Random forest predictive model development with uncertainty analysis capability for the estimation of evapotranspiration in an arid oasis region

Min Wu, Qi Feng, Xiaohu Wen, Ravinesh C. Deo, Zhenliang Yin, Linshan Yang and Danrui Sheng

ABSTRACT

The study evaluates the potential utility of the random forest (RF) predictive model used to simulate daily reference evapotranspiration (ET_0) in two stations located in the arid oasis area of northwestern China. To construct an accurate RF-based predictive model, ET_0 is estimated by an appropriate combination of model inputs comprising maximum air temperature (T_{max}), minimum air temperature (T_{min}), sunshine durations (S_{un}), wind speed (U_2), and relative humidity (R_h). The output of RF models are tested by ET_0 calculated using Penman–Monteith FAO 56 (PMF-56) equation. Results showed that the RF model was considered as a better way to predict ET_0 for the arid oasis area with limited data. Besides, R_h was the most influential factor on the behavior of ET_0, except for air temperature in the proposed arid area. Moreover, the uncertainty analysis with a Monte Carlo method was carried out to verify the reliability of the results, and it was concluded that RF model had a lower uncertainty and can be used successfully in simulating ET_0. The proposed study shows RF as a sound modeling approach for the prediction of ET_0 in the arid areas where reliable weather data sets are available, but relatively limited.

Key words | arid areas, evapotranspiration, Monte Carlo, predict, random forest

HIGHLIGHTS

- Evapotranspiration is an essential hydrological property used for the computation of water balance, including the scheduling of irrigation systems, water resources planning, and management for agricultural purposes, especially in an arid region.
- Random forest model is designed for estimation of evapotranspiration in an arid oasis region.
- The Monte-Carlo method is carried to analyze uncertainty of simulation results.
- The model can be used successfully in simulating evapotranspiration in arid regions where weather data are limited.
- Model has a lower uncertainty and can provide reliable tool of modeling evapotranspiration under the same climatic conditions.

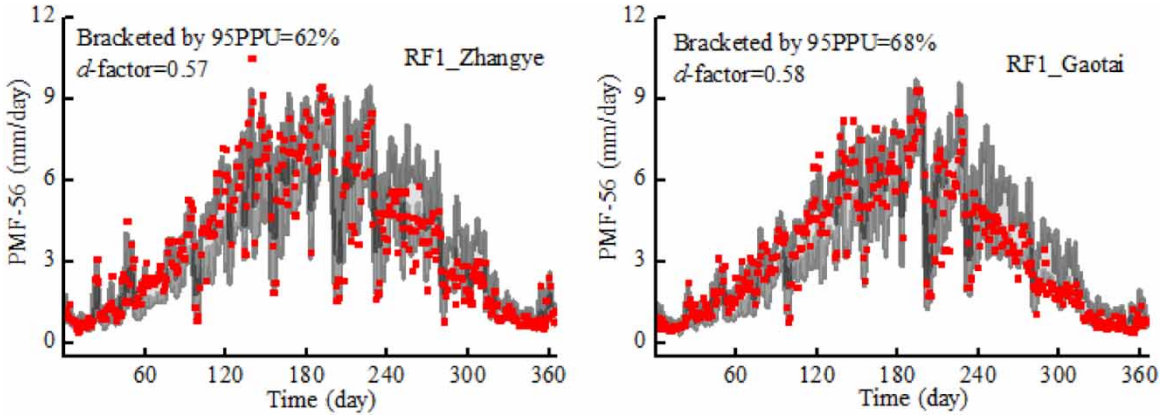
Min Wu
Qi Feng
Xiaohu Wen (corresponding author)
Zhenliang Yin
Linshan Yang
Danrui Sheng
Key Laboratory of Ecohydrology of Inland River Basin,
Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences,
Lanzhou 730000,
China
E-mail: *xhwen@lzb.ac.cn*

Min Wu
Danrui Sheng
University of Chinese Academy of Sciences,
Beijing 100049,
China

Ravinesh C. Deo
School of Agricultural, Computational and Environmental Sciences, Centre for Sustainable Agricultural Systems,
University of Southern Queensland,
Springfield, QLD 4300,
Australia

doi: 10.2166/nh.2020.012

GRAPHICAL ABSTRACT



INTRODUCTION

Evapotranspiration (ET) is the process of transfer of water from the surface of the earth to the atmosphere including evaporation and transpiration (Shiri *et al.* 2014; Nourani *et al.* 2019), and often used to estimate actual evapotranspiration in water balance studies and water resources management (Tao *et al.* 2015). In arid oasis conditions, crops are a material basis on which human beings depend for their survival as well as being an ecological protection barrier in such areas. Knowledge of crop-water demands is an important practical consideration for improved water-use efficiency (Benli *et al.* 2006). This is because ET is a primary source of water loss, so its accurate evaluation can provide valuable information for water balance, irrigation system design, and water resources management (Torres *et al.* 2011; Wen *et al.* 2015). This is especially true for arid regions, such as the northwest region in China, where population growth, expansion of agriculture, and other socio-economic activities are significantly constraining the available water resources.

Due to the lack of observation data, the precise estimation of ET has produced the need for another comprehensive concept called reference evapotranspiration (ET_0) (Abdullah *et al.* 2014). ET_0 can be measured directly using lysimeters which are characterized by providing accurate measurement results; however, the application of the methods is limited due to their cost and complexity (Ferreira *et al.* 2019), which increases the requirements of employing data-based methods to predict ET_0. Several conventionally empirical models like Hargreaves equation, Priestley–Taylor equation, and Ritchie equation have been developed to estimate ET_0 using meteorological data. Because the PMF-56 equation takes into account moisture availability, mass transfer, and required energy for the process (Granata 2019), it has been recommended for the computation of ET_0 by the Food and Agricultural Organization of the United Nations (FAO) as the only standard equation which is usually applied to validate other models and has been accepted in many regions across the world. PMF-56 equation can be broadly applied in various environments and climate conditions due to its good precision and stability (Huang *et al.* 2019). However, some restrictions still exist in the application of PMF-56 equation, for example, it is difficult to obtain all meteorological data required in the estimation process, particularly in a developing country, where the number of meteorological stations is limited and weather data records could be scarce (Abdullah *et al.* 2015). Within this context, an alternative data-driven model which

requires easily available input variables is necessary and significant.

As the ET_0 depends on several interacting meteorological factors, such as temperature, humidity, wind speed, and radiation, it is difficult for the ordinary formula to express all the related physical processes (Yassin *et al.* 2016; Yin *et al.* 2016). In this context, artificial intelligence or data-driven models are considered as efficient tools to deal with non-linear relationships between independent and dependent variables. In the past few decades, artificial intelligence models, including artificial neural network (ANN), extreme learning machine (ELM), support vector machine (SVM), and so on, have been extensively used in the area of predicting and forecasting (Kisi & Cimen 2009; Yoon *et al.* 2011; Tabari *et al.* 2012; Acharya *et al.* 2013; He *et al.* 2014; Deo & Şahin 2015). In terms of ET_0 prediction, Traore *et al.* (2010) assessed the performance of feed forward backpropagation neural network (BPNN) algorithm (a type of ANN) based on different inputs in estimating ET_0 in the Bobo-Dioulasso region. The results showed that the BPNN algorithm had a better performance than conventional Hargreaves equation and that wind was found to be the most effective variable significantly required for modeling with high accuracy when added into inputs. Huo *et al.* (2012) compared the performance of ANN models with multiple linear regressions, the Penman equation, and two empirical equations for calculation of ET_0 in northwest China, concluding that ANN models exhibited higher accuracy than the others, and they also concluded that temperature, R_h, was the most important input affecting ET_0. Abdullah *et al.* (2015) proved that ELM was efficient, simple in application, of high speed, and had a very good generalization performance at predicting Penman–Monteith (P-M) ET_0 using four different complete and incomplete meteorological input combinations in Iraq. Patil & Deka (2016) developed the ELM model utilizing three different input combinations to calculate ET_0 in the Thar Desert, India, and Hargreaves equation, ANN and least-square support vector machine (LS-SVM) models were used for a contrast. The results revealed that ELM is a simple yet efficient algorithm and superior to the other two methods. Tabari *et al.* (2012) estimated the performances of SVM, adaptive neuro-fuzzy inference system (ANFIS), multiple linear regression (MLR), and multiple non-linear regression (MNLR) for estimating ET_0 using six input vectors of climatic data in a semi-arid highland environment in Iran. The results displayed that the capability of SVM and ANFIS models for ET_0 prediction was better than those achieved using the regression and climate-based models. Kisi & Cimen (2009) used the SVM approach for modeling ET_0 in three stations in central California. The results were compared with empirical models and ANN model and revealed that the SVM method could be employed successfully in simulating the ET_0 process. These models have demonstrated promising prediction ability of ET_0 in many parts of the world, but some deficiencies exist. ANN models become easily stuck in a local minimum, and the optimization process is effortlessly influenced by initial point selection. SVM and numerous ELM models are machine learning methods based on kernel function, and generalization abilities depend largely on the choice of the kernel function.

Random forest (RF) is another emerging machine learning technique and a natural non-linear modeling tool, the superiority of which is good tolerance for outliers and noise, difficulty in producing an over-fitting phenomenon. As well, it can overcome the 'black-box' limitations of ANN and provides evaluation of the importance degree of input variables (Rodriguez-Galiano *et al.* 2014). RF with its merits has been widely used in classification and prediction (Gislason *et al.* 2006; Cutler *et al.* 2007; Heung *et al.* 2014; Gong *et al.* 2018). Wang *et al.* (2015) proposed the RF model to evaluate flood hazard risk and implemented the method in Dongjiang River Basin, China; consequently, the capacity of the RF model was similar to the SVM model with a correlation coefficient of 0.916, but the RF method had a better performance with its advantages including providing credible assessment consequences of importance degree of input variables. Dong *et al.* (2013) classified whether rockburst will happen and the intensity of rockburst in underground rock projects utilizing RF method, and selected some main control factors of rockburst, including the values of *in-situ* stresses, uniaxial compressive strength and tensile strength of rock, and the elastic energy index of rock to analysis. The results indicated that the RF model exhibited high classification accuracy compared with the ANN and SVM approach with misjudgment ratios of 0, 10%, and 20%, respectively. In RF modeling of ET_0, Fukuda *et al.* (2013) accessed the

applicability of RF model for estimating mango fruit yields using 10-day rainfall and irrigation data in response to water supply under different irrigation regimes. The RF models accurately estimated the maximum and mean values of mango fruit yields, and the results displayed the applicability of RF in the field of agricultural engineering. Feng *et al.* (2017) proposed RF and generalized regression neural networks (GRNN) models for daily ET_0 estimation in southwest China, and the result revealed that the RF model was slightly better than GRNN model for estimating daily ET_0. Although the RF model demonstrated significant potential in many studies, the use of RF model for evaluating ET_0 has been rarely recorded by research, especially in the arid environment of northwest China. It is, thus, important to predict ET_0 using the RF model to provide a reliable method in data-limited areas.

Despite these advantages, there is a deficiency in the application of the RF model for ET_0 predictions. Almost all the artificial intelligence models are stochastic algorithms, the RF approach is no exception, and running the model will not reproduce the same result even in an identical situation. Uncertainty analysis is an indispensable procedure for getting reliable results in model simulations. For uncertainty analysis, two primarily different aspects of uncertainty include uncertain input variables, model parameters, and model structure. By means of its general applicability, the Monte Carlo simulation technique is a widely used method for uncertainty analysis in hydrological modeling (Shrestha *et al.* 2009; Antanasijević *et al.* 2014). However, one remarkable issue is that the uncertainty of the model in estimation is usually ignored by most studies, and no such studies have been reported adding uncertainty analysis in predicting ET_0 so far. In this condition, uncertainty analysis is conducted in the paper for assessing the precision of the RF model.

The present study was carried out in an arid oasis area of the middle reaches of the Heihe River Basin, northwest China (Figure 1), where water resources play an important role in the sustainable development of the ecological environment. Besides, the study area is a typical irrigated agricultural area as well as an important commodity grain base of Heihe River. Agriculture consumes most water, plus water resources are in severely short supply in this region. As a vital component to describe the hydrological cycle, estimate water balance, and schedule irrigation (Rawat *et al.* 2019), ET_0 determination with reliable accuracy is significant in such a water-scarce region (Nourani *et al.* 2020). Hence, the objective of this paper was to investigate the precision of the RF model by using different variables' combination of meteorological data including T_{max}, T_{min}, U_2, R_h, S_{un} in an arid region, northwest China. The results obtained from the RF modes for various input combinations are compared to each other, and subsequently determine the effects of different meteorological arguments on ET_0 according to the importance degree of variables. Moreover, the uncertainty analysis is performed for the RF model by Monte Carlo simulations for the purpose of a better accurate result applying to arid areas.

MATERIALS AND METHODS

PMF-56 equation

As a standard method to estimate ET_0, PMF-56 equation was used to be a RF target output to train and test the model in this paper and proposed by Allen *et al.* (1998) as follows:

$$ET_{0-\mathrm{PMF}-56} = \frac{0.408\Delta(R_n - G) + \gamma \dfrac{900}{T+273} U_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34U_2)} \quad (1)$$

where $ET_{0\text{-PMF-}56}$ is the reference evapotranspiration (mm day^{-1}); R_n is the net radiation at the crop surface (MJ m^{-2} day^{-1}); G is the soil heat flux (MJ m^{-2} day^{-1}); γ is the psychrometric constant (kPa $°C^{-1}$); T is the mean daily air temperature at 2 m height (°C); U_2 is the mean daily wind speed at 2 m height (m s^{-1}); e_s is the saturation vapor pressure (kPa), e_a is the actual vapor pressure (kPa), $e_s - e_a$ is the saturation vapor pressure deficit (kPa); Δ is the slope of the saturation vapor pressure-temperature curve (kPa $°C^{-1}$). Allen *et al.* (1998) described the calculation process of each parameter required to compute ET_0 in detail, and all parameters could be calculated by meteorological data obtained directly by weather stations.

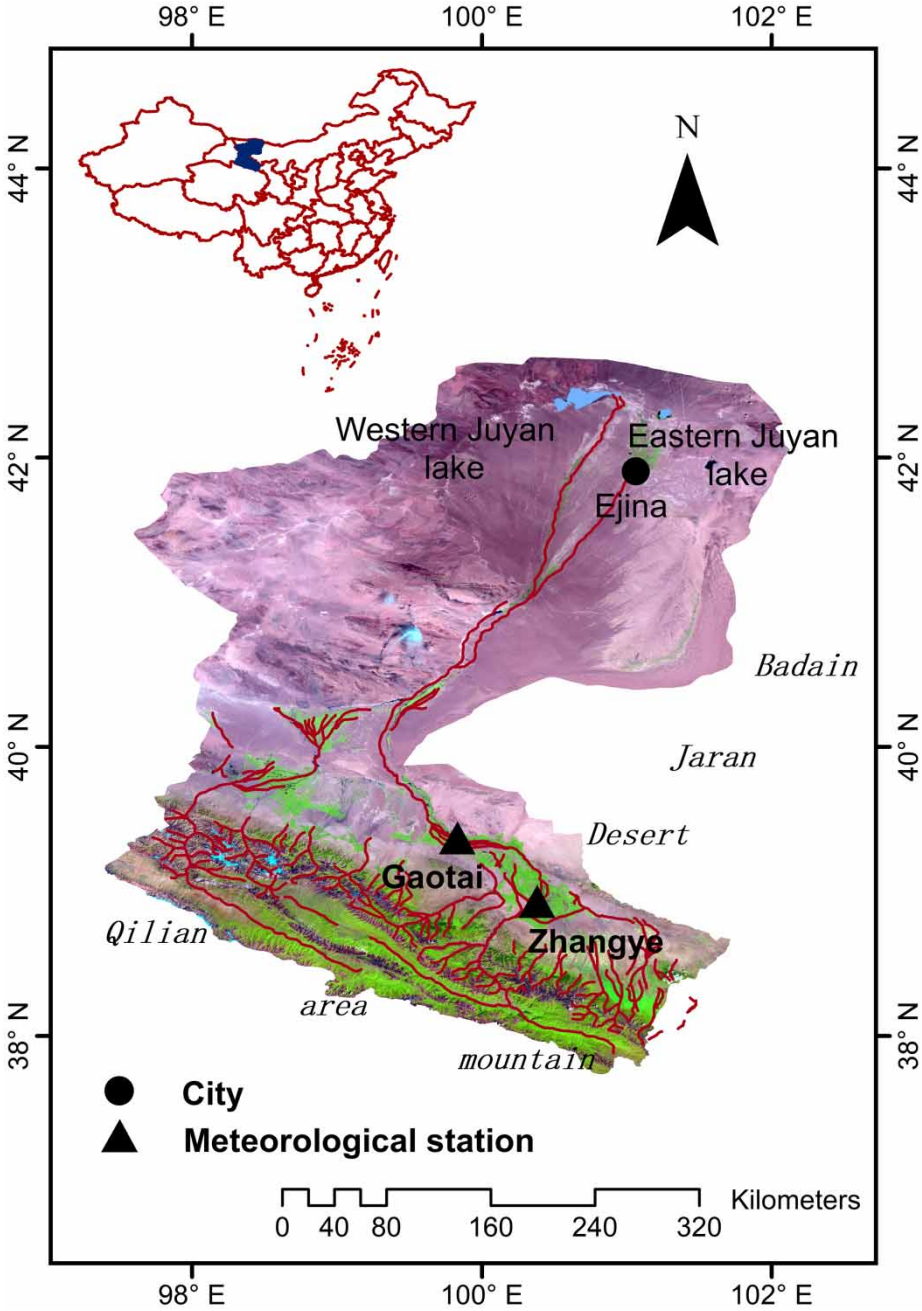


Figure 1 | Location study area and the climate data measured sites.

RF

The RF was developed by Breiman (2001) based on a CART decision tree model, including regression (RFR) and classification (RFC) algorithm. The basic idea based on statistical theory is that extracting repeatedly and randomly K samples from the original training sample set N for generating a new set of training samples through the bootstrap resampling method, then producing K decision trees and comprising random forest according to the bootstrap sample set. In terms of the classification model, the classified results of new data depend on the number of votes obtained by

classification tree votes, and for the regression model, all the averages of the predictive value of decision trees are regarded as final prediction outcomes (Figure 2).

This paper uses the regression algorithm whose calculation processes are as follows.

First, randomly generate k training samples (Θ_1, Θ_2, …, Θ_k) from the total training sample using the bootstrap sampling method, corresponding to K decision trees can be constructed.

Second, at each node of the decision tree, the m features are randomly selected from the M features as the splitting features set of the current nodes, then selecting one node from the m features to split according to the principle of node purity minimum, each decision tree is grown to the largest extent possible, no pruning.

Third, for new data, the predictive value of a single decision tree can be obtained through the average of the observations of the leaf node 1(x, Θ). If an observation value Xi is a leaf node 1(x, Θ) and not 0, the weight $\omega_i(x,\Theta)$ is set as:

$$\omega_i(x\Theta) = \frac{1\{x_i \epsilon R_l(x,\Theta)\}}{\#\{j : x_j \epsilon R_l(x,\Theta)\}} \quad (2)$$

where the sum of weights equals 1.

Fourth, the prediction of a single decision tree gained by the weighted average of the observations of dependent variables is defined as:

$$\mu(x) = \sum_{i=1}^{n} \omega_i(x\Theta) Y_i \quad (3)$$

where Y_i (i = 1,2,…,n) is the observation of the dependent variable.

Finally, given weight of decision tree $\omega_i(x,\Theta_t)$ (t = 1, 2, …, k), the weight of each observation as Equation (4):

$$\omega_i(x) = \frac{1}{k}\sum_{i=1}^{k} \omega_i(x\Theta_t)\ Y \quad (4)$$

thus, the final predicted value of RFR is:

$$\mu(x) = \sum_{i=1}^{n} \omega_i(x) Y_i \quad (5)$$

the flowchart of RF for regression is shown as follows.

In addition, index importance assessment is a prominent advantage of the RF algorithm, the purpose of which lies in evaluating the effect of each variable on the accuracy of the RF model. IncNodePurity index adopted in this research was used to assess the importance of each parameter, and compare that by calculating the reduced values of impurity of the nodes of all tree variables. That higher index importance measurement can intuitively reflect the main factors affecting estimated ET_0. Besides, in this research, we applied the randomForest package to train data and access variable importance in the R environment.

Uncertainty analysis

Uncertainty analysis by Monte Carlo simulations is used for evaluating the analysis of final models. Input parameter uncertainty considered in this paper is related to the precision and representativeness of the input data applied for predictions (Antanasijević *et al.* 2014). In this method, the input parameter is described using a probability distribution and a single input data set involves the generation of random input respecting this distribution, then running the model and obtaining

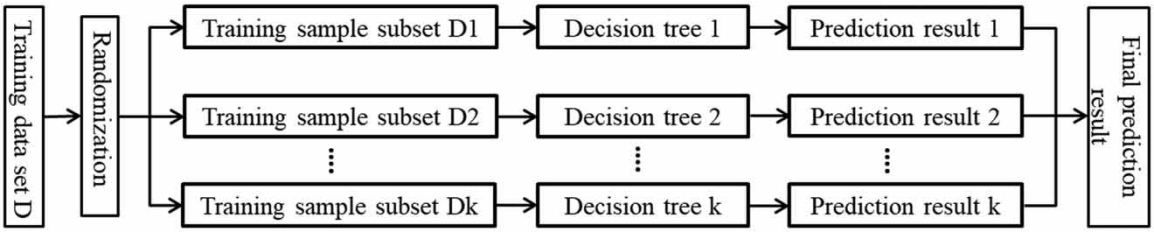


Figure 2 | Schematic of random forest workflow.

output (Noori *et al.* 2010). In the present work, we randomly resample the input data set without replacement for 1,000 times, keeping the ratio between the training and validation sets unchanged (Dehghani *et al.* 2014; Gao *et al.* 2018). Finally, the 95% confidence intervals are determined by finding the 2.5th (X_L) and 97.5th (X_U) percentiles of the cumulative distribution consisting of 1,000 data. The ratio of observed values that lie within the 95% confidence interval is calculated as judging the robustness metric of the final model; the higher the ratio is, the stronger the robustness is, and vice versa. The 95% prediction uncertainties (95PPU) are represented as:

$$Bracketed\ by\ 95PPU = \frac{1}{n} Count(N | X_L \le N \le X_U) \times 100 \quad (6)$$

where the n indicates the number of observed data points. N is increasing with the value of PMF-56 ET_0 falling between corresponding X_L and X_U increase, the 'Bracketed by 95PPU' is 100 when all of the PMF-56 ET_0 values are within the range of $X_L \le N \le X_U$.

In addition, *d-factor* (Ghorbani *et al.* 2016) is applied for computing the average width of the confidence interval, and can be evaluated according to Equation (7):

$$d - factor = \bar{d}_x / \bar{\sigma}_x \quad (7)$$

$$\bar{d}_x = \frac{1}{n} \sum_{i=1}^{n} (X_U - X_L) \quad (8)$$

where $\bar{d}_x$ is the average distance between the upper (97.5th) and lower (2.5th) bands, $\bar{\sigma}_x$ is the standard deviation of the observed data. It is relevant to note that the better results would have a *d-factor* value which is close to 0.

CASE STUDY

Observation data and statistical analysis

The weather data for this study were obtained from two sites in Zhangye (100°17′E, 39°05′N) and Gaotai (99°50′E, 39°22′N), as shown in Figure 1. In this study, five years of meteorological data was sourced from the National Climatic Centre of the China Meteorological Administration. The duration of the data is from 2013 to 2017 at daily timescales, which includes T_{max}, T_{min}, U_2, R_h, S_{un}. There were 1,826 records and these were divided into two parts: the training part composed of 1,461 daily records which account for about 80% of the total data set, and the testing part, the remaining 365 records, which accounts for about 20% of the total data set. The statistical characteristics of daily weather data and the PMF-56 ET_0 for each station are shown in Table 1. In terms of the skewness values, T_{max}, T_{min}, and R_h showed lower skewed distribution than other variables. Also, it can be seen that U_2 shows a higher skewed feature than the other variables (1.06 and 1.21 for the two sites, respectively). T_{max}, T_{min}, and PMF-56 ET_0 demonstrate a strong variability and the CV values exceed 0.62, which principally resulted from seasonal changes. The

Table 1 | Statistical parameters of climatic data and the PMF-56 ET_0 at two stations

Station	Climatic data and the PMF − 56 ET_0	Maximum	Minimum	Mean	Std.	SK	CV
Zhangye	T_{max} (°C)	39.60	−13.50	17.09	11.42	−0.31	0.67
	T_{min} (°C)	22.80	−28.60	1.91	11.84	−0.26	6.20
	S_{un} (h)	14.00	0	8.58	3.38	−0.96	0.39
	U_2 (m/s)	8.00	0.90	2.84	0.98	1.06	0.35
	R_h (%)	100.00	10.00	45.91	16.96	0.47	0.37
	PMF − 56 ET_0 (mm/day)	11.67	0.11	3.63	2.52	0.59	0.69
Gaotai	T_{max} (°C)	39.80	−11.70	17.64	11.61	−0.32	0.66
	T_{min} (°C)	25.90	−26.70	2.65	11.20	−0.18	4.27
	S_{un} (h)	13.80	0	8.40	3.35	−0.91	0.40
	U_2 (m/s)	7.20	0.50	2.07	0.84	1.21	0.41
	R_h (%)	100.00	13.00	46.27	15.68	0.38	0.34
	PMF − 56 ET_0 (mm/day)	11.62	0.16	3.30	2.30	0.55	0.70

Std., standard deviation; SK, skewness; CV, coefficient of variation.

variation intensity of the rest of parameters is intermediate (CV is bounded between 0.25 and 0.75). Additionally, there are no significant differences for other data between the two weather stations.

Model development

The selection of appropriate input variables has a direct impact on the performance of the model; moreover, finding suitable inputs can provide an efficient way of estimating ET_0 for many regions where weather data are not always available. For the development of the RF model, this study selected different combinations of various daily climatic data as input, and ET_0 computed by daily PMF-56 equation as output for training and testing the models. Eight different combinations were considered in the present study and are referred to in the short form as shown in Table 2. Temperature is the most influential variable on ET_0 and predominant physical factor in the evaporation process (Jain *et al.* 2008; Wen *et al.* 2015). Thus, combination 1, as the base inputs, consists of T_{max} and T_{min} and the other combinations are formed by integrating S_{un}; R_h and U_2; R_h, U_2 and S_{un}; R_h, U_2, and S_{un} into combination 1, respectively. Each combination was trained and tested by the RF model.

In order to eliminate the influence of the dimension, the input and output data were normalized to obtain data with a mean of 0 and a variance of 1 before running models; the equation is used as follows:

$$x_{new} = (x - \mu)/\sigma \tag{9}$$

where χ_{new} is the normalized dimensionless data, μ is the average data and σ is the standard deviation.

Table 2 | Input combinations of RF models used in the study

Input combination	Model	Inputs
Combination 1	RF1	T_{max}, T_{min}
Combination 2	RF2	T_{max}, T_{min}, S_{un}
Combination 3	RF3	T_{max}, T_{min}, U_2
Combination 4	RF4	T_{max}, T_{min}, R_h
Combination 5	RF5	T_{max}, T_{min}, S_{un}, U_2
Combination 6	RF6	T_{max}, T_{min}, S_{un}, Rh
Combination 7	RF7	T_{max}, T_{min}, U_2, R_h
Combination 8	RF8	T_{max}, T_{min}, S_{un}, U_2, R_h

Models' performance criteria

For the assessment of the performances of the RF model, statistical indices such as coefficient of correlation (r), root mean squared error (RMSE), mean absolute error (MAE), and Nash–Sutcliffe efficiency coefficient (NS) were applied in this research. r measures the correlation between estimated and observed values; the smaller the differences between r and 1.0 are, the stronger the correlation is. RMSE and MAE provide different types of information about the measurement of the prediction capability of the models. RMSE demonstrates the goodness-of-fit relevant to high values whereas MAE yields a more balanced perspective of the goodness-of-fit at moderate values (Citakoglu *et al.* 2014). The small RMSE and MAE values indicate that the error between the estimated and calculated values is small and the performance of the models is good. The r, RMSE, MAE, and NS are computed by the following equations:

$$r = \frac{\sum_{i=1}^{n}(E^p(i) - \overline{E^p(i)})(E^o(i) - \overline{E^o(i)})}{\sqrt{\sum_{i=1}^{n}(E^p(i) - \overline{E^p(i)})^2 (E^o(i) - \overline{E^o(i)})^2}} \tag{10}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(E^p(i) - E^o(i))^2}{n}} \tag{11}$$

$$\text{MAE} = \frac{\sum_{i=1}^{n}|(E^p(i) - E^o(i))/E^o(i)|}{n} \tag{12}$$

$$\text{NS} = 1 - \frac{\sum_{i=1}^{n}(E^o(i) - E^p(i))^2}{\sum_{i=1}^{n}(E^o(i) - \overline{E^o(i)})^2} \tag{13}$$

where $E^p(i)$ and $E^o(i)$ are the ith ET_0 values computed through different models and PMF-56 equation, respectively; $\overline{E^p(i)}$ and $\overline{E^o(i)}$ are the average of $E^p(i)$ and $E^o(i)$; and n is the number of data. In terms of these metrics, the model is denoted as a perfect fit when $r = 1$, RMSE and MAE $= 0$, and NS $= 1$, respectively.

RESULTS AND DISCUSSION

Model performance

The performance of RF model for PMF-56 ET_0 applied to the studied stations for the training and testing periods are summarized in Tables 3 and 4, which demonstrate the precision of the proposed RF model by the formulae of *r*, RMSE, MAE, and NS. As can be seen, there were no significant changes in respect to all of the metrics of these models in training as well as testing periods. This research selected the criteria during the testing phase to compare the capabilities of these models in the prediction of PMF-56 ET_0 and all of the following analyses were performed in the testing period.

Considering all models for the two stations, it can be observed that the RF8 model outperforms all of the other models in the four-estimation norm, with the highest *r* and NS values as well as the lowest RMSE and MAE values. We also clearly see all of the *r* and NS values surpass 0.8 for both stations, indicating the performances of RF models in PMF-56 ET_0 prediction were encouraging. Thus, it was selected as the best-fit model for estimating the PMF-56 ET_0 at the two stations. In the remaining models, RF5, RF6, and RF7 models including four parameters as inputs had a higher *r* and NS values, lower RMSE and MAE values and were found to be better than RF2, RF3, and RF4 models with three input parameters at each individual site. Alternatively, RF1 with only T_{max} and T_{min} as inputs had the biggest errors rates compared to other models. This demonstrated that the performance of the models relied on the number of input parameters. However, weather factors were usually incomplete in data-limited regions, especially in arid environments, such as northwest China. The

Table 3 | Performance analysis of the RF models at Zhangye station during the training and testing periods

	Training				Testing			
Models	***r***	**RMSE**	**MAE**	**NS**	***r***	**RMSE**	**MAE**	**NS**
RF1	0.962	0.705	0.480	0.923	0.909	1.019	0.747	0.823
RF2	0.970	0.621	0.404	0.940	0.946	0.795	0.570	0.893
RF3	0.975	0.571	0.387	0.949	0.934	0.873	0.634	0.871
RF4	0.988	0.395	0.277	0.976	0.965	0.638	0.453	0.931
RF5	0.985	0.451	0.287	0.968	0.964	0.645	0.472	0.929
RF6	0.992	0.319	0.213	0.984	0.973	0.561	0.394	0.947
RF7	0.995	0.256	0.177	0.990	0.978	0.509	0.361	0.956
RF8	0.996	0.238	0.156	0.991	0.990	0.339	0.255	0.981

Table 4 | Performance analysis of the RF models at Gaotai station during the training and testing periods

	Training				Testing			
Models	***r***	**RMSE**	**MAE**	**NS**	***r***	**RMSE**	**MAE**	**NS**
RF1	0.967	0.594	0.412	0.935	0.907	0.967	0.716	0.814
RF2	0.980	0.474	0.329	0.958	0.951	0.694	0.524	0.904
RF3	0.982	0.448	0.320	0.963	0.952	0.689	0.510	0.905
RF4	0.986	0.394	0.270	0.971	0.951	0.698	0.505	0.903
RF5	0.989	0.334	0.231	0.979	0.974	0.515	0.389	0.947
RF6	0.991	0.319	0.213	0.981	0.968	0.562	0.416	0.937
RF7	0.993	0.271	0.185	0.986	0.971	0.540	0.389	0.942
RF8	0.996	0.206	0.142	0.992	0.987	0.352	0.267	0.975

selection of the model should be decided according to the available meteorological parameters. The models whose input comprised T_{max} and T_{min} are needed and can be used in this study for practical application.

Concretely, in terms of Zhangye station, the corporation of R_h can significantly improve the performance of RF models. Adding R_h to temperature-based inputs, the RF4 model improved r, RMSE, MAE, and NS by 6.2%, 37.4%, 39.4%, and 13.1%, respectively. Likewise, the RF6 and RF7 models introducing R_h as input variable achieved higher simulation precision (with the higher r and NS values, the lower RMSE and MAE values) than the RF5 model with the absence of R_h. From these results, it was shown that the addition of R_h was more sensitive to output relative to the S_{un} and U_2. The RF2 model performed the second best in ET_0 estimation among the RF2, RF3, and RF4 models. Note that RF3 including U_2 on the basis of RF1 improved r, RMSE, MAE, and NS value by 4.1%, 22%, 23.7%, and 8.5%, respectively. It was an objective fact that RF6 was superior to RF5 according to four evaluation criteria. This result showed that inserting R_h was more effective than U_2 to the estimation of ET_0. The RF3 model, whose inputs included T_{max}, T_{min}, and U_2 were found to be worse than RF2 and RF4 models among the three models. As a result, the ET_0 is most easily affected by R_h, followed by S_{un} and U_2. This conclusion is in disagreement with the findings of many studies (Dai *et al.* 2009; Petkovic *et al.* 2015; Tao *et al.* 2015; Xing *et al.* 2016), namely, the S_{un} is considered as the most effective parameter for simulating PMF-56 ET_0. Generally, the results depend on the selected geographical location and climate type of the study area.

For the case of Gaotai station, there were different results compared with those of Zhangye station. It was shown that the RF3 model performed slightly better than RF2 and RF4 models in terms of four statistical indicators, and it can be stated that PMF-56 ET_0 was easily influenced by U_2. This was also confirmed by RF5 and RF7 models with the insertion of U_2 into the inputs presented in Tables 3 and 4. The RF5 and RF7 models remarkably increased the r and NS values of 0.6% and 1.1%, and 0.3% and 0.5%, respectively, and decreased the RMSE and MAE values of 8.4% and 6.5%, and 3.9% and 6.5%, respectively, relative to the RF6 model, exhibiting the superiority of RF5 and RF7 models to the RF6 model significantly. The results of this comparison revealed that integrating U_2 improved the accuracy of the model significantly. Accordingly, adding U_2 is found to be more influential than S_{un} and R_h on ET_0 simulation, which is the same outcome obtained by Traore *et al.* (2010) and Karimaldini *et al.* (2012). It is observed that the input scenarios listed in Table 2 have a distinct performance for the two stations due to the different geographical locations.

To compare the performance of the temperature-based models and the other models with the absence of temperature, the Supplementary material lists performance statistics of four input combinations including: (1) S_{un} and U_2; (2) S_{un} and R_h; (3) U_2 and R_h; (4) S_{un}, U_2 and R_h, and the four inputs are expressed as RF9, RF10, RF11, and RF12, respectively. It is apparent that all the RF models produced higher RMSE (more than 1.38) and MAE (more than 1.08) as well as lower r (less than 0.81) and NS (less than 0.66), and were inferior to combinations 1–8 inserting T_{max} and T_{min} into inputs for PMF-56 ET_0 forecasting (Supplementary material, Tables A1 and A2). As the best fitting models, RF12 had r values of 0.809 and 0.790, RMSE values of 1.427 and 1.382, MAE values of 1.097 and 1.082, and NS values of 0.654 and 0.619 for Zhangye and Gaotai stations, respectively, which cannot meet the prediction standards of PMF-56 ET_0. In such circumstances, RF9–RF12 models should not be selected as techniques to estimate PMF-56 ET_0. Therefore, the following does not elaborate on the four models, but mainly focuses on RF1–RF8 models.

Figures 3–6 exhibit the hydrograph and scatter plots of the ET_0 values computed by the PMF-56 equation and the values estimated by different combinations of the RF model of the validation period for the two stations. A total of eight combinations of RF model displayed a good prediction of ET_0. In addition, it is obviously seen that the ET_0 values estimated by the RF8 model were closer to the PMF-56 ET_0 values and followed the same trend than the other models while the RF1 model performed the worst in this area. From the fit line with the form of $y = ax + b$, the coefficients a and b of the RF8 model were closer to 1 and 0 than the other models, because the lowest values of b (equal to 0), and the highest values of the slope (equal to 1) denote the best fit of models. These were confirmed by r, RSME, MAE, and NS values shown in Tables 3 and 4. As well, it was observed that the fitting performance of the maximum and minimum PMF-56 ET_0 was not very good, especially that of peaks of the first few models.

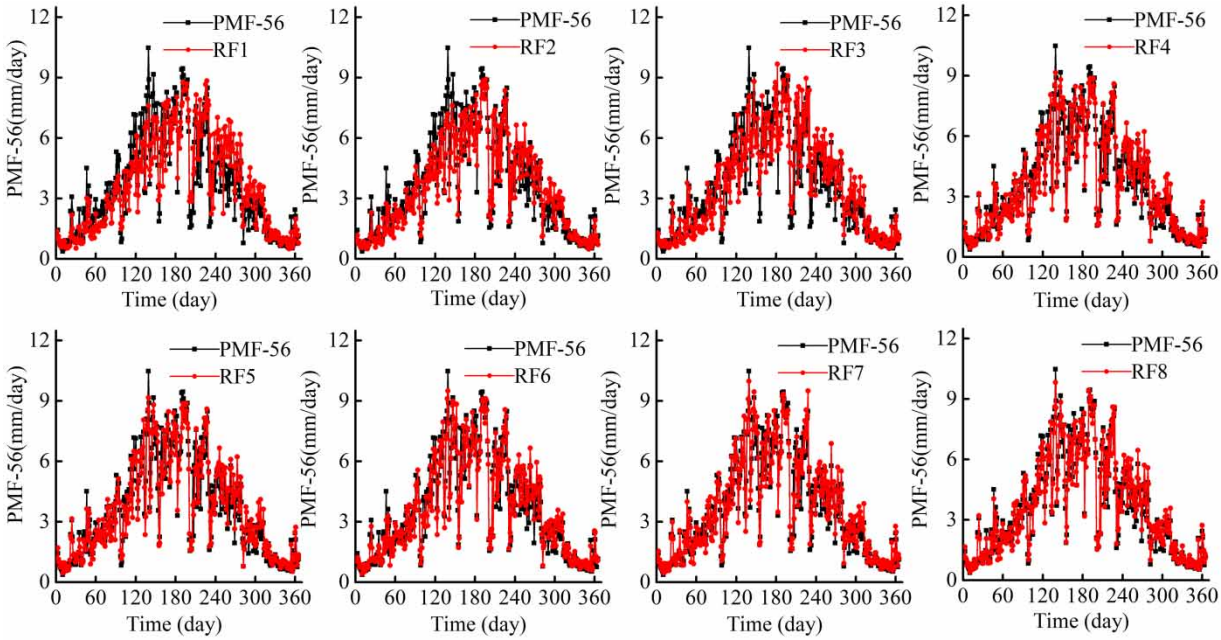


Figure 3 | Comparison of the ET_0 values estimated by the PMF-56 equation and the RF models for Zhangye station during the testing period.

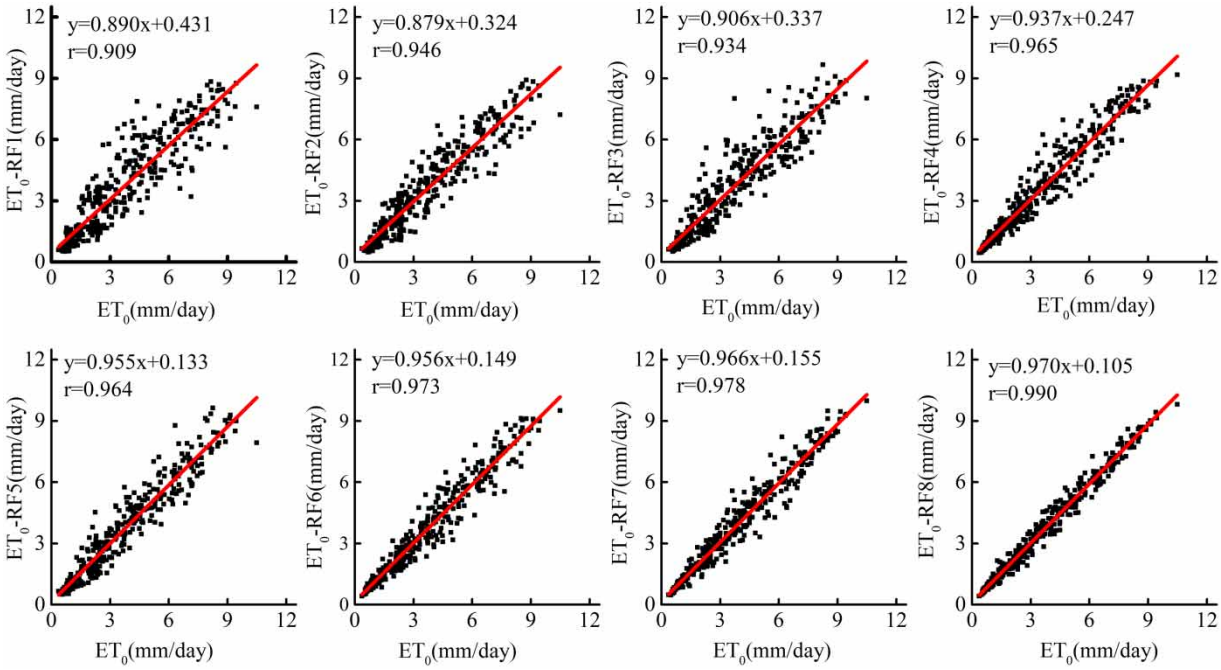


Figure 4 | Relationship between ET_0 values estimated by the PMF-56 equation and the RF models for Zhangye station during the testing period.

Due to the importance of PMF-56 ET_0 in irrigation and agricultural water use, water resources planning and management, the estimation of total PMF-56 ET_0 obtained by different combinations of RF model was also considered in this paper. The total ET_0 amounts calculated by PMF-56 and RF models in the testing phase are given in Table 5. It

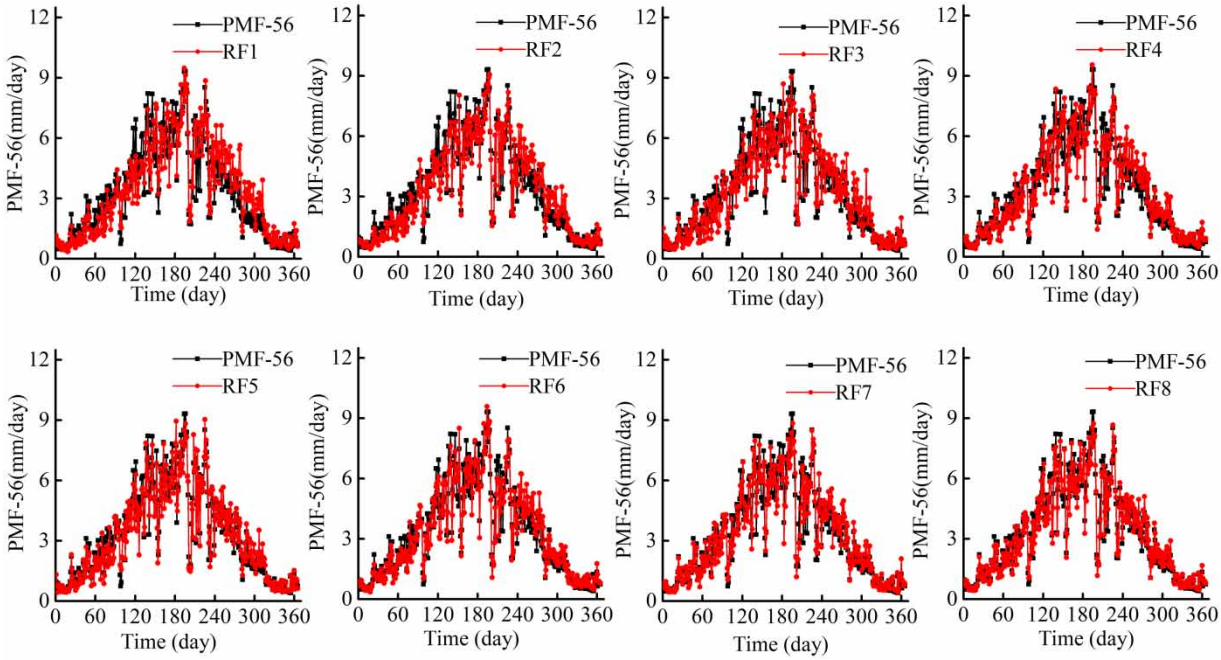


Figure 5 | Comparison of the ET_0 values estimated by the PMF-56 equation and the RF models for Gaotai station during the testing period.

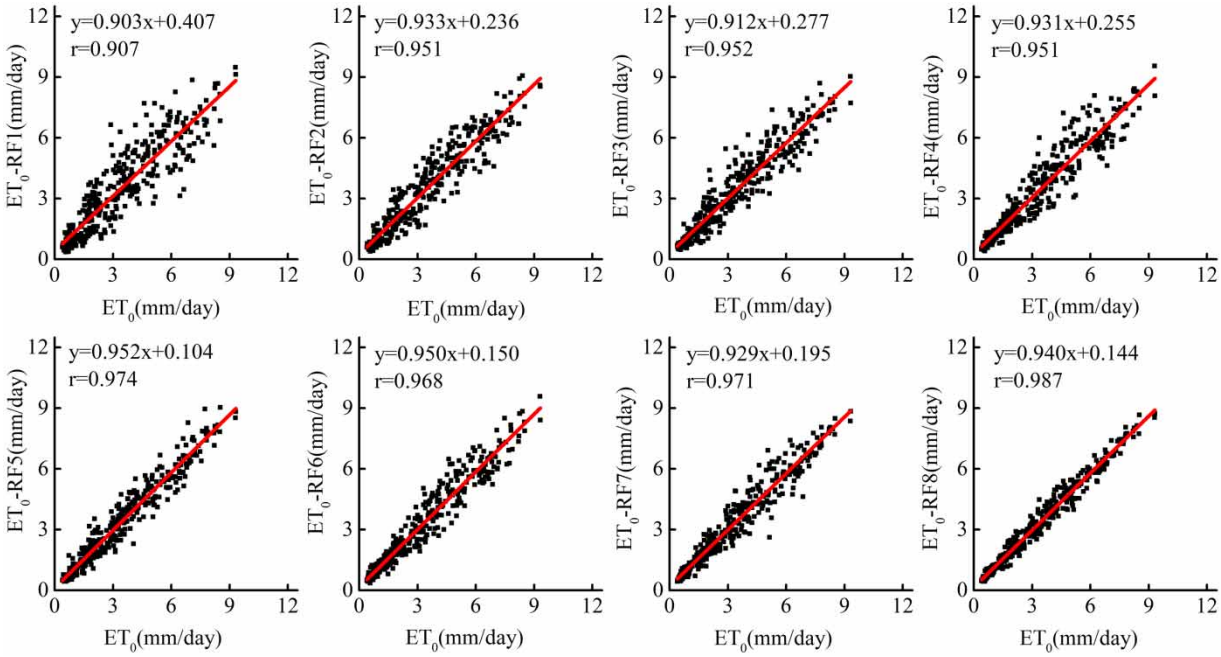


Figure 6 | Relationship between ET_0 values estimated by the PMF-56 equation and the RF models for Gaotai station during the testing period.

is remarkable that all models had a quite good estimation of total PMF-56 ET_0 value since there was a smaller relative error (all values less than 3.5%) for both sites, especially the RF1 model, whose input parameters were only T_{max} and T_{min} at Zhangye station, with a relative error of −0.2%. In addition, noting the fact that the RF8 model

Table 5 | Total ET_0 values and relative error calculated by different combinations of RF models during the testing period

	Zhangye		Gaotai	
Input	Total ET_0 (mm)	relative error (%)	Total ET_0 (mm)	relative error (%)
PMF-56	1,332.40	–	1,208.60	–
RF1	1,329.65	−0.2	1,239.70	2.6
RF2	1,289.10	−3.2	1,213.55	0.4
RF3	1,329.98	−0.2	1,202.84	−0.5
RF4	1,338.79	0.5	1,218.12	0.8
RF5	1,320.52	−0.9	1,188.09	−1.7
RF6	1,327.80	−0.3	1,203.06	−0.5
RF7	1,343.31	0.8	1,194.60	−1.2
RF8	1,330.26	−0.2	1,188.33	−1.6

with all the variables as inputs did not have the lowest relative error (−1.6%) among all the models at Gaotai station, it still performed well and its value fell within the reliable range. Although generally, reliable weather data sets such as R_h, U_2 and S_{un} are limited in the arid regions, combining the above calculation results and demands of practical use, the RF model can be employed to predict PMF-56 ET_0 where restricted data are available.

Evaluation of the importance of variables

Index importance assessment is an advantage of the RF model which can directly obtain an order of all of the weather parameters. As shown in Figure 7, temperature is the most relevant variable connected with the estimation of PMF-56 ET_0 in the two stations, with IncNodePurity values individually accounting for 69% and 71%, which illustrates that T_{max} and T_{min} can be employed to predict PMF-56 ET_0 combined with Tables 3 and 4; on the contrary, the RF model cannot simulate ET_0 with higher accuracy when temperature is missing. In addition, for the importance degree of the other three factors there existed a similarity at the two sites; R_h can be considered as an influential index due to higher IncNodePurity values at Zhangye and Gaotai stations. It is relevant to note that this result is inconsistent with the consequence obtained at Gaotai station (described by Table 4). The aforementioned outcomes are achieved by different combinations among all parameters, indicating that inserting U_2 into inputs has higher precision compared to adding other variables for Gaotai site. Nevertheless, IncNodePurity value of random forest explains each index's contribution to ET_0, therefore, in terms of importance of variables, there is no doubt that R_h is the most relavant factor affecting ET_0 in this area. Similar results were also carried out in Shiyang River Basin, northwest China by Huo *et al.* (2012), where R_h has a large effect on daily PMF-56 ET_0 except for air temperature in an arid region, northwest China.

Uncertainty analysis

The techniques of Monte Carlo simulations were used to corroborate the applicability of RF models in modeling

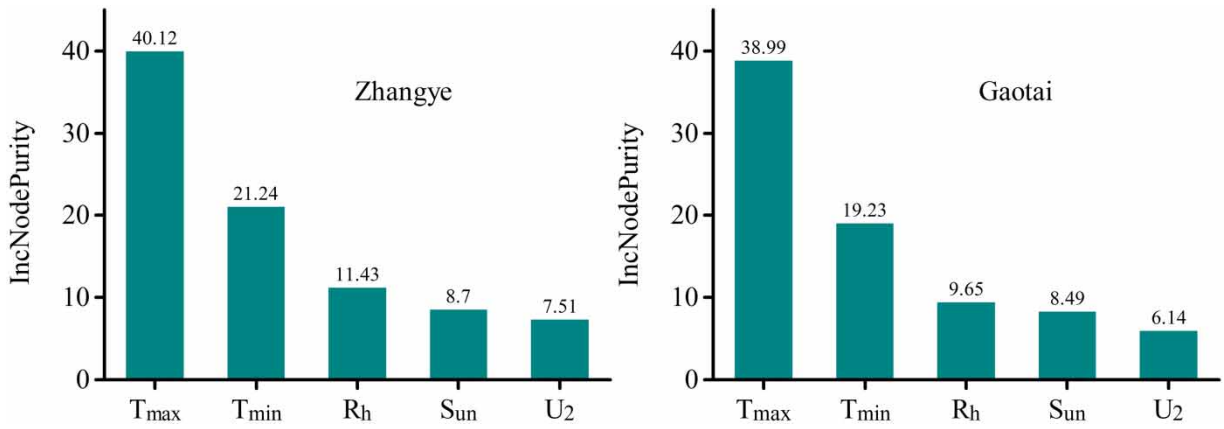


Figure 7 | Importance degree of evaluation indicators for the two stations.

PMF-56 ET_0, an important hydro-meteorological parameter for agriculture, ecosystems, and several other socio-economic activities. The method proposed here has been used to quantify the uncertainty by predicting the confidence intervals of the simulation results. Figures 8 and 9 illustrate 95% confidence intervals for the estimates of daily PMF-56 ET_0 applying the RF model for Zhangye and Gaotai stations during the testing period. From these two figures, we find that there was a good match between 95% confidence intervals and results obtained by the RF model, and most of the observed ET_0 data lay within the confidence intervals at the two stations. Results of Monte Carlo analysis of the RF model for the two stations are given in the upper left corner of the panels. In general, satisfactory results indicate that more observed data were bracketed within the 95PPU (all values are over 60%), while a lower *d-factor* value can be obtained (*d-factor* values less than 1 are considered appropriate). Remarkably, the RF8 model produced acceptable *d-factor* values, noting that the *d-factor* values were 0.26 and 0.27 at Zhangye and Gaotai station, respectively. Also, the PMF-56 ET_0 values bracketed by 95PPU were more significant for both stations; it was observed that 91% and 87% of the PMF-56 ET_0 data were bracketed by the 95PPU at Zhangye and Gaotai station, respectively. In addition, 95%

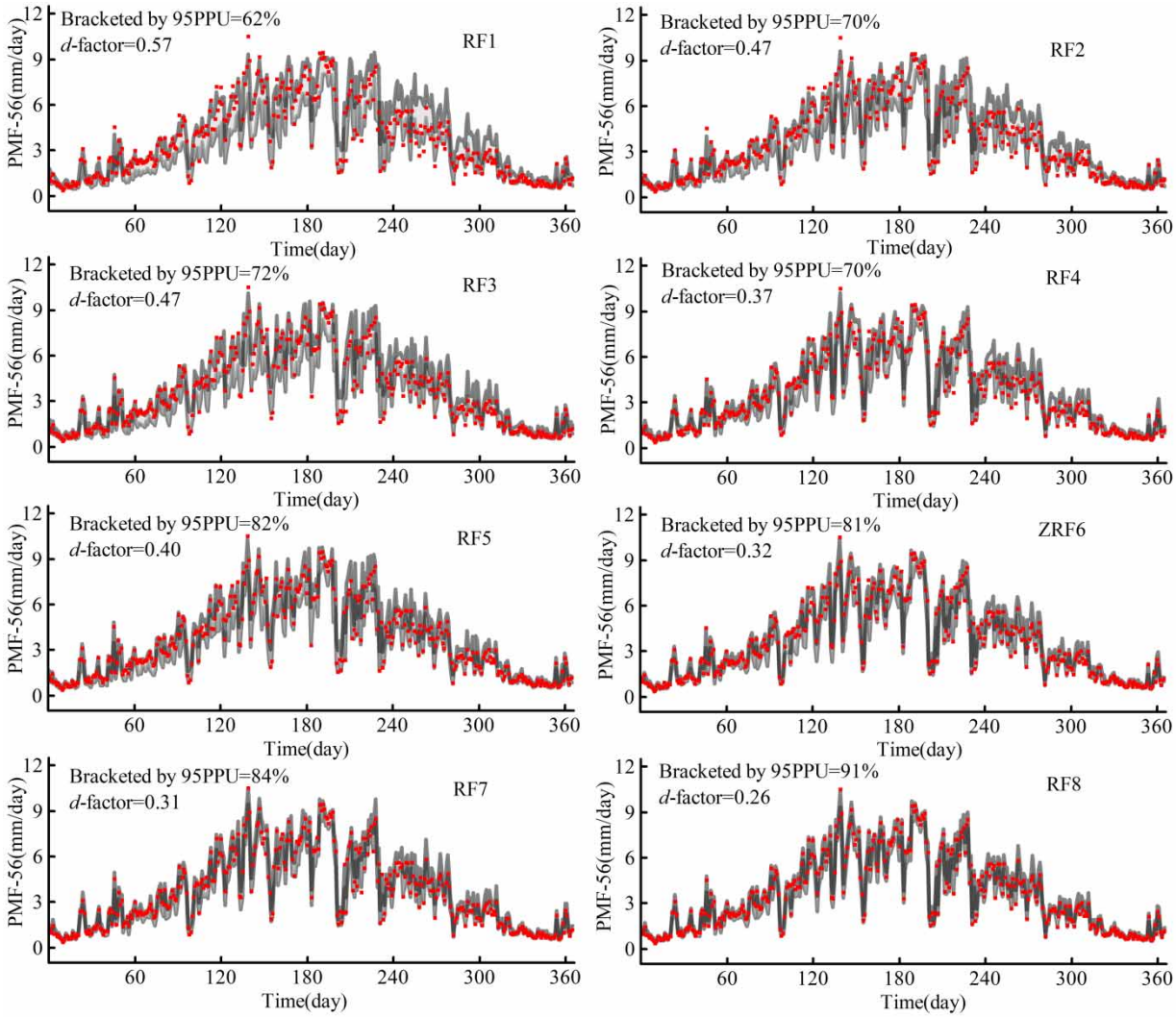


Figure 8 | The ET_0 values estimated by the PMF-56 equation and the 95% confidence intervals estimated by Monte Carlo simulation of RF models with randomly sampled input vectors in Zhangye station.

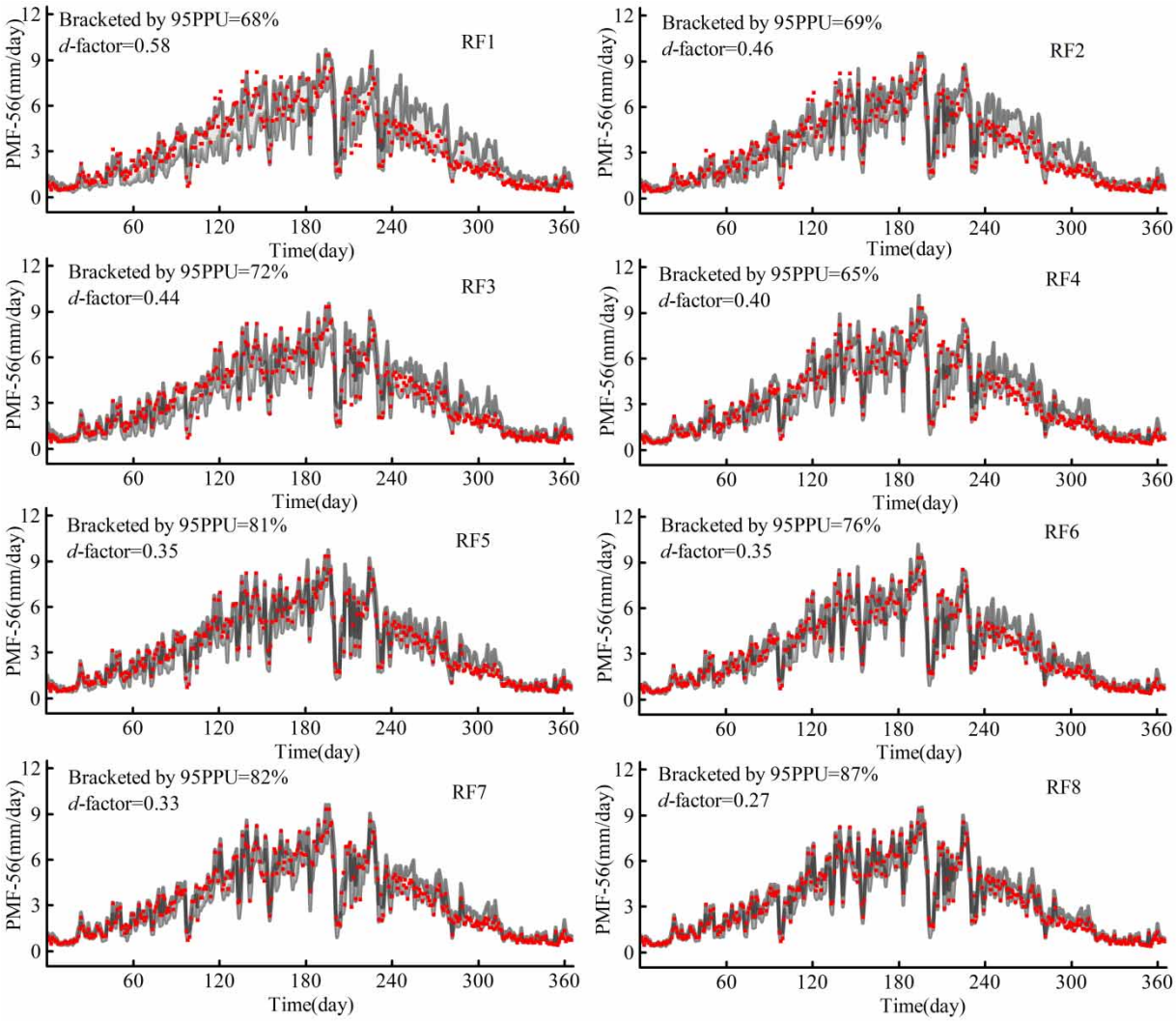


Figure 9 | The ET_0 values estimated by the PMF-56 equation and the 95% confidence intervals estimated by Monte Carlo simulation of RF models with randomly sampled input vectors in Gaotai station.

confidence intervals of RF3 to RF8 models were found to be a relatively good fit to the peak of PMF-56 ET_0 for the two sites. Although the RF1 model has wider 95% confidence and higher *d-factor* values than other combinations for the two stations, the uncertainty is still within acceptable limits. As can be seen from these figures, the trend of 95% confidence intervals calculated by the RF1 model is basically close to that of PMF-56 ET_0. The values bracketed by 95PPU were more than 62%, and the *d-factor* values were less than 0.58, indicating that the RF1 model is able to predict daily PMF-56 ET_0 with smaller uncertainties. Considering the purpose of this paper and the above discussion, the prediction uncertainties of the RF1 model were determined to further illustrate the application of this model and therefore find a reliable model in an arid area with a lack of sufficient meteorological data. Besides, it is worth noting that the maximum and minimum ET_0 values cannot be simulated perfectly by models, consistent with the previous conclusion (as shown in Figures 3–6), which also further illustrates the differences between the total ET_0 amounts computed by PMF-56 and RF models as displayed by Table 5. In spite of some errors, taking the discussion of the section on evaluation of the importance of variables into account, we find that the RF model with

only T_{max} and T_{min} as inputs is still considered as an appropriate technique to simulate daily PMF-56 ET_0 in arid conditions.

CONCLUSIONS

Water resources play an essential role in arid environments, so new modeling and water assessment methods are crucial for maintaining sustainability of water resources, strategies for water quality and usage. ET_0 provides a vital parameter of water resources calculation, regional water resources management, and irrigation plan development. This research discussed the performance of the RF model to predict PMF-56 ET_0 using different combinations of daily climatic data, including maximum air temperature (T_{max}), minimum air temperature (T_{min}), sunshine duration (S_{un}), wind speed (U_2), and relative humidity (R_h) for Zhangye and Gaotai stations, in an arid region, northwest China. It was found that the precision of the models was respectively improved when adding S_{un}, U_2, and R_h into the temperature-based model. Moreover, the importance evaluation of indices indicated PMF-56 ET_0 was more readily influenced by R_h with the exception of air temperature in this region. The best performance was achieved by the RF8 model with all the meteorological arguments as inputs. Although the precision of the model depends on the number of input climatic variables, all of the combinations of RF model turned out to be capable of producing reliable precision in ET_0 modeling, as mentioned above. Thus, the RF model should be the recommended model for PMF-56 ET_0 modeling in arid regions where weather data are limited. The Monte Carlo simulation technique was also employed for quantifying RF model uncertainty. The results of uncertainty analysis indicated that the PMF-56 ET_0 values bracketed by 95% confidence interval (95PPU) were larger, namely, most of the PMF-56 ET_0 values fell within 95PPU, and *d-factor* values calculated by upper and lower limits of the confidence interval were smaller in the studied area. In summary, the RF model is considered as an appropriate way of forecasting PMF-56 ET_0 and can provide an alternative tool under a minimal amount of climate data, as well as a reference for water resources management.

It should be noted that there have been extensive studies comparing the performance of RF and other artificial intelligence models for simulating ET_0, regarding the fact that almost all studies confirmed that RF model achieved higher simulation precision than others. Under these circumstances, this article did not conduct comparative research. In addition, although the RF model provides significant potential for more accurate estimation of the ET_0 with a lack of appropriate weather data in arid regions, certain drawbacks still persist. In this investigation, the selected sites are insufficient and the amounts of data size used to develop the model are smaller; besides, maximum and minimum ET_0 values simulated by RF model cannot accurately reflect the observed data. Therefore, further study can potentially focus on choosing more studied points with different climate types and combining RF method and another algorithm, such as Kalman filtering and wavelet transform techniques, for obtaining more reliable and practical results.

ACKNOWLEDGEMENTS

This research was funded by the National Key Research and Development Program of China (2017YFC0404305), Key Research Program of Frontier Sciences, CAS (QYZDJ-SSW-DQC031), Natural Science Foundation of Gansu province, China (18JR4RA002,18JR3RA393), CAS 'Light of West China' Program and the USQ-CAS collaborative research agreement USQ943692018 (2019–2021). The authors thank all the anonymous reviewers for their constructive comments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this paper is available online at https://dx.doi.org/10.2166/nh.2020.012.

REFERENCES

Abdullah, S. S., Malek, M. A., Mustapha, A. & Aryanfar, A. 2014 Hybrid of artificial neural network-genetic algorithm for

prediction of reference evapotranspiration (ET_0) in arid and semiarid regions. *J. Agr. Sci.* **6** (3), 191–200.

Abdullah, S. S., Malek, M. A., Abdullah, N. S., Kisi, O. & Yap, K. S. 2015 Extreme learning machines: a new approach for prediction of reference evapotranspiration. *J. Hydrol.* **527**, 184–195.

Acharya, N., Shrivastava, N. A., Panigrahi, B. K. & Mohanty, U. C. 2013 Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: an application of extreme learning machine. *Clim. Dyn.* **43**, 1303–1310.

Allen, R. G., Pereira, L. S., Raes, D. & Smith, M. 1998 Crop evapotranspiration-guidelines for computing crop water requirement. In *Food and Agricultural Organisation of the United Nations*, Rome, Italy.

Antanasijević, D., Pocajt, V., Perić-Grujić, A. & Ristić, M. 2014 Modelling of dissolved oxygen in the Danube river using artificial neural networks and Monte Carlo simulation uncertainty analysis. *J. Hydrol.* **519**, 1895–1907.

Benli, B., Kodal, S., Ilbeyi, A. & Ustun, H. 2006 Determination of evapotranspiration and basal crop coefficient of alfalfa with a weighing lysimeter. *Agric. Water Manage.* **81**, 358–370.

Breiman, L. 2001 Random forests. *Mach. Learn.* **45**, 5–32. doi:10.1023/A:1010933404324.

Citakoglu, H., Cobaner, M., Haktanir, T. & Kisi, O. 2014 Estimation of monthly mean reference evapotranspiration in Turkey. *Water Resour. Manage.* **28**, 99–113.

Cutler, R. D., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J. & Lawler, J. J. 2007 Random forests for classification in ecology. *Ecology* **88**, 2783–2792.

Dai, X. Q., Shi, H., Li, Y. S., Zhu, O. Y. & Huo, Z. L. 2009 Artificial neural network models for estimating regional reference evapotranspiration based on climate factors. *Hydrol. Process.* **23**, 442–450.

Dehghani, M., Saghafian, B., Saleh, F. N., Farokhnia, A. & Noori, R. 2014 Uncertainty analysis of streamflow drought forecast using artificial neural networks and Monte-Carlo simulation. *Int. J. Climatol.* **34**, 1169–1180.

Deo, R. C. & Şahin, M. 2015 Application of the extreme learning machine algorithm for the prediction of monthly effective drought index in eastern Australia. *Atmos. Res.* **153**, 512–525.

Dong, L. J., Li, X. B. & Peng, K. 2013 Prediction of rockburst classification using Random Forest. *T. Nonferr. Metal. Soc.* **23** (2), 472–477.

Feng, Y., Cui, N. B., Gong, D. Z., Zhang, Q. W. & Zhao, L. 2017 Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modeling. *Agric. Water Manage.* **193**, 163–173.

Ferreira, L. B., Cunha, F. F., Oliveira, R. A. & Filho, E. I. F. 2019 Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM – A new approach. *J. Hydrol.* **572**, 556–570. https://doi.org/10.1016/j.jhydrol.2019.03.028.

Fukuda, S., Spreer, W., Yasunaga, E., Yuge, K., Sardsud, V. & Müller, J. 2013 Random Forests modelling for the estimation of mango (Mangifera indica L. cv. Chok Anan) fruit yields under different irrigation regimes. *Agric. Water Manage.* **116**, 142–150.

Gao, M., Yin, L. & Ning, J. C. 2018 Artificial neural network model for ozone concentration estimation and Monte Carlo analysis. *Atmos. Environ.* **184**, 129–139.

Ghorbani, M. A., Zadeh, H. A., Isazadeh, M. & Terzi, O. 2016 A comparative study of artificial neural network (MLP, RBF) and support vector machine models for river flow prediction. *Environ. Earth Sci.* **75**. doi:10.1007/s12665-015-5096-x.

Gislason, P. O., Benediktsson, J. A. & Sveinsson, J. R. 2006 Random Forests for land cover classification. *Pattern Recogn. Lett.* **27** (4), 294–300.

Gong, H. R., Sun, Y. R., Shu, X. & Huang, B. S. 2018 Use of random forests regression for predicting IRI of asphalt pavements. *Constr. Build. Mater.* **189**, 890–897.

Granata, F. 2019 Evapotranspiration evaluation models based on machine learning algorithms – a comparative study. *Agric. Water Manage.* **217**, 303–315.

He, Z. B., Wen, X. H., Liu, H. & Du, J. 2014 A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *J. Hydrol.* **509**, 379–386.

Heung, B., Bulmer, C. E. & Schmidt, M. G. 2014 Predictive soil parent material mapping at a regional-scale: a random forest approach. *Geoderma* **214–215**, 141–154.

Huang, G. M., Wu, L. F., Ma, X., Zhang, W. Q., Fan, J. L., Yu, X., Zeng, W. Z. & Zhou, H. M. 2019 Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions. *J. Hydrol.* **574**, 1029–1041. https://doi.org/10.1016/j.jhydrol.2019.04.085.

Huo, Z., Feng, S., Kang, S. & Dai, X. 2012 Artificial neural network models for reference evapotranspiration in an arid area of northwest China. *J. Arid. Environ.* **82**, 81–90.

Jain, S. K., Nayak, P. C. & Sudheer, K. P. 2008 Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. *Hydrol. Process.* **22**, 2225–2234.

Karimaldini, F., Shui, L. T., Mohamed, T. A., Abdollahi, M. & Khalili, N. 2012 Daily evapotranspiration modelling from limited weather data by using neuro-fuzzy computing technique. *J. Irrig. Drain. Eng.* **138** (1), 21–34.

Kisi, O. & Cimen, M. 2009 Evapotranspiration modeling using support vectormachines. *Hydrol. Sci. J.* **54** (5), 918–928.

Noori, R., Hoshyaripour, G., Ashrafi, K. & Araabi, B. N. 2010 Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration. *Atmos. Environ.* **44**, 476–482.

Nourani, V., Elkiran, G. & Abdullahi, J. 2019 Multi-station artificial intelligence based ensemble modeling of reference evapotranspiration using pan evaporation measurements. *J. Hydrol.* **577**, 123958.

Nourani, V., Elkiran, G. & Abdullahi, J. 2020 Multi-step ahead modeling of reference evapotranspiration using a multi-model approach. *J. Hydrol.* **581**, 124434.

Patil, A. P. & Deka, P. C. 2016 An extreme learning machine approach for modeling evapotranspiration using extrinsic inputs. *Comput. Electron. Agric.* **121**, 385–392.

Petkovic, D., Gocic, M., Trajkovic, S., Shamshirband, S., Motamedi, S., Hashim, R. & Bonakdari, H. 2015 Determination of the most influential weather parameters on reference evapotranspiration by adaptive neuro-fuzzy methodology. *Comput. Electron. Agric.* **114**, 277–284.

Rawat, K. S., Singh, S. K., Bala, A. & Szabó, S. 2019 Estimation of crop evapotranspiration through spatial distributed crop coefficient in a semi-arid environment. *Agric. Water Manage.* **213**, 922–933.

Rodriguez-Galiano, V., Mendes, M. P., Garcia-Soldado, M. J., Chica-Olmo, M. & Ribeiro, L. 2014 Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: a case study in an agricultural setting (Southern Spain). *Sci. Total. Environ.* **476–477**, 189–206.

Shiri, J., Nazemi, A. H., Sadraddini, A. A. & Landeras, G. 2014 Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. *Comput. Electron. Agric.* **108**, 230–241.

Shrestha, D. L., Kayastha, N. & Solomatine, D. P. 2009 A novel approach to parameter uncertainty analysis of hydrological models using neural networks. *Hydrol. Earth Syst. Sci.* **13**, 1235–1248.

Tabari, H., Kisi, O., Ezani, A. & Talaee, P. H. 2012 SVM, ANFIS, regression and climatebased models for reference evapotranspiration modeling using limitedclimatic data in a semi-arid highland environment. *J. Hydrol.* **444–445**, 78–89.

Tao, X. E., Chen, H., Xu, C. Y., Hou, Y. K. & Jie, M. X. 2015 Analysis and prediction of reference evapotranspiration with climate change in Xiangjiang River Basin, China. *Water Sci. Eng.* **8** (4), 273–281.

Torres, A. F., Walker, W. R. & McKee, M. 2011 Forecasting daily potential evapotranspiration using machine learning and limited climatic data. *Agric. Water Manage.* **98** (4), 553–562.

Traore, S., Wang, Y. M. & Kerh, T. 2010 Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone. *Agric. Water Manage.* **97** (5), 707–714.

Wang, Z. L., Lai, C. J., Chen, X. H., Yang, B., Zhao, S. W. & Bai, X. Y. 2015 Flood hazard risk assessment model based on random forest. *J. Hydrol.* **527**, 1130–1141.

Wen, X. H., Si, J. H., He, Z. B., Wu, J., Shao, H. B. & Yu, H. J. 2015 Support-vector-machine-based models for modeling daily reference evapotranspiration with limited climatic data in extreme arid regions. *Water Resour. Manage.* **29** (9), 3195–3209.

Xing, X. G., Liu, Y., Zhao, W. G., Kang, D. J., Yu, M. & Ma, X. Y. 2016 Determination of dominant weather parameters on reference evapotranspiration by path analysis theory. *Comput. Electron. Agric.* **120**, 10–16.

Yassin, M. A., Alazba, A. A. & Mattar, M. A. 2016 Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. *Agric. Water Manage.* **163**, 110–124.

Yin, Z. L., Wen, X. H., Feng, Q., He, Z. B., Zou, S. B. & Yang, L. S. 2016 Integrating genetic algorithm and support vector machine for modeling daily reference evapotranspiration in a semi-arid mountain area. *Hydrol. Res.* **48**. doi:10.2166/nh.2016.205.

Yoon, H., Jun, S. C., Hyun, Y., Bae, G. O. & Lee, K. K. 2011 A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *J. Hydrol.* **396** (1–2), 128–138.

First received 17 January 2020; accepted in revised form 3 April 2020. Available online 3 June 2020