. In addition, Figs. 5, 6 and 7 show the predicted values of ET versus the calculated values of ET.

To evaluate the robust regression model, the amount of ET estimated by this model was plotted versus the calculated values. These results are shown for Esfahan, Kashan, and Ardestan stations in Figs. 5, 6 and 7, respectively. The maximum value of the correlation coefficient was obtained for Ardestan; however, the minimum value of RMSE belongs to Kashan.

Fig. 1 Position of Esfahan, Kashan, and Ardestan in Iran

Table 1 Geographic characteristics of the stations

Station	Latitude	Longitude	Elevation m.a.s.l	Record duration	Annual precipitation (mm)	Climate (de Martonne)
Esfahan	32°37′N	51°40′E	1550.4	1961–2005	122.8	Arid
Kashan	33°59′N	51° 27′E	982.3	1966–2005	138.4	Arid
Ardestan	33°23′N	52° 23′E	1252.4	1992–2005	115.8	Arid

The formula obtained from robust regression model for Esfahan, Kashan, and Ardestan stations are shown in Eqs. 15, 16, and 17, respectively.

$$ET = -19.6047 + 2.8297T_{mean} - 0.7866(T_{max} - T_{min}) \\ + 0.0753Ra \tag{15}$$

$$ET = -51.9819 + 2.022T_{mean} - 0.0465(T_{max} - T_{min}) \\ + 0.1061Ra \tag{16}$$

$$ET = -87.4207 + 3.9657T_{mean} + 2.381(T_{max} - T_{min}) \\ + 0.1195Ra \tag{17}$$

Table 2 Statistical parameters of the MFP model for ET prediction

Stations	R^2	RMSE
Esfahan	0.96	9.48
Kashan	0.98	7.58
Ardestan	0.98	8.74
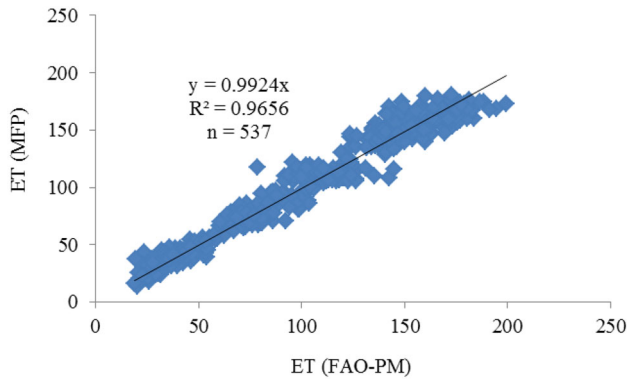
Fig. 2 Comparison of the ET predicted with MFP model versus the calculated values by FAO-PM for Esfahan station
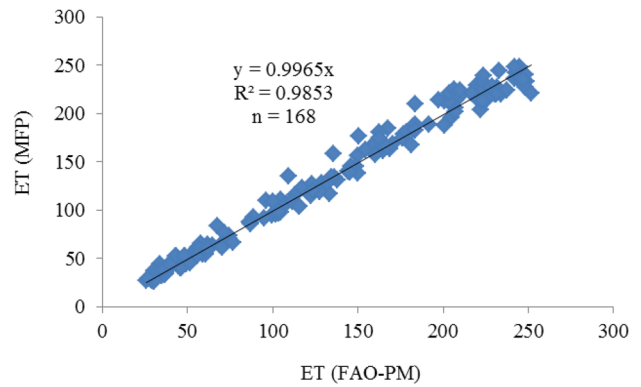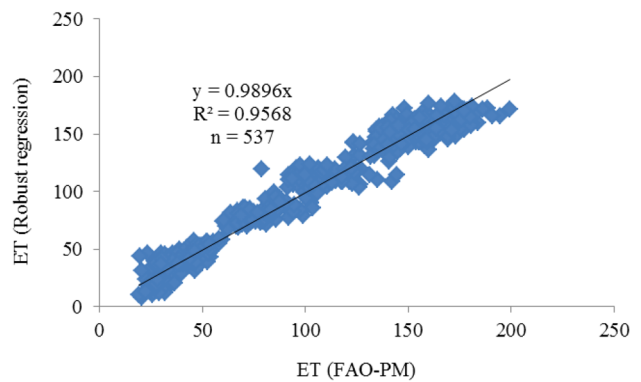

Fig. 3 Comparison of the ET predicted with MFP model versus the calculated values by FAO-PM for Kashan station

Fig. 4 Comparison of the ET predicted with MFP model versus the calculated values by FAO-PM for Ardestan station

Table 3 Statistical parameters of the robust regression model for ET prediction

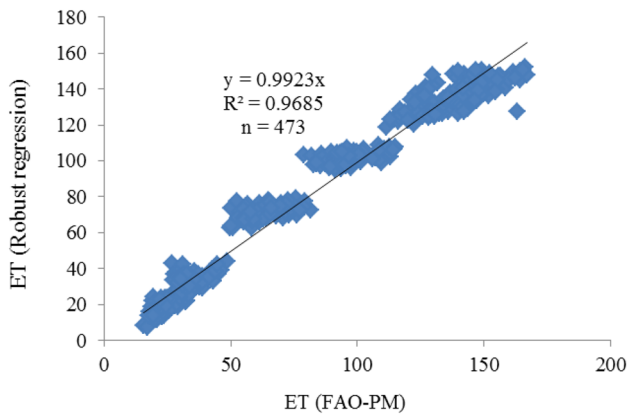| Stations	R^2	RMSE
Esfahan	0.95	10.59
Kashan	0.96	8.24
Ardestan	0.97	12.07

Fig. 5 Comparison of the ET predicted with robust regression model versus the calculated values by FAO-PM for Esfahan station

Bayesian regression

The results of statistical parameters of Bayesian regression model are given in Table 4. In addition, Figs. 8, 9, and 10 show the predicted values of ET versus the calculated values of ET.

To evaluate the Bayesian regression model, the amount of ET estimated by this model was plotted versus the calculated values. These results are shown for Esfahan, Kashan, and Ardestan stations in Figs. 8, 9, and 10, respectively. Similar to the robust regression analysis, the maximum value of the correlation coefficient was obtained

for Ardestan; however, the minimum value of RMSE belongs to Kashan. It should be noted that performance of the Bayesian model for Kashan (Fig. 9) is the only over-estimating observed in all models among all stations.
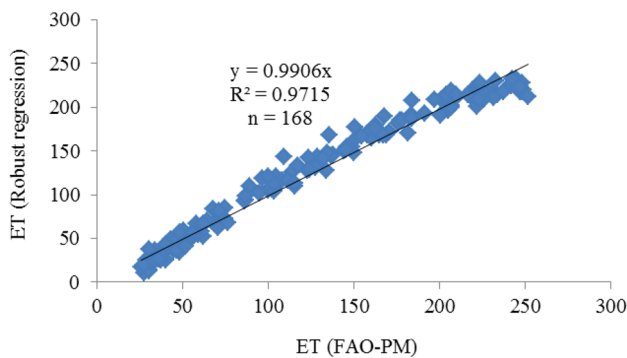
The formula obtained from Bayesian regression model for Esfahan, Kashan, and Ardestan stations are shown in Eqs. 18, 19, and 20, respectively.

$$ET = -20.2773 + 0.0769Ra - 0.7886(T_{max} - T_{min}) + 2.7749T_{mean} \tag{18}$$

$$ET = -61.0469 + 0.0961Ra + 0.9314(T_{max} - T_{min}) + 2.2568T_{mean} \tag{19}$$

Fig. 6 Comparison of the ET predicted with robust regression model versus the calculated values by FAO-PM for Kashan station

Fig. 7 Comparison of the ET predicted with robust regression model versus the calculated values by FAO-PM for Ardestan station

**Table 4** Statistical parameters of the Bayesian regression model for ET prediction

Station	R^2	RMSE
Esfahan	0.95	10.59
Kashan	0.96	8.84
Ardestan	0.97	12.05

$$ET = -87.9839 + 0.1166Ra + 2.4713(T_{max} - T_{min}) + 4.0873T_{mean} \tag{20}$$

According to Tables 2, 3, and 4, the best results obtained from the MFP model with the highest regression coefficient and the lowest RMSE. The MFP model is more flexible than the ordinary polynomial methods and more resistant than nonparametric regression methods.

Figures 11 and 12 are plotted to better comparison of the performance of the models used with respect to the average values of the evaluation indices.

According to Fig. 11, once again the MFP model is introduced as the best model for estimation of reference ET in the study areas due to the highest correlation coefficient
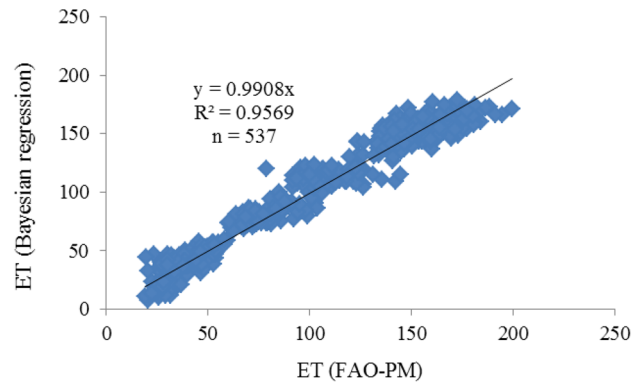
Fig. 8 Comparison of the ET predicted with Bayesian regression model versus the calculated values by FAO-PM for Esfahan station
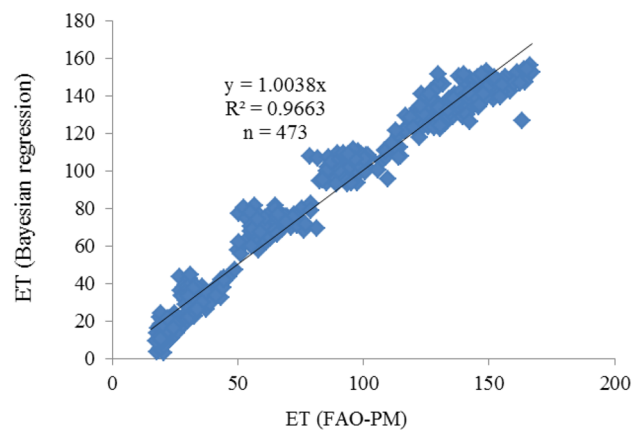
Fig. 9 Comparison of the ET predicted with Bayesian regression model versus the calculated values by FAO-PM for Kashan station
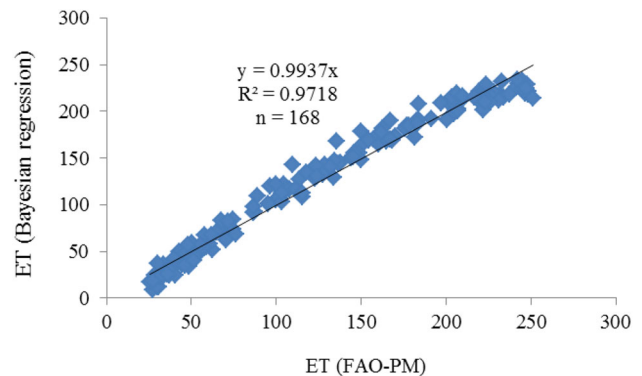

Fig. 10 Comparison of the ET predicted with Bayesian regression model versus the calculated values by FAO-PM for Ardestan station

and the lowest RMSE and slope error. It should be noted that the slope error was obtained based on the coefficient of the suggested equations in Figs. 2, 3, 4, 5, 6, 7, 8, 9, and 10.

According to Fig. 12, the models used estimate reference ET in Kashan better than Esfahan and Ardestan.

Fig. 11 Comparison of the models used according to the average of the evaluation indices
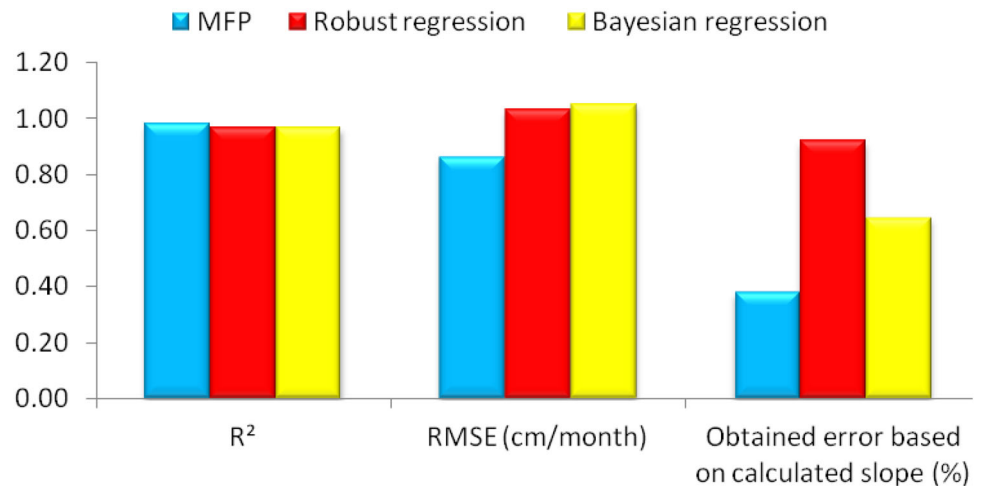
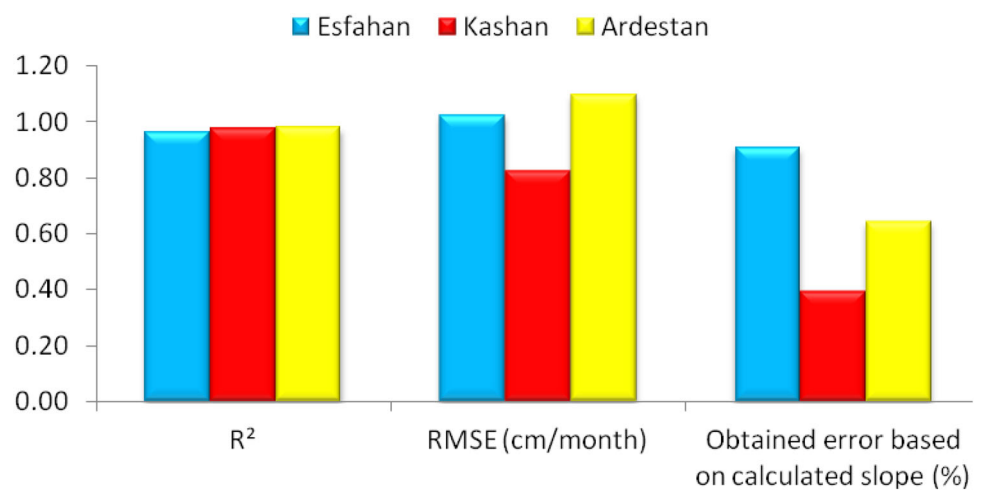Fig. 12 Comparison of the average of evaluation indices for area studies

Table 1 confirms this result because the annual precipitation of Kashan (138.4 mm) is more than both Esfahan (122.8 mm) and Ardestan (115.8 mm). In addition, compared with Esfahan (1550.4 masl) and Ardestan (1252.4 masl), the lowest elevation belongs to Kashan (982.3 masl). Therefore, application of the models is proposed for all regions in the world with weather conditions close to Kashan, and this leads to the best performance as well as reduction of uncertainty for estimation of reference ET.

Compared with the previous works (in these three regions), the obtained results underline better performance of MFP, robust regression, and Bayesian regression in Esfahan, Kashan, and Ardestan. Azizi et al. (2010) applied multivariations regression model for estimating reference ET in Esfahan province. The authors used four weather parameters (temperature, air vapor pressure, relative humidity, and wind speed) to estimate reference ET and claimed that use of this approach led to correlation coefficients equal to 0.998 and 0.992, in Esfahan and Kashan,

respectively. In the other study, Heydari et al. (2013) evaluated different methods include Penman FAO, Blaney–Criddle, Linacre, Thorenthwaite, Hargreaves–Samani, Irmak, Turc, and Jensen–Haise in Ardestan, and then their results were compared with FAO-PM method. The results indicated that the Blaney–Criddle model is the most appropriate for this particular study area, with an average RMSE of 1.34 mm day^{-1} and MAE of 1.00 mm day^{-1} and a correlation coefficient of 0.952. The obtained results can be expanded for other arid regions with similar wheatear conditions.

Summary and conclusions

In this research, the capacity of different models for evapotranspiration prediction was investigated. Evapotranspiration is a complex and nonlinear phenomenon because it depends on several interacting climatological factors, such as temperature, humidity, wind speed,

radiation, type, and growth stage of the crop (Jain et al. 2008; Ojha and Bhakar 2012; Valipour 2012a, b, 2013c, d, e, f, g, h, Valipour 2014b). The MFP model is an effective tool to model nonlinear systems. The obtained results of predictions are in good agreement with the calculated ET at all stations. The results showed that the accuracy of MFP model was greater than the other two models in estimation of the amount of ET. Namely, the MFP model has more regression coefficient values and the less prediction errors than the other two models. Therefore, the suggestion is that these models, especially the MFP, could apply for other climatic and hydrologic modeling.

One of the most important results of the current investigation is estimation of reference ET by using only two weather variables including temperature and radiation, with an acceptable accuracy. Therefore, the formulae obtained are preferred to (1) other models in which reference ET was estimated using only these to variables with a future error as well as (2) other models in which reference ET was estimated using more weather variables such as relative humidity, wind speed, elevation, and precipitation.

References

Abtew W (1996) Evapotranspiration measurements and modeling for three wetland systems in South Florida. J Water Resour Bull 32(3):465–473

Abtew W, Obeysekera J (1995) Lysimeter study of evaporation of cattails and comparison of 3 estimation methods. Trans ASAE 38(1):121–129

Allen RG (1986) A Penman for all seasons. J Irrig Drain Eng 112(4):348–368

Allen RG, Jensen ME, Wright JL, Burman RD (1989) Operational estimates of reference evapotranspiration. J Agron 81:650–662

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration guidelines for computing crop water requirements. FAO irrigation and drainage. Paper no. 56, FAO, Rome

Azizi G, Hanafi A, Soltani M (2010) Evaluation of potential evapotranspiration via multi-variations regression in Esfahan province. Geography 11(3):77–92 **(in Persian)**

Bachour R, Walker WR, Ticlavilca AM, McKee M, Maslova I (2014) Estimation of spatially distributed evapotranspiration using remote sensing and a relevance vector machine. J Irrig Drain Eng 140(8):04014029

Banihabib ME, Valipour M, Behbahani SMR (2012) Comparison of autoregressive static and artificial dynamic neural network for the forecasting of monthly inflow of Dez reservoir. J Environ Sci Technol 13(4):1–14

Belsley DA, Kuh E, Welsch RE (1980) Regression diagnostics. Wiley, New York

Block P, Rajagopalan B (2009) Statistical-dynamical approach for streamflow modeling at malakal, sudan, on the white Nile river. J Hydrol Eng 14(2):185–196

Box GEP, Tidwell PW (1962) Transformation of the independent variables. Technometrics 4:531–550

Chiew FHS, Kamaladassa NN, Malano HM, McMahon TA (1995) Penman-Monteith, FAO-24 reference crop evapotranspiration and class-A pan data in Australia. J Agr Water Manage 28:9–21

Chulani S, Boehm B, Steece B (1998) Calibrating software cost models using bayesian analysis. University of Southern California. Henry Salvatori, Los Angeles, CA 90089-0781. USC-CSE

Coleman D, Holland P, Kaden N, Klema V (1980) A system of subroutines for iteratively reweighted least squares computations. ACM Trans Math Softw 6:327–336

De Souza F, Yoder RE (1994) ET estimation in the northeast of Brazil: Hargreaves or penman-monteith equation? In: Proceedings of technical paper ASAE international winter meeting, American Society of Agricultural Engineers, St. Joseph, Mich

Demoster AP, Laird NM, Rubin DB (1980) Iteratively reweighted least squares for linear regression when errors are normal/independent distributed. Multivariate analysis V. Krishnaiah PR, North-Holland, New York, pp 35–37

Douglas EM, Jacobs JM, Sumner DM, Ray RL (2009) A comparison of models for estimating potential evapotranspiration for Florida land cover types. J Hydrol 373:366–376

Ershadi A, McCabe MF, Evans JP, Mariethoz G, Kavetski D (2013) A Bayesian analysis of sensible heat flux estimation: quantifying uncertainty in meteorological forcing to improve model prediction. Water Resour Res 49:2343–2358

Eslamian SS, Abedi-Koupai J, Amiri MJ, Gohari SA (2009) Estimation of daily reference evapotranspiration using support vector machines and artificial neural networks in Greenhouse. Res J Environ Sci 3:439–447

Ghosh A, Joshi PK (2014) Hyperspectral imagery for disaggregation of land surface temperature with selected regression algorithms over different land use land cover scenes. ISPRS J Photogramm Remot Sens 96:76–93

Heydari MM, Abbasi A, Heydari M (2013) Estimation of evapotranspiration in Ardestan, center of Iran. World Appl Sci J 21(2):230–236

Iizumi T, Nishimori M, Yokozawa M, Kotera A, Duy Khang N (2012) Statistical downscaling with Bayesian inference: estimating global solar radiation from reanalysis and limited observed data. Int J Climatol 32:464–480

Izadifar Z, Elshorbagy A (2010) Prediction of hourly actual evapotranspiration using neural networks, genetic programming, and statistical models. Hydrol Process 24:3413–3425

Jain SK, Nayak PC, Sudheer KP (2008) Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. Hydrol Process 22:2225–2234

Jensen ME, Burman RD, Allen RG (1990) Evapotranspiration and irrigation water requirements. ASCE manual and report on engineering practice no. 70. ASCE, New York

Jones JW, Keating BA, Porter CH (2001) Approaches to modular model development. Agr Syst 70:421–443

Joshi D, St-Hilaire A, Daigle A, Ouarda TB (2013) Databased comparison of sparse bayesian learning and multiple linear regression for statistical downscaling of low flow indices. J Hydrol 488:136–149

Kumar M, Raghuwanshi NS, Singh R, Wallender WW, Pruitt WO (2002) Estimating evapotranspiration using artificial neural network. J Irrig Drain Eng 128(4):224–233

Liang Z, Wang D, Guo Y, Zhang Y, Dai R (2013) Application of bayesian model averaging approach to multimodel ensemble hydrologic forecasting. J Hydrol Eng 18(11):1426–1436

Mahdizadeh Khasraghi M, Gholami Sefidkouhi MA, Valipour M (2015) Simulation of open- and closed-end border irrigation systems using SIRMOD. Arch Agron Soil Sci 61(7):929–941

Marshall L, Nott D, Sharma A (2007) Towards dynamic catchment modelling: a Bayesian hierarchical mixtures of experts framework. Hydrol Process 21:847–861

Matlab Statistics Toolbox (2010) User manual MATLAB STATISTICS TOOLBOX 7. http://somemanuals.com/MATLAB,STATISTICS–TOOLBOX–7,ICDFDDF.PDF

McNaughton KG, Jarvis PG (1984) Using Penman-Monteith equation predictively. Agr Water Manage 8:263–278

Ojha S, Bhakar SR (2012) Estimation of evapotranspiration for wheat crop using artificial neural network. J Indian Water Resour Soci 32:13–21

Pejic B, Aksic M, Mackic K, Sekularac G (2015) Response of potato to water stress in Southern Serbia. Austin J Irrig 1(1):1001

Rahimi S, Gholami Sefidkouhi MA, Raeini-Sarjaz M, Valipour M (2014) Estimation of actual evapotranspiration by using MODIS images (a case study: tajan catchment). Arch Agron Soil Sci 61(5):695–709

Rahimkhoob A (2009) Estimating daily pan evaporation using artificial neural network in a semi-arid environment. Theor Appl Climatol 98:101–105

Rathnayake RC (2010) Bayesian inference and computation. Research papers. Paper 29. http://opensiuc.lib.siu.edu/gs_rp/29

Rosenberry DO, Winter TC, Buso DC, Likens GE (2007) Comparison of 15 evaporation methods applied to a small mountain lake in the Northeastern USA. J Hydrol 340:149–166

Royston P, Altman DG (1994) Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling (with discussion). Appl Stat 43:429–467

Royston P, Sauerbrei W (2005) Building multivariable regression models with continuous covariates in clinical epidemiology with an emphasis on fractional polynomials. Meth Inf Med 44:561–571

Royston P, Sauerbrei W (2007) Improving the robustness of fractional polynomial models by preliminary covariate transformation: a pragmatic approach. Comput Stat Data Anal 51:4240–4253

Sauerbrei W, Royston P (1999) Building multivariable prognostic and diagnostic models: transformation of the predictors using fractional polynomials. J Roy Stat Soc (Series A) 162:71–94 Corrigendum: J Roy Stat Soc (2002) (Series A) 165:399–400

Schuch UK, Burger DW (1997) Water use and crop coefficients of woody ornamentals in containers. J Am Soc Hort Sci 122:727–734

Singh K (2014) A review on software cost estimation models and techniques. IJETI Int J Eng Technol Innovat 1(3):17–26

Tabari H, Kisi O, Ezani A, Hosseinzadeh Talaee P (2012) SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. J Hydrol 444:78–89

Ticlavilca AM, McKee M, Walker WR (2013) Real-time forecasting of short-term irrigation canal demands using a robust multivariate Bayesian learning model. Irrig Sci 31(2):151–167

Tzimopoulos C, Mpallas L, Papaevangelou G (2007) Estimation of evapotranspiration using fuzzy rules and comparison with the Blaney-Criddle method. In: Water Resour Manag: new approaches and technologies. Edited by: European Water Resources Association, pp 245–252 Chania, Crete, Greece, June 24–28

Valipour M (2012a) Hydro-module determination for Vanaei Village in Eslam Abad Gharb. Iran. ARPN J Agr Biol Sci 7(12):968–976

Valipour M (2012b) Ability of box-Jenkins models to estimate of reference potential evapotranspiration (A case study: Mehrabad synoptic station, Tehran, Iran). IOSR J Agr Veter Sci 1(5):1–11

Valipour M (2012c) Comparison of Surface Irrigation Simulation Models: full Hydrodynamic, Zero Inertia. Kinematic Wave. J Agr Sci 4(12):68–74

Valipour M (2012d) Scrutiny of pressure loss, friction slope, inflow velocity, velocity head, and reynolds number in Center Pivot. Int J Adv Sci Technic Res 2(5):703–711

Valipour M (2012e) Determining possible optimal values of required flow, nozzle diameter, and wetted area for linear traveling laterals. Int J Eng Sci 1(1):37–43

Valipour M (2012f) Sprinkle and trickle irrigation system design using tapered pipes for pressure loss adjusting. J Agr Sci 4(12):125–133

Valipour M (2013a) Use of surface water supply index to assessing of water resources management in colorado and oregon, US. Adv Agr Sci Eng Res 3(2):631–640

Valipour M (2013b) Estimation of surface water supply index using snow water equivalent. Adv Agr Sci Eng Res 3(1):587–602

Valipour M (2013c) Necessity of irrigated and rainfed agriculture in the world. Irrig Drain Sys Eng S9:e001. http://www.omicsgroup.org/journals/2168-9768/pdfdownload.php?download=2168-9768-S9-e001.pdf&&aid=12800

Valipour M (2013d) Increasing irrigation efficiency by management strategies: cutback and surge irrigation. ARPN J Agr Biol Sci 8(1):35–43

Valipour M (2013e) Evolution of irrigation-equipped areas as share of cultivated areas. Irrig Drain Sys Eng 2(1):e114. http://www.omicsgroup.org/journals/evolution-of-irrigationequipped-areas-as-share-of-cultivated-areas-2168-9768.1000e114.php?aid=12797

Valipour M (2013f) Need to update of irrigation and water resources information according to the progresses of agricultural knowledge. Irrig Drain Sys Eng S10:e001. http://www.omicsgroup.org/journals/need-to-update-of-irrigation-and-water-resources-information-according-to-the-progresses-of-agricultural-knowledge-2168-9881.S10-e001.php?aid=15774

Valipour M (2013g) Comparison of different drainage systems usable for solution of environmental crises in soil. In: The 1st international conference on environmental crises and its solutions, Kish Island, Iran

Valipour M (2013h) Scrutiny of inflow to the drains applicable for improvement of soil environmental conditions. In: The 1st international conference on environmental crises and its solutions, Kish Island, Iran

Valipour M (2014a) Analysis of potential evapotranspiration using limited weather data. Appl Water Sci. Assessed 27 September 2014. doi:10.1007/s13201-014-0234-2

Valipour M (2014b) Application of new mass transfer formulae for computation of evapotranspiration. J Appl Water Eng Res 2(1):33–46

Valipour M (2014c) Variations of irrigated agriculture indicators in different continents from 1962 to 2011. Adv Water Sci Technol 1(1):1–14

Valipour M (2014d) Assessment of different equations to estimate potential evapotranspiration versus FAO Penman Monteith method. Acta Adv Agr Sci 2(11):14–27

Valipour M (2014e) Irrigation status of Americas. Adv Appl Agr Sci 2(12):56–72

Valipour M (2014f) Handbook of irrigation engineering problems. OMICS Group eBooks, Foster City. http://www.esciencecentral.org/ebooks/handbook-of-irrigation-engineering-problems/pdf/handbook-of-irrigation-engineering-problems.pdf

Valipour M (2014g) Handbook of drainage engineering problems. OMICS Group eBooks, Foster City. http://www.esciencecentral.org/ebooks/handbook-of-drainage-engineering-problems/pdf/handbook-of-drainage-engineering-problems.pdf

Valipour M (2014h) Handbook of hydraulic engineering problems. OMICS Group eBooks, Foster City. http://www.esciencecentral.

org/ebooks/handbook-of-hydraulic-engineering-problems/pdf/handbook-of-hydraulic-engineering-problems.pdf

Valipour M (2014i) Handbook of hydrologic engineering problems. OMICS Group eBooks, Foster City. http://www.esciencecentral.org/ebooks/handbook-of-hydrologic-engineering-problems/pdf/handbook-of-hydrologic-engineering-problems.pdf

Valipour M (2014j) Handbook of water engineering problems. OMICS Group eBooks, Foster City. http://www.esciencecentral.org/ebooks/handbook-of-water-engineering-problems/pdf/handbook-of-water-engineering-problems.pdf

Valipour M (2014k) Prediction of irrigated agriculture in Asia Pacific using FAO indices. Acta Adv Agr Sci 2(9):40–53

Valipour M (2014l) Pressure on renewable water resources by irrigation to 2060. Acta Adv Agr Sci 2(8):32–42

Valipour M (2014m) Future of agricultural water management in Europe based on socioeconomic indices. Acta Adv Agr Sci 2(7):1–18

Valipour M (2015a) What is the tendency to cultivate plants for designing cropping intensity in irrigated area? Adv Water Sci Technol 2(1):1–12

Valipour M (2015b) Assessment of important factors for water resources management in European agriculture. J Water Res Hydraul Eng 4(2):169–178

Valipour M (2015c) Handbook of environmental engineering problems. OMICS Group eBooks, Foster City. http://www.esciencecentral.org/ebooks/handbook-of-environmental-engineering-problems/pdf/handbook-of-environmental-engineering-problems.pdf

Valipour M (2015d) Calibration of mass transfer-based models to predict reference crop evapotranspiration. Appl Water Sci. Assessed 5 March 2015. doi:10.1007/s13201-015-0274-2

Valipour M (2015e) Optimization of neural networks for precipitation analysis in a humid region to detect drought and wet year alarms. Met Appl. Assessed 26 November 2015. doi:10.1002/met.1533

Valipour M, Eslamian S (2014) Analysis of potential evapotranspiration using 11 modified temperature-based models. Int J Hydrol Sci Technol 4(3):192–207

Valipour M, Mousavi SM, Valipour R, Rezaei E (2012) SHCP: soil heat calculator program. IOSR J Appl Phys 2(3):44–50

Valipour M, Gholami Sefidkouhi MA, Eslamian S (2015) Surface irrigation simulation models: a review. Int J Hydrol Sci Technol 5(1):51–70

Verstraeten WW, Veroustraete F, Feyen J (2008) Assessment of evapotranspiration and soil moisture content across different scales of observations. Sensor 8:70–117

Vigiak O, Bende-Mich U (2013) Estimating bootstrap and Bayesian prediction intervals for constituent load rating curves. Water Resour Res 49:8565–8578

Walter IA, Allen RG, Elliot R, Itenfisu D, Brown P (2002) The ASCE standardized reference evapotranspiration equation. In: Proceeding of the Rep. Task Com. on standardized reference evapotranspiration, July 9, American Society of Civil Engineering, Reston, VA, USA, pp 1–70

Weib M, Menzel L (2008) A global comparison of four potential evapotranspiration equations and their relevance to stream flow modelling in semi-arid environments. Adv Geosci 18:15–23

Wellen C, Arhonditsis GB, Labencki T, Boyd D (2012) A Bayesian methodological framework for accommodating interannual variability of nutrient loading with the SPARROW model. Water Resour Res 48(10):W10505

Wrachien DD, Mambretti S (2015) Irrigation and drainage systems in flood-prone areas: the role of mathematical models. Austin J Irrig 1(1):1002

Wu IP (1997) A Simple Evapotranspiration Model for Hawaii: The Hargreaves Model. Department of Biosystems Engineering. Cooperative Extension Service. CTAHR Fact Sheet Engeineer's Notebook no. 106

Yannopoulos SI, Lyberatos G, Theodossiou N, Li W, Valipour M, Tamburrino A, Angelakis AN (2015) Evolution of water lifting devices (pumps) over the centuries worldwide. Water 7(9):5031–5060

Yao Y, Liang S, Li X, Hong Y, Fisher JB, Zhang N, Jiang B, Sun L, Jia K, Wang K, Chen Y, Mu Q, Feng F (2014) Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations. J Geophys Res Atmos 119:4521–4545

Zhu G, Su Y, Li X, Zhang K, Li C (2013) Estimating actual evapotranspiration from an alpine grassland on Qinghai-Tibetan plateau using a two-source model and parameter uncertainty analysis by Bayesian approach. J Hydrol 476:42–51