

## RESEARCH ARTICLE

# Comparison of neuron-based, kernel-based, tree-based and curve-based machine learning models for predicting daily reference evapotranspiration

Lifeng Wu<sup>1</sup>, Junliang Fan<sup>2\*</sup>

**1** School of Hydraulic and Ecological Engineering, Nanchang Institute of Technology, Nanchang, China, **2** College of Water Resources and Architectural Engineering, Northwest A&F University, Yangling, China

\* [nwwfjl@163.com](mailto:nwwfjl@163.com)

## OPEN ACCESS

**Citation:** Wu L, Fan J (2019) Comparison of neuron-based, kernel-based, tree-based and curve-based machine learning models for predicting daily reference evapotranspiration. PLoS ONE 14(5): e0217520. <https://doi.org/10.1371/journal.pone.0217520>

**Editor:** Paweł Pławiak, Politechnika Krakowska im Tadeusza Kosciuszki, POLAND

**Received:** March 11, 2019

**Accepted:** May 13, 2019

**Published:** May 31, 2019

**Copyright:** © 2019 Wu, Fan. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The statistical results are within the manuscript and its Supporting Information file. However, raw data underlying the results presented in the study are available from the National Meteorological Information Center (NMIC) of China Meteorological Administration (CMA) (<http://data.cma.cn/>).

**Funding:** This study was jointly supported by the Central Public-interest Scientific Institution Basal Research Fund (Farmland Irrigation Research Institute, CAAS, FIRI2018-01), the National Natural

## Abstract

Accurately predicting reference evapotranspiration ( $ET_0$ ) with limited climatic data is crucial for irrigation scheduling design and agricultural water management. This study evaluated eight machine learning models in four categories, i.e. neuron-based (MLP, GRNN and ANFIS), kernel-based (SVM, KNEA), tree-based (M5Tree, XGBoost) and curve-based (MARS) models, for predicting daily  $ET_0$  with maximum/maximum temperature and precipitation data during 2001–2015 from 14 stations in various climatic regions of China, i.e., arid desert of northwest China (NWC), semi-arid steppe of Inner Mongolia (IM), Qinghai-Tibetan Plateau (QTP), (semi-)humid cold-temperate northeast China (NEC), semi-humid warm-temperate north China (NC), humid subtropical central China (CC) and humid tropical south China (SC). The results showed machine learning models using only temperature data obtained satisfactory daily  $ET_0$  estimates (on average  $R^2 = 0.829$ ,  $RMSE = 0.718$  mm  $day^{-1}$ ,  $NRMSE = 0.250$  and  $MAE = 0.508$  mm  $day^{-1}$ ). The prediction accuracy was improved by 7.6% across China when information of precipitation was further considered, particularly in (sub)tropical humid regions (by 9.7% in CC and 12.4% in SC). The kernel-based SVM, KNEA and curve-based MARS models generally outperformed the others in terms of prediction accuracy, with the best performance by KNEA in NWC and IM, by SVM in QTP, CC and SC, and very similar performance by them in NEC and NC. SVM (1.9%), MLP (2.0%), MARS (2.6%) and KNEA (6.4%) showed relatively small average increases in RMSE during testing compared with training RMSE. SVM is highly recommended for predicting daily  $ET_0$  across China in light of best accuracy and stability, while KNEA and MARS are also promising powerful models.

## Introduction

Accurate prediction of reference evapotranspiration ( $ET_0$ ) is significant for irrigation schedules design, crop growth modeling and agricultural water management [1–5]. Various

Science Foundation of China (No. 51709144). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

mathematical models have been proposed to estimate  $ET_0$  from meteorological variables, among which the FAO-56 Penman–Monteith (FAO-56 PM) equation is suggested by the Food and Agriculture Organization of the United Nations as a reference model in various regions and climates [6], because it considers both the thermodynamic and aerodynamic items. However, the FAO-56 PM model needs a variety of climatic parameters as model inputs for calculation, e.g., maximum and minimum ambient temperatures, wind speed, relative humidity and net radiation [7–11], which significantly restricts the application of the FAO-56 PM model in many worldwide regions. Therefore, the simplified empirical models with fewer climatic variables is becoming increasingly popular in the absence of complete data [12–15], such as temperature-based models [16], mass transfer-based models [17] and radiation-based models [18]. However, evapotranspiration is a complex and highly nonlinear phenomenon dependent on several climatic parameters. Therefore, it is difficult to establish empirical models that can consider all those complicated processes. In recent years, much attention has been drawn to use alternative techniques such as machine learning models for  $ET_0$  prediction as a result of their excellent performance in tackling the nonlinear relationship between the model inputs and output [19–21].

The neuron-based machine learning models, i.e. artificial neural networks (ANNs), are the earliest and most widely used models for  $ET_0$  prediction. [22] compared the multi-layer perceptron (MLP) model and (semi) empirical models to estimate daily  $ET_0$  in the Basque Country of Spain with different input combinations, and revealed that the MLP model obtained better  $ET_0$  estimates than the locally calibrated empirical models. [23] also investigated the potential of three ANNs models, including the MLP model, radial basis function neural networks (RBNN) along with generalized regression neural networks (GRNN) for estimating  $ET_0$  at two weather stations in USA using ambient temperatures, relative humidity, wind speed and solar radiation. It was concluded that the MLP and RBNN models attained satisfactory  $ET_0$  estimates. [14] evaluated the potential of four ANNs models, including the generalized feedforward (GFF), linear regression (LR), probabilistic neural network (PNN) and MLP models for short-term  $ET_0$  prediction using forecasted weather data (minimum/maximum temperatures and net solar radiation). It was also found that the MLP model was generally superior to the other three ANNs models. [24] assessed the ANFIS and ANNs models for predicting daily  $ET_0$  in two locations in South Korea based on temperature, sunshine hours, wind speed and relative humidity. They found that the proposed machine learning models performed well for  $ET_0$  estimation. [25] compared the capability of three ANNs models (MLP, RBNN and GRNN) with the adaptive neuro fuzzy inference system (ANFIS) and gene expression programming (GEP) for monthly  $ET_0$  estimation in two locations in the Mediterranean Region of Turkey with only temperature data. They found that the GRNN and GEP models performed better than the other models at one station, while the performance of the RBNN and ANFIS models was best at the other station.

The kernel-based machine learning models have become increasingly popular for  $ET_0$  prediction. [26] compared the capability of the kernel-based support vector machine (SVM) with the multiple linear regression (MLR), multiple nonlinear regression (MNL), ANFIS and empirical models for predicting  $ET_0$  in the semi-arid regions of Iran. The results indicated that the SVM and ANFIS models performed better than the other models. [13] Wen et al. (2015) predicted daily  $ET_0$  using the SVM model with few climatic variables in arid China. It was found that the SVM model outperformed the ANNs and empirical models. [27] evaluated the capabilities of the least-squares support vector machine (LS-SVM), multivariate adaptive regression spline (MARS) and M5 model tree (M5Tree) for the estimation of  $ET_0$  in Mediterranean Turkey, and found that the kernel-based LS-SVM model was superior to the other two models. [28] assessed the potential of the kernel-based extreme learning machine (ELM) to

predict daily  $ET_0$  in Iraq, and found it exhibited good efficiency and generalization performances. [29] estimated daily  $ET_0$  with the ELM and GRNN models based on air temperatures along in southwestern China. They found that the ELM model was superior to the GRNN model. The kernel-based SVM, LS-SVM and ELM models are also coupled with pre-treatment or optimization algorithms such as wavelet transform (WT) [30] and genetic algorithm (GA) [31] to improve the prediction accuracy. Recently, an improved version of kernel-based machine learning models, i.e., kernel-based nonlinear extension of Arps decline model (KNEA) has been developed and successfully applied in various fields [32–33]. However, this new powerful model has not yet been tested in evapotranspiration studies.

The tree-based machine learning models have recently started to draw researchers' attention due to their relative simplicity but powerful capability in time-series prediction [34]. [35] evaluated the performance of the M5Tree model for estimating daily  $ET_0$  in California of USA, and found the M5Tree model gave satisfactory  $ET_0$  estimates. [36] evaluated the M5Tree and feedforward ANNs models to predict  $ET_0$  in the arid regions. It was concluded that  $ET_0$  values predicted by the M5Tree and ANNs models agreed well with the FAO-56 PM values. [37] compared the M5Tree and ANNs models to predict  $ET_0$  at two sites in USA. They found that the M5Tree model outperformed the ANNs models for estimation of  $ET_0$  when the input and output data at the target station were not available. [38] compared the random forest (RF) to the GRNN model for prediction of daily  $ET_0$  in southwestern China, and concluded that the RF model gave better daily  $ET_0$  estimates than the GRNN model. [39] further used a hybrid RF model with the wavelet algorithm for daily  $ET_0$  estimation in Southern Iran. They indicated that the new coupled RF model outperformed the classic RF model. [40] explored the performance of two kernel-based models and four tree-based models for daily  $ET_0$  estimation with limited meteorological data across China. They found that the extreme gradient boosting (XGBoost) and gradient boosting decision tree (GBDT) exhibited similar accuracy and stability compare with the kernel-based SVM and ELM models.

Other machine learning models, such as the curve-based MARS model, have also been applied to predict  $ET_0$ . [27] modeled monthly  $ET_0$  in Mediterranean Turkey using the MARS, LSSVM and M5Tree models. It was found that the MARS model outperformed the least-squares support vector regression (LSSVR) and M5Tree models. [41] predicted monthly  $ET_0$  in Iran using the MARS, SVM, GEP and empirical models. The results showed that the MARS and SVM-RBF models were generally superior to the GEP and SVM-Poly models. [42] has also evaluated the performances of the GEP model and the (semi)empirical models for predicting daily  $ET_0$  in the hyper-arid regions of Iran, and revealed that the superiority of the GEP model for  $ET_0$  estimation over the (semi)empirical models.

Although the neuron-based, kernel-based, tree-based and curve-based machine learning models have been widely used to predict  $ET_0$  around the world, their performances are inconsistent in various  $ET_0$  studies. Particularly, there is still lack of direct and comprehensive comparison of various categories of machine learning models for prediction of  $ET_0$  in a specific region or country such as China with a vast territory and diverse climates. Due to the limited availability of complete climatic variables, the applicability of these machine learning models for estimation of  $ET_0$  with more cheaply and reliably measured meteorological variables (e.g. temperature and precipitation) should be explored. In addition, although the utmost attention is usually paid to prediction accuracy when applying machine learning models, model stability is another major factor to consider because unstable models may produce inaccurate  $ET_0$  estimates if new data are included [34]. Thus, the objectives of this study are to: (1) determine the effects of temperature and precipitation (a variable representing relative humidity to some extent) on the prediction accuracy of daily  $ET_0$  in different climatic zones of China, and (2) further compare both the prediction accuracy and model stability of eight machine learning

models in four categories (MLP, GRNN, ANFIS, SVM, KNEA, M5Tree, XGBoost and MARS) for predicting daily  $ET_0$  across China as a case study.

## Materials and methods

### Case study and site description

According to multiple-year mean temperatures, precipitation and altitude (Table 1), China is classified as seven climatic regions (Fig 1), i.e. the arid desert of northwest China (NWC), the semi-arid steppe of Inner Mongolia (IM), the Qinghai-Tibetan Plateau (QTP), the (semi-) humid cold-temperate northeast China (NEC), the semi-humid warm-temperate north China (NC), the humid subtropical central China (CC), and the humid tropical south China (SC) [43–44]. The mean daily air temperature in NWC, IM, QTP, NEC, NC, CC and SC varied between  $-13.2^{\circ}\text{C}$  and  $26.2^{\circ}\text{C}$ ,  $-10.5^{\circ}\text{C}$  and  $21.1^{\circ}\text{C}$ , between  $-9.6^{\circ}\text{C}$  and  $19.8^{\circ}\text{C}$ , between  $-19.3^{\circ}\text{C}$  and  $25.7^{\circ}\text{C}$ , between  $-4.0^{\circ}\text{C}$  and  $28.4^{\circ}\text{C}$ , between  $1.5^{\circ}\text{C}$  and  $30.7^{\circ}\text{C}$  and between  $3.5^{\circ}\text{C}$  and  $35.6^{\circ}\text{C}$ , respectively. The average annual precipitation were 269 mm in NWC, 302 mm in IM, 382 mm in QTP, 637 mm in NEC, 658 mm in NC, 1538 mm in CC and 1964 mm in SC.

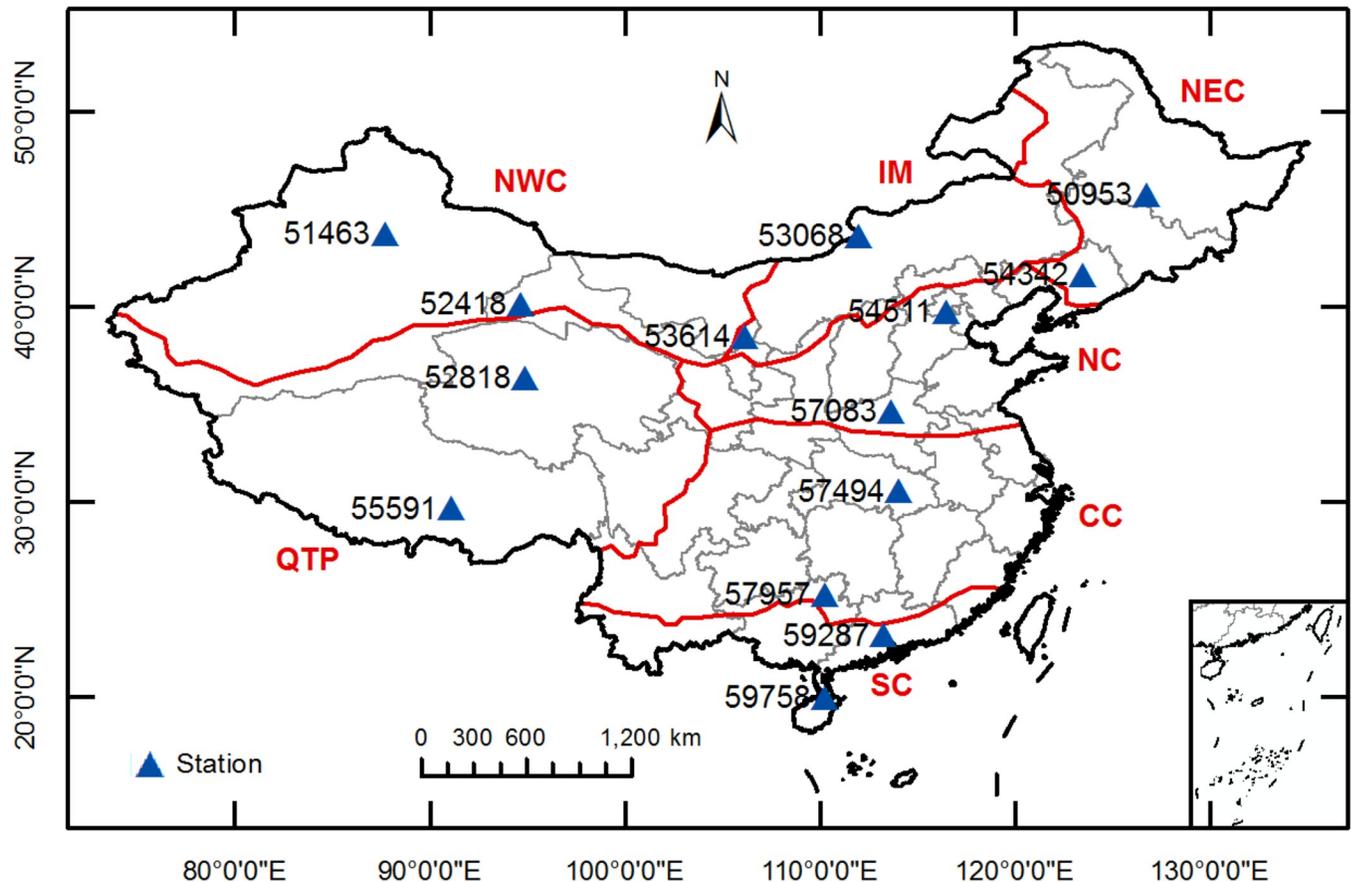
### Data collection and analysis

Long-term daily maximum ( $T_{\max}$ ) and minimum ( $T_{\min}$ ) temperatures, relative humidity ( $H_r$ ), wind speed ( $U$ ) and horizontal global solar radiation ( $R_s$ ) from 2001–2015 were obtained at 14 representative stations across various climatic regions of China (Fig 1). The geographical locations and meteorological values of the 14 stations are presented in Table 1. These meteorological data are provided by the National Meteorological Information Center (NMIC) of China Meteorological Administration (CMA), who has rigorously examined the data quality. The daily data were further excluded if any of the above meteorological data were missing. Overall, missing data accounted for only 0.08% of the database, ranging from 0 to 0.53% at various weather stations.

**Table 1. The geographical locations and annual mean values ( $\pm$  standard deviation) of meteorological data during 2001–2015 for each of the 14 weather stations used in the present study.**

Station ID	Station name	Latitude (N)	Longitude (E)	Altitude (m)	$T_{\max}$ ( $^{\circ}\text{C}$ )	$T_{\min}$ ( $^{\circ}\text{C}$ )	P ( $\text{mm yr}^{-1}$ )	$ET_0$ ( $\text{mm d}^{-1}$ )
51463	Urumqi	43.8	87.7	935	12.9 $\pm$ 13.6	2.8 $\pm$ 12.8	271 $\pm$ 75.7	2.9 $\pm$ 2.3
52418	Dunhuang	40.2	94.7	1139	18.3 $\pm$ 14.6	2.2 $\pm$ 11.5	42 $\pm$ 5.1	3.3 $\pm$ 1.7
53614	Yinchuan	38.5	106.2	1111	16.2 $\pm$ 10.5	3.6 $\pm$ 9.3	203 $\pm$ 58.6	2.9 $\pm$ 1.8
53068	Erenhot	43.7	112.0	965	11.9 $\pm$ 11.4	-2.2 $\pm$ 10.7	142 $\pm$ 31.2	3.3 $\pm$ 1.7
50953	Harbin	45.8	126.8	142	10.2 $\pm$ 15.2	-0.9 $\pm$ 14.8	479 $\pm$ 132.1	2.3 $\pm$ 1.8
54342	Shenyang	41.7	123.5	49	14.1 $\pm$ 13.4	3.1 $\pm$ 12.4	656 $\pm$ 150.3	2.4 $\pm$ 1.6
54511	Beijing	39.8	116.5	31	18.1 $\pm$ 11.4	7.4 $\pm$ 10.5	510 $\pm$ 145.7	2.9 $\pm$ 1.9
57083	Zhengzhou	34.7	113.7	110	20.4 $\pm$ 13.5	9.8 $\pm$ 11.6	626 $\pm$ 154.5	2.9 $\pm$ 1.9
52818	Geermu	36.4	94.9	2808	13.1 $\pm$ 9.7	-1.2 $\pm$ 9.6	44 $\pm$ 5.6	3.3 $\pm$ 1.8
55591	Lasa	29.7	91.1	3649	16.0 $\pm$ 6.0	1.8 $\pm$ 7.1	416 $\pm$ 132.8	3.4 $\pm$ 1.4
57494	Wuhan	30.6	114.1	23	21.4 $\pm$ 9.7	13.2 $\pm$ 9.2	1241 $\pm$ 201.4	2.5 $\pm$ 1.7
57957	Guilin	25.3	110.3	164	23.3 $\pm$ 8.2	16.0 $\pm$ 8.7	1856 $\pm$ 248.3	2.7 $\pm$ 1.4
59287	Guangzhou	23.2	113.3	41	26.5 $\pm$ 6.3	19.0 $\pm$ 6.4	1921 $\pm$ 270.1	2.7 $\pm$ 1.3
59758	Haikou	20.0	110.3	64	28.1 $\pm$ 5.7	21.6 $\pm$ 6.0	2067 $\pm$ 401.7	3.2 $\pm$ 1.3

<https://doi.org/10.1371/journal.pone.0217520.t001>



**Fig 1. The distribution of different climatic regions across China and the geographical locations of 14 weather stations used in this study.** NWC: arid desert of northwest China, IM: semi-arid steppe of Inner Mongolia, QTP: Qinghai-Tibetan Plateau, NEC: (semi-)humid cold-temperate northeast China, NC: semi-humid warm-temperate north China, CC: humid subtropical central China, and SC: humid tropical south China; the South China Sea Islands are presented in the bottom right-hand corner; the same below.

<https://doi.org/10.1371/journal.pone.0217520.g001>

### FAO-56 Penman–Monteith equation

The FAO-56 Penman–Monteith model was utilized to calculate daily  $ET_0$  ( $\text{mm day}^{-1}$ ) and provide the reference data for the training and testing of machine learning models in this paper:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T_{\text{mean}} + 273} U_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34U_2)} \quad (1)$$

where,  $R_n$ : net radiation ( $\text{MJ m}^{-2} \text{day}^{-1}$ );  $G$ : soil heat flux ( $\text{MJ m}^{-2} \text{day}^{-1}$ );  $T_{\text{mean}}$ : average ambient temperature ( $^{\circ}\text{C}$ );  $U_2$ : wind speed at 2 m height ( $\text{m s}^{-1}$ );  $e_s$ : saturation vapor pressure (kPa);  $e_a$ : actual vapor pressure (kPa);  $\Delta$ : slope of the vapor pressure curve ( $\text{kPa } ^{\circ}\text{C}^{-1}$ );  $\gamma$ : psychrometric constant ( $\text{kPa } ^{\circ}\text{C}^{-1}$ ). Detailed calculation procedures can be found in [6].

### Machine learning models for predicting daily reference evapotranspiration

**Multilayer perceptron neural networks (MLP).** The MLP model is one of the widely used ANNs models, which is a feed-forward neural network for nonlinear function approximation. The MLP model consists of three layers: the input, hidden and output layers. The hidden layer often has only one layer, and the number of neural unit has to be determined by the trial and error approach. In the present study, a three-layer neural network was developed: the

first layer was input layer, the neural number was equal to the input numbers and the output layer has one neural unit. The MLP model is trained by the Levenberg–Marquardt algorithm, which interpolates between the Gauss–Newton algorithm (GNA) and the gradient descent algorithm. It is more robust than the GNA but still can stick with local, rather than global minima. Further details on the MLP model refer to [4].

**Generalized regression neural network (GRNN).** The GRNN model is proposed by [45] and is one of the radial basis function neural network (RBF) models. This model can approximate non-linear function of the input and output vectors with a function estimate obtained from the training dataset. It shows a parallel structure and no iterative process is required for model learning between the inputs and outputs. The GRNN model does not need iterative training procedures compared with the back propagation method. Further details on the GRNN model is given by [46].

**Adaptive neuro-fuzzy inference system (ANFIS).** The ANFIS model is proposed by [47], which is a multi-layer adaptive network coupled with neural networks with the fuzzy inference system. The first-order Sugeno fuzzy model with two fuzzy if-then rules is applied in the ANFIS model to approximate the nonlinear function in this study. The ANFIS model is consisted of five layers: the fuzzification, product, normalization, de-fuzzification and output layers. The model uses different node functions to learn and adjust the parameters in a fuzzy inference system, where the forward and backward passes are applied to decrease the computed errors. More details about the ANFIS model is given by [47].

**M5 model tree (M5Tree).** The M5Tree model is firstly established by [48], which is a powerful learning method to estimate the true values in a large dataset. It has a series of linear regression functions at the terminal nodes, which develops relationships between the independent variables and a dependent variable. The model firstly constructs a regression tree by splitting the instance space in a recursive manner, and selects the one maximizing the expected errors reduction following all the potential splits. The over-grown trees are then pruned and the sub-trees are then replaced by the linear regression functions. Further details of the M5Tree model refer to [49].

**Extreme gradient boosting (XGBoost).** The XGBoost model is proposed by [50] and is originated from the idea of "boosting". The XGBoost model integrates all the predictions of a series of "weak" learners to develop a "strong" learner via an additive training process. It is supposed to avoid the over-fitting issue and reduce the computational time. This can be obtained by simplifying the objective functions and combining the predictive and regularization terms, while it maintains optimum computation efficiency at the same time. Parallel calculation is also automatically implemented during the training period. More details about the XGBoost model refer to [50].

**Support vector machine (SVM).** The SVM model is developed by [51], which is widely used for classification, pattern recognition and regression analysis. The SVM model can estimate the regression on the basis of a set of kernel functions, which are capable of implicitly converting the original, lower-dimensional input dataset to a higher-dimensional feature space. The SVM model has been successfully applied in predicting  $ET_0$  [31, 41]. The radial basis function (RBF) nonlinear kernel function was utilized in the present study as a result of its outstanding performance for predicting  $ET_0$  relative to other kernel functions [19](Kisi, 2015), such as linear, polynomial and sigmoid functions. Further information about the SVM model is given by [51].

**Kernel-based nonlinear extension of Arps decline model (KNEA).** The KNEA model is a newly nonlinear model initially proposed by [32] based on the Arps decline model and kernel method. Compared with the non-parametric and "Black-Box" kernel-based models such as least-squares SVM, the KNEA model is based on the idea of "Grey-Box" and uses the semi-

parametric formulation to build the nonlinear models [52]. The kernel-based grey system models are more efficient with small samples [53–54], while the KNEA model performed better with larger samples since samples are not accumulated in the model.

The KNEA model can be described as:

$$f(x) = af(x - 1) + g(u(x)) + \mu \tag{2}$$

where  $f(x)$  is the output at this time,  $f(x-1)$  is the output at the last step time.  $u(x)$  represents the factors affecting the output,  $g(u(x))$  can be interpreted as the relationship between  $u(x)$  and  $f(x)$ ,  $\mu$  is the bias. From this model, we can see that the output of this time is the result of joint action between the output from last time step and the influencing factors at this time.

The nonlinear function  $g(u(x))$  is difficult to determine and can be translated to:

$$g(u(x)) = \omega^T \varphi(u(x)) \tag{3}$$

This means mapping the original influencing factors into the new space. Therefore, the formula (2) can be written as:

$$f(x) = af(x - 1) + \omega^T \varphi(u(x)) + \mu \tag{4}$$

Although we still can't solve Eq (4), we can find a very small value so that the difference between the left and right side of the equation is as small as possible:

$$e_x = f(x) - af(x - 1) - \omega^T \varphi(u(x)) - \mu \tag{5}$$

$$\min_{a, \omega, e} \zeta(a, \omega, e) = \frac{1}{2} a^2 + \frac{1}{2} \|\omega\|^2 + \frac{\gamma}{2} \sum_{x=2}^n e_x^2 \tag{6}$$

$$\text{s.t. } f(x) = af(x - 1) + \omega^T \varphi(u(x)) + \mu + e_x \tag{7}$$

where  $\gamma$  is the regularization term, it can control the model smoothness. Like SVM, this optimization problem can be solved by Lagrangian multiplier method:

$$\mathcal{L}(a, \omega, \mu, e, \lambda) = \zeta(a, \omega, e) - \sum_{x=2}^n \lambda_x \{bf(x - 1) + \omega^T \varphi(u(x)) + \mu + e_x - f(x)\} \tag{8}$$

where  $\lambda_x$  is the Lagrangian multiplier. The Karush–Kuhn–Tucker (KKT) conditions for

optimality of the Lagrangian multiplier method are as follows:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial a} = 0 & a = \sum_{x=2}^n \lambda_x f(x-1) \\ \frac{\partial \mathcal{L}}{\partial \omega} = 0 & \omega = \sum_{x=2}^n \lambda_x \varphi(u(x)) \\ \frac{\partial \mathcal{L}}{\partial \mu} = 0 & \sum_{x=2}^n \lambda_x = 0 \\ \frac{\partial \mathcal{L}}{\partial e_x} = 0 & e_x = \lambda_x \gamma^{-1} \\ \frac{\partial \mathcal{L}}{\partial \lambda_x} = 0 & f(x) = af(x-1) + \omega^T \varphi(u(x)) + \mu + e_x \end{cases} \quad (9)$$

$$\begin{pmatrix} 0 & 1_{n-1}^T \\ 1_{n-1} & \Omega + Q + \frac{1}{\gamma} I_{n-1} \end{pmatrix} \begin{pmatrix} \mu \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ x_{2:n} \end{pmatrix} \quad (10)$$

where

$$\begin{aligned} I_{n-1} &= [1, 1, \dots, 1]_{n-1}^T, \\ \lambda_{n-1} &= [\lambda_1, \lambda_2, \dots, \lambda_{n-1}]_{n-1}^T, \\ x_{2:n} &= [x(2), x(3), \dots, x(n)]^T, \end{aligned}$$

$\Omega_{ij} = \varphi(u(i)) \times \varphi(u(j)) = K(u(i), u(j))$ , where  $I_{n-1}$  is an  $n-1$  dimensional identity matrix with all the diagonal elements as 1 and others as 0.  $\lambda, \mu$  and  $a$  can be obtained by Eq (9). The  $\Omega_{ij}$  can be employed a kernel function  $K(\cdot, \cdot)$ , which satisfies the Mercer's theorem, and a Gauss-type kernel function was selected in the present study. Further information about is given by [32].

$$K(u(i), u(j)) = \exp\left[\frac{-(u(i) - u(j))^2}{2\sigma^2}\right] \quad (11)$$

**Multivariate adaptive regression spline (MARS).** The MARS model is a non-parametric regression approach proposed by [55], which needs no assumption on the relationships between the independent and dependent variables. In the MARS model, a series of coefficients and functions defined as basis functions are used for modeling. The basis function of the MARS model is the outcome of a truncated spline function or multiple spline functions. The number of basis functions and the determination of basis functions are automatically determined by data. Meanwhile, the MARS model integrates the merits of the recursive auto-fractional regression method in dividing spatial regions, projection tracking method in processing high-dimensional data and the advantages of accumulative regression node self-adaptation. Further details of the MARS model refer to [10].

### Input combinations and K-fold cross-validation

Precipitation is not directly correlated to  $ET_0$ , but it is a manifestation of relative humidity to some extent and may correct the temperature-based  $ET_0$  models. However, the real amount of precipitation may underestimate or exaggerate its effect on the reduction of daily  $ET_0$  due to large variation range from 0 mm to even hundreds of mm in humid regions. Therefore, a simple transformed precipitation ( $P_t$ , 1 for precipitation > 0 and 0 for precipitation = 0) was

**Table 2. The five cross-validation stages involved in the present study.**

Cross validation	Training dataset	Testing dataset
S1	2004–2015	2001–2003
S2	2001–2003 and 2007–2015	2004–2006
S3	2001–2006 and 2010–2015	2007–2009
S4	2001–2009 and 2013–2015	2010–2012
S5	2001–2012	2013–2015

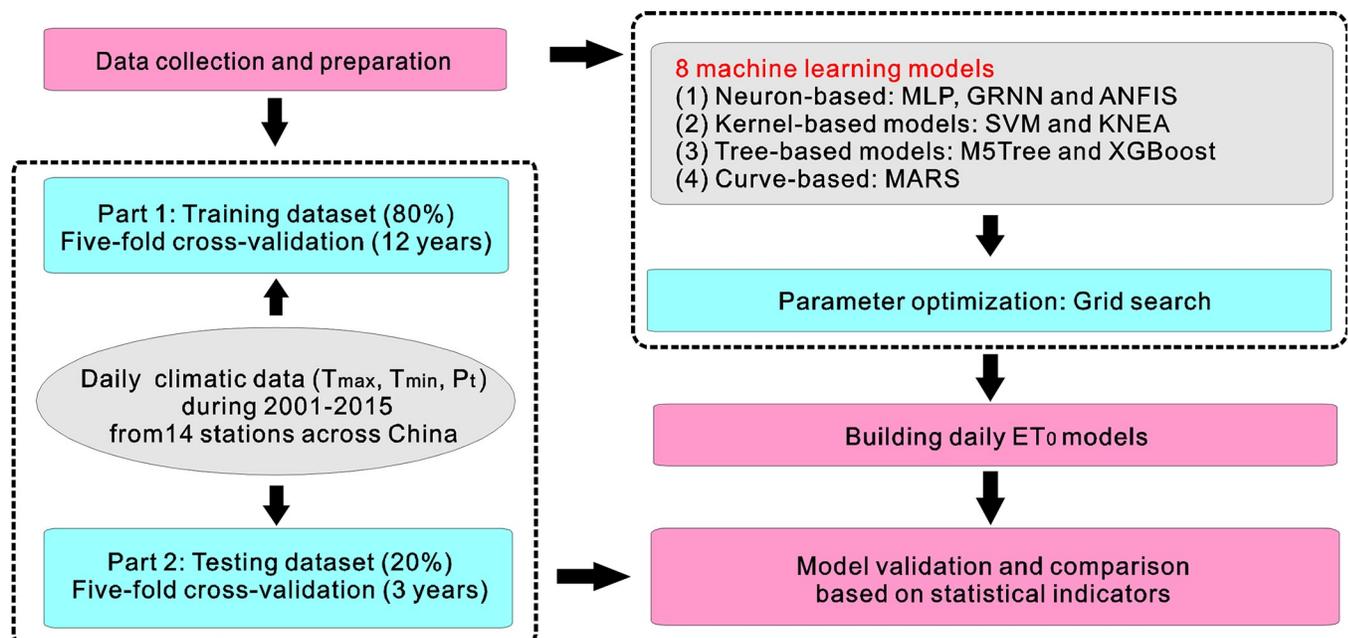
<https://doi.org/10.1371/journal.pone.0217520.t002>

applied here to represent the general effect of precipitation on ET<sub>0</sub> prediction. Two input combinations of meteorological variables were thus used in the present study to assess the temperature and precipitation effects on daily ET<sub>0</sub> prediction, i.e., C1: T<sub>max</sub>, T<sub>min</sub> and R<sub>a</sub>; C2: T<sub>max</sub>, T<sub>min</sub>, and R<sub>a</sub>. The K-fold cross-validation method was applied, where the obtained temperature and precipitation data during 2001–2015 were equally partitioned into five periods. Four periods were used for model training and the last one was used to test the models, which was run over the five various stages (Table 2). The main parameters of the eight machine learning models were optimized by using the grid-search method. Fig 2 presents the simple flowchart of the proposed methodology in the present study.

### Statistical evaluation

Four common statistical indicators were used in this study to evaluate the models, i.e., RMSE, R<sup>2</sup>, MAE and NRMSE, which can be expressed as [56–57]:

$$R^2 = \frac{\left[ \sum_{i=1}^n (Y_{i,m} - \bar{Y}_{i,m})(Y_{i,e} - \bar{Y}_{i,e}) \right]^2}{\sum_{i=1}^n (Y_{i,m} - \bar{Y}_{i,m})^2 \sum_{i=1}^n (Y_{i,e} - \bar{Y}_{i,e})^2} \tag{12}$$



**Fig 2. Simple flowchart of the proposed methodology in the present study.**

<https://doi.org/10.1371/journal.pone.0217520.g002>

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{i,m} - Y_{i,e})^2} \tag{13}$$

$$NRMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{i,m} - Y_{i,e})^2} / \bar{Y}_{i,m} \times 100\% \tag{14}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_{i,m} - Y_{i,e}| \tag{15}$$

where  $Y_{i,m}$ ,  $Y_{i,e}$ ,  $\bar{Y}_{i,m}$  and  $\bar{Y}_{i,e}$  are the measured, estimated, mean of measured, mean of estimated reference evapotranspiration, respectively;  $n$  is the number of observed data. Higher  $R^2$  values indicate high prediction accuracy, whereas lower values of RMSE, MAE and NRMSE suggest better model performance. Considering the requirements of the MLP and KNEA models, the raw climatic data were normalized between 0 and 1 as follows:

$$z_n = \frac{z_i - z_{\min}}{z_{\max} - z_{\min}} \tag{16}$$

where  $z_n$  and  $z_i$  is the normalized and raw data;  $z_{\max}$  and  $z_{\min}$  are the minimum and maximum raw data.

## Results and discussion

### Comparison of prediction accuracy of eight machine learning models across China

The statistical results of the eight machine learning models (MLP, GRNN, ANFIS SVM, KNEA, M5Tree, XGBoost and MARS) for estimating daily  $ET_0$  in NWC (Urumqi and Dunhuang), IM (Yinchuan and Erenhot), NEC (Harbin and Shenyang), NC (Beijing and Zhengzhou), QTP (Geermu and Lasa), CC (Wuhan and Guilin) and SC (Guangzhou and Haikou) under the two input combinations of climatic variables (C1:  $T_{\max}$ ,  $T_{\min}$  and  $R_a$ ; C2:  $T_{\max}$ ,  $T_{\min}$ ,  $P_t$  and  $R_a$ ) during both training and testing periods, are presented in Tables 3–9, respectively.

In terms of  $R^2$ , RMSE, NRMSE and MAE averaged over the fourteen weather stations across China, the GRNN model (on average 0.888, 0.578 mm day<sup>-1</sup>, 0.201 and 0.394 mm day<sup>-1</sup>, respectively), XGBoost model (on average 0.882, 0.612 mm day<sup>-1</sup>, 0.207 and 0.424 mm day<sup>-1</sup>, respectively) and M5Tree model (on average 0.873, 0.626 mm day<sup>-1</sup>, 0.217 and 0.432 mm day<sup>-1</sup>, respectively) generally produced better prediction accuracy in predicting daily  $ET_0$  than the other machine learning models during training in the whole China. The MLP model (on average  $R^2 = 0.850$ , RMSE = 0.685 mm day<sup>-1</sup>, NRMSE = 0.257 and MAE = 0.476 mm day<sup>-1</sup>) produced the worst performance in all the climatic zones, followed by the MARS model (on average  $R^2 = 0.855$ , RMSE = 0.668 mm day<sup>-1</sup>, NRMSE = 0.232 and MAE = 0.469 mm day<sup>-1</sup>) and SVM model (on average  $R^2 = 0.857$ , RMSE = 0.662 mm day<sup>-1</sup>, NRMSE = 0.232 and MAE = 0.456 mm day<sup>-1</sup>). On the contrary, the SVM model (on average  $R^2 = 0.860$ , RMSE = 0.674 mm day<sup>-1</sup>, NRMSE = 0.234 and MAE = 0.470 mm day<sup>-1</sup>), KNEA model (on average  $R^2 = 0.857$ , RMSE = 0.676 mm day<sup>-1</sup>, NRMSE = 0.236 and MAE = 0.474 mm day<sup>-1</sup>) and MARS model (on average  $R^2 = 0.855$ , RMSE = 0.685 mm day<sup>-1</sup>, NRMSE = 0.237 and MAE = 0.483 mm day<sup>-1</sup>) generally performed better than the other models for daily  $ET_0$  estimation during the testing period in the whole China. The M5Tree model (on average  $R^2 =$

Table 3. Statistical values of the eight machine learning models with different input parameters during training and testing at Urumqi and Dunhuang in the arid desert of northwest China.

Station	Inputs / Model	Training				Testing				(RMSE <sub>tes</sub> —RMSE <sub>trn</sub> )%	(RMSEC <sub>2</sub> —RMSEC <sub>1</sub> ) <sub>tes</sub> %
		R <sup>2</sup>	RMSE	NRMSE	MAE	R <sup>2</sup>	RMSE	NRMSE	MAE		
			(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		
Urumqi	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.893	0.833	0.283	0.452	0.889	0.841	0.289	0.491	1.0%	-
	GRNN1	<b>0.930</b>	<b>0.667</b>	<b>0.224</b>	<b>0.373</b>	0.897	0.829	0.279	0.486	24.3%	-
	SVM1	0.893	0.824	0.278	0.458	0.895	0.837	0.281	0.490	1.6%	-
	KNEA1	0.911	0.750	0.252	0.429	<b>0.898</b>	<b>0.822</b>	<b>0.277</b>	<b>0.473</b>	9.6%	-
	M5Tree1	0.904	0.779	0.262	0.452	0.889	0.863	0.290	0.511	10.8%	-
	XGBoost1	0.914	0.738	0.248	0.445	0.895	0.836	0.281	0.485	13.3%	-
	ANFIS1	0.907	0.766	0.258	0.443	0.893	0.840	0.282	0.478	9.7%	-
	MARS1	0.897	0.807	0.272	0.476	0.897	0.829	0.278	0.492	2.7%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>0</sub>, R<sub>a</sub>)</b>										
	MLP2	0.915	0.785	0.252	0.432	0.917	0.799	0.256	0.440	1.8%	5.0%
	GRNN2	<b>0.945</b>	<b>0.628</b>	<b>0.198</b>	<b>0.344</b>	0.905	0.799	0.269	0.465	27.2%	3.6%
	SVM2	0.903	0.782	0.263	0.439	0.903	0.806	0.271	0.457	3.1%	3.7%
	KNEA2	0.926	0.706	0.231	0.395	<b>0.924</b>	<b>0.781</b>	<b>0.242</b>	<b>0.421</b>	10.6%	5.0%
	M5Tree2	0.925	0.732	0.232	0.404	0.915	0.812	0.256	0.453	10.9%	5.9%
	XGBoost2	0.940	0.699	0.208	0.364	0.923	0.784	0.244	0.436	12.2%	<b>6.2%</b>
	ANFIS2	0.933	0.723	0.220	0.386	0.830	0.789	0.372	0.459	9.1%	6.1%
	MARS2	0.917	0.764	0.244	0.434	0.918	0.784	0.250	0.447	2.6%	5.4%
Dunhuang	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.889	0.741	0.224	0.468	0.883	0.750	0.233	0.501	1.2%	-
	GRNN1	<b>0.923</b>	<b>0.603</b>	<b>0.181</b>	<b>0.392</b>	0.884	0.766	0.224	0.502	27.0%	-
	SVM1	0.889	0.725	0.218	0.482	0.889	0.732	0.220	0.491	1.0%	-
	KNEA1	0.901	0.683	0.205	0.455	<b>0.890</b>	<b>0.729</b>	<b>0.219</b>	<b>0.490</b>	6.7%	-
	M5Tree1	0.901	0.681	0.205	0.453	0.881	0.757	0.228	0.516	11.2%	-
	XGBoost1	0.904	0.669	0.201	0.458	0.887	0.737	0.222	0.502	10.2%	-
	ANFIS1	0.897	0.695	0.209	0.465	0.886	0.742	0.223	0.501	6.8%	-
	MARS1	0.888	0.727	0.219	0.487	0.886	0.737	0.222	0.497	1.4%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>0</sub>, R<sub>a</sub>)</b>										
	MLP2	0.893	0.703	0.222	0.487	0.892	0.712	0.223	0.486	1.3%	5.1%
	GRNN2	<b>0.926</b>	<b>0.578</b>	<b>0.177</b>	<b>0.382</b>	0.886	0.730	0.223	0.499	26.3%	4.7%
	SVM2	0.899	0.688	0.207	0.460	<b>0.899</b>	<b>0.694</b>	<b>0.210</b>	<b>0.471</b>	<b>0.9%</b>	5.2%
	KNEA2	0.902	0.649	0.204	0.457	0.897	0.703	0.211	0.477	8.3%	3.6%
	M5Tree2	0.905	0.649	0.201	0.448	0.884	0.726	0.225	0.510	11.9%	4.1%
	XGBoost2	0.923	0.633	0.181	0.418	0.895	0.710	0.214	0.486	12.2%	3.7%
	ANFIS2	0.907	0.662	0.199	0.447	0.885	0.698	0.224	0.498	5.4%	<b>5.9%</b>
	MARS2	0.895	0.693	0.212	0.478	0.894	0.705	0.215	0.487	1.7%	4.3%

<https://doi.org/10.1371/journal.pone.0217520.t003>

0.840, RMSE = 0.716 mm day<sup>-1</sup>, NRMSE = 0.249 and MAE = 0.504 mm day<sup>-1</sup>), MLP model (on average R<sup>2</sup> = 0.851, RMSE = 0.698 mm day<sup>-1</sup>, NRMSE = 0.259 and MAE = 0.487 mm day<sup>-1</sup>) and GRNN model (on average R<sup>2</sup> = 0.845, RMSE = 0.696 mm day<sup>-1</sup>, NRMSE = 0.245 and MAE = 0.492 mm day<sup>-1</sup>) performed worst among the machine learning models. These results are in good agreement with other machine learning-based ET<sub>0</sub> studies. For instance, [58] found that the LS-SVM model yielded accurate ET<sub>0</sub> estimation in the Changwu County, China. [13] indicated that the SVM model outperformed the ANNs model for the estimation

Table 4. Statistical values of the eight machine learning models with different input parameters during training and testing at Yinchuan and Erenhot in the semi-arid steppe of Inner Mongolia.

Station	Inputs / Model	Training				Testing				(RMSE <sub>tes</sub> —RMSE <sub>trn</sub> )%	(RMSEC2—RMSE <sub>CI</sub> ) <sub>tes</sub> %
		R <sup>2</sup>	RMSE	NRMSE	MAE	R <sup>2</sup>	RMSE	NRMSE	MAE		
			(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		
Yinchuan	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.873	0.693	0.249	0.477	0.875	0.696	0.251	0.479	0.4%	-
	GRNN1	<b>0.914</b>	<b>0.566</b>	<b>0.192</b>	<b>0.370</b>	0.873	0.704	0.238	0.477	24.4%	-
	SVM1	0.881	0.669	0.226	0.444	0.881	0.680	0.230	0.462	1.6%	-
	KNEA1	0.893	0.629	0.213	0.418	<b>0.882</b>	<b>0.680</b>	<b>0.230</b>	<b>0.456</b>	8.1%	-
	M5Tree1	0.893	0.632	0.214	0.421	0.863	0.727	0.246	0.494	15.0%	-
	XGBoost1	0.896	0.625	0.212	0.427	0.878	0.688	0.233	0.469	10.1%	-
	ANFIS1	0.888	0.647	0.219	0.436	0.880	0.685	0.232	0.460	5.9%	-
	MARS1	0.879	0.673	0.228	0.458	0.878	0.689	0.233	0.471	2.4%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.892	0.645	0.232	0.449	0.894	0.661	0.238	0.462	2.5%	5.0%
	GRNN2	<b>0.924</b>	<b>0.536</b>	<b>0.181</b>	<b>0.354</b>	0.880	0.674	0.231	0.467	25.7%	4.3%
	SVM2	0.885	0.637	0.222	0.439	0.873	0.644	0.238	0.475	<b>1.1%</b>	5.3%
	KNEA2	0.903	0.601	0.203	0.413	<b>0.900</b>	<b>0.624</b>	<b>0.211</b>	<b>0.431</b>	3.8%	<b>8.2%</b>
	M5Tree2	0.906	0.590	0.200	0.406	0.878	0.687	0.232	0.481	16.4%	5.5%
	XGBoost2	0.923	0.567	0.182	0.378	0.898	0.632	0.214	0.440	11.5%	8.1%
	ANFIS2	0.906	0.591	0.200	0.410	0.884	0.651	0.227	0.456	10.2%	5.0%
	MARS2	0.895	0.627	0.212	0.437	0.895	0.642	0.217	0.450	2.4%	6.8%
Erenhot	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.896	0.846	0.254	0.555	0.901	0.866	0.259	0.569	2.4%	-
	GRNN1	<b>0.937</b>	<b>0.656</b>	<b>0.193</b>	<b>0.404</b>	0.899	0.863	0.253	0.568	31.6%	-
	SVM1	0.900	0.826	0.243	0.538	0.904	0.842	0.247	0.554	<b>1.9%</b>	-
	KNEA1	0.916	0.752	0.221	0.486	<b>0.907</b>	<b>0.836</b>	<b>0.245</b>	<b>0.553</b>	11.2%	-
	M5Tree1	0.912	0.774	0.228	0.502	0.891	0.896	0.263	0.588	15.8%	-
	XGBoost1	0.916	0.751	0.221	0.499	0.904	0.846	0.248	0.560	12.6%	-
	ANFIS1	0.907	0.788	0.232	0.516	0.901	0.857	0.252	0.558	8.8%	-
	MARS1	0.901	0.820	0.241	0.539	0.905	0.842	0.247	0.556	2.7%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.908	0.795	0.239	0.523	0.914	0.802	0.240	0.532	<b>0.9%</b>	7.4%
	GRNN2	<b>0.943</b>	<b>0.624</b>	<b>0.184</b>	<b>0.385</b>	0.904	0.826	0.248	0.560	32.4%	4.3%
	SVM2	0.912	0.774	0.228	0.507	0.917	0.791	0.232	0.526	2.2%	6.1%
	KNEA2	0.920	0.716	0.217	0.486	<b>0.920</b>	<b>0.781</b>	<b>0.229</b>	<b>0.520</b>	9.1%	6.6%
	M5Tree2	0.919	0.740	0.218	0.485	0.902	0.856	0.251	0.574	15.7%	4.5%
	XGBoost2	0.936	0.686	0.193	0.447	0.917	0.794	0.233	0.532	15.7%	6.1%
	ANFIS2	0.921	0.730	0.215	0.486	0.909	0.792	0.244	0.547	8.5%	<b>7.6%</b>
	MARS2	0.910	0.778	0.229	0.517	0.915	0.801	0.235	0.535	3.0%	4.9%

<https://doi.org/10.1371/journal.pone.0217520.t004>

of daily ET<sub>0</sub> in an extreme arid region of China. [27] found that the LS-SVM model outperformed the MARS and M5Tree models, while the MARS model was superior to the LS-SVM and M5Tree models in cross-station applications. [59] suggested that the SVM model gave better daily ET<sub>0</sub> estimation than the tree-based assemble models (RF, M5Tree, GBDT and XGBoost) under various input combinations across China.

Specifically, the GRNN model outperformed all the other machine learning models in daily ET<sub>0</sub> modeling in the seven climatic zones of China during the training period. However,

**Table 5. Statistical values of the eight machine learning models with different input parameters during training and testing at Harbin and Shenyang in the (semi-) humid cold-temperate northeast China.**

Station	Inputs / Model	Training				Testing				(RMSE <sub>tes</sub> —RMSE <sub>trn</sub> )%	(RMSEC2—RMSEC1) <sub>tes</sub> %
		R <sup>2</sup>	RMSE	NRMSE	MAE	R <sup>2</sup>	RMSE	NRMSE	MAE		
			(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		
Harbin	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.861	0.675	0.313	0.439	0.868	0.701	0.326	0.457	3.9%	-
	GRNN1	<b>0.915</b>	<b>0.528</b>	<b>0.241</b>	<b>0.322</b>	0.873	0.670	0.306	0.434	26.9%	-
	SVM1	0.875	0.643	0.294	0.409	<b>0.881</b>	<b>0.652</b>	<b>0.298</b>	<b>0.422</b>	<b>1.4%</b>	-
	KNEA1	0.892	0.595	0.272	0.378	0.877	0.663	0.303	0.427	11.4%	-
	M5Tree1	0.891	0.595	0.271	0.377	0.861	0.699	0.319	0.448	17.5%	-
	XGBoost1	0.896	0.586	0.267	0.385	0.879	0.660	0.302	0.431	12.6%	-
	ANFIS1	0.887	0.608	0.278	0.391	0.880	0.658	0.300	0.423	8.2%	-
	MARS1	0.876	0.638	0.291	0.416	0.878	0.661	0.302	0.432	3.6%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.884	0.618	0.286	0.397	0.888	0.639	0.287	0.399	3.4%	<b>8.8%</b>
	GRNN2	<b>0.926</b>	<b>0.494</b>	<b>0.226</b>	<b>0.300</b>	0.881	0.631	0.297	0.423	27.7%	5.8%
	SVM2	0.893	0.598	0.273	0.376	<b>0.895</b>	<b>0.611</b>	<b>0.279</b>	<b>0.389</b>	<b>2.2%</b>	6.3%
	KNEA2	0.899	0.556	0.263	0.369	0.894	0.612	0.279	0.396	10.1%	7.7%
	M5Tree2	0.902	0.546	0.258	0.358	0.877	0.654	0.298	0.421	19.8%	6.4%
	XGBoost2	0.923	0.534	0.230	0.335	0.893	0.615	0.281	0.401	15.2%	6.8%
	ANFIS2	0.904	0.560	0.256	0.359	0.888	0.612	0.289	0.404	9.3%	7.0%
	MARS2	0.889	0.601	0.274	0.391	0.890	0.626	0.286	0.407	4.2%	5.3%
Shenyang	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.829	0.730	0.320	0.487	0.833	0.743	0.325	0.498	1.8%	-
	GRNN1	<b>0.896</b>	<b>0.571</b>	<b>0.238</b>	<b>0.367</b>	0.837	0.724	0.301	0.490	26.8%	-
	SVM1	0.838	0.716	0.298	0.468	0.840	0.725	0.301	0.480	<b>1.3%</b>	-
	KNEA1	0.862	0.653	0.272	0.433	<b>0.849</b>	<b>0.699</b>	<b>0.291</b>	<b>0.474</b>	7.0%	-
	M5Tree1	0.865	0.648	0.270	0.429	0.827	0.749	0.311	0.504	15.6%	-
	XGBoost1	0.865	0.647	0.270	0.441	0.840	0.719	0.299	0.487	11.1%	-
	ANFIS1	0.857	0.666	0.277	0.448	0.841	0.718	0.299	0.484	7.8%	-
	MARS1	0.840	0.703	0.293	0.478	0.841	0.720	0.299	0.491	2.4%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.851	0.682	0.299	0.459	0.854	0.687	0.301	0.465	<b>0.7%</b>	7.5%
	GRNN2	<b>0.907</b>	<b>0.539</b>	<b>0.225</b>	<b>0.346</b>	0.846	0.685	0.293	0.478	27.1%	5.4%
	SVM2	0.859	0.666	0.278	0.438	0.861	0.674	0.280	0.449	1.2%	7.0%
	KNEA2	0.871	0.614	0.264	0.424	<b>0.865</b>	<b>0.660</b>	<b>0.275</b>	<b>0.445</b>	7.5%	5.6%
	M5Tree2	0.878	0.615	0.256	0.410	0.842	0.703	0.297	0.485	14.3%	6.1%
	XGBoost2	0.898	0.585	0.235	0.391	0.860	0.671	0.279	0.458	14.7%	6.7%
	ANFIS2	0.879	0.614	0.256	0.415	0.860	0.661	0.279	0.454	7.7%	<b>7.9%</b>
	MARS2	0.858	0.664	0.277	0.453	0.856	0.685	0.285	0.470	3.2%	4.9%

<https://doi.org/10.1371/journal.pone.0217520.t005>

during the testing period, the statistical indicators indicated that the KNEA model had the lowest values of average RMSE, NRMSE, MAE but the highest R<sup>2</sup> values in NWC (0.757 mm day<sup>-1</sup>, 0.241, 0.476 mm day<sup>-1</sup> and 0.899, respectively) and IM (0.732 mm day<sup>-1</sup>, 0.229, 0.491 mm day<sup>-1</sup> and 0.902, respectively). The lowest values of average RMSE, NRMSE, MAE but the highest R<sup>2</sup> values in QTP (0.477 mm day<sup>-1</sup>, 0.145, 0.339 mm day<sup>-1</sup> and 0.914, respectively), CC (0.757 mm day<sup>-1</sup>, 0.241, 0.476 mm day<sup>-1</sup> and 0.899, respectively) and SC (0.757 mm day<sup>-1</sup>, 0.241, 0.476 mm day<sup>-1</sup> and 0.899, respectively) were obtained by the SVM model. However,

Table 6. Statistical values of the eight machine learning models with different input parameters during training and testing at Beijing and Zhengzhou in the semi-humid warm-temperate north China.

Station	Inputs / Model	Training				Testing				(RMSE <sub>tes</sub> —RMSE <sub>trn</sub> )%	(RMSEC <sub>2</sub> —RMSEC <sub>1</sub> ) <sub>tes</sub> %
		R <sup>2</sup>	RMSE	NRMSE	MAE	R <sup>2</sup>	RMSE	NRMSE	MAE		
			(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		
Beijing	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.792	0.840	0.311	0.603	0.792	0.841	0.311	0.603	<b>0.1%</b>	-
	GRNN1	<b>0.853</b>	<b>0.704</b>	<b>0.244</b>	<b>0.495</b>	0.787	0.849	0.294	0.613	20.6%	-
	SVM1	0.795	0.836	0.289	0.580	<b>0.797</b>	<b>0.831</b>	<b>0.287</b>	<b>0.588</b>	0.6%	-
	KNEA1	0.817	0.783	0.271	0.558	0.795	0.835	0.289	0.603	6.6%	-
	M5Tree1	0.819	0.779	0.269	0.552	0.774	0.877	0.303	0.633	12.6%	-
	XGBoost1	0.819	0.779	0.270	0.566	0.791	0.842	0.291	0.611	8.1%	-
	ANFIS1	0.809	0.801	0.277	0.574	0.794	0.842	0.292	0.599	5.1%	-
	MARS1	0.797	0.826	0.286	0.596	0.793	0.838	0.290	0.608	1.5%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.818	0.787	0.291	0.576	0.816	0.788	0.292	0.578	<b>0.1%</b>	6.3%
	GRNN2	<b>0.869</b>	<b>0.667</b>	<b>0.231</b>	<b>0.473</b>	0.799	0.805	0.286	0.601	20.7%	5.2%
	SVM2	0.823	0.774	0.268	0.548	<b>0.822</b>	<b>0.780</b>	<b>0.270</b>	<b>0.560</b>	0.8%	6.1%
	KNEA2	0.829	0.758	0.262	0.554	0.821	0.785	0.272	0.573	3.6%	6.0%
	M5Tree2	0.840	0.733	0.254	0.528	0.798	0.829	0.287	0.606	13.1%	5.5%
	XGBoost2	0.860	0.705	0.237	0.508	0.820	0.783	0.271	0.577	11.1%	<b>7.0%</b>
ANFIS2	0.837	0.740	0.256	0.541	0.810	0.784	0.278	0.585	5.9%	6.9%	
MARS2	0.819	0.780	0.270	0.575	0.816	0.791	0.274	0.586	1.4%	5.6%	
Zhengzhou	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.785	0.808	0.310	0.594	0.784	0.822	0.316	0.606	1.7%	-
	GRNN1	<b>0.845</b>	<b>0.683</b>	<b>0.237</b>	<b>0.482</b>	0.781	0.812	0.282	0.593	18.9%	-
	SVM1	0.791	0.796	0.277	0.561	0.789	0.806	0.280	0.579	<b>1.3%</b>	-
	KNEA1	0.811	0.752	0.261	0.541	<b>0.790</b>	<b>0.797</b>	<b>0.277</b>	<b>0.570</b>	6.0%	-
	M5Tree1	0.815	0.743	0.258	0.534	0.767	0.840	0.292	0.617	13.1%	-
	XGBoost1	0.814	0.748	0.260	0.549	0.784	0.809	0.281	0.594	8.2%	-
	ANFIS1	0.804	0.766	0.266	0.556	0.787	0.803	0.279	0.583	4.8%	-
	MARS1	0.790	0.793	0.276	0.577	0.788	0.800	0.278	0.584	0.9%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.810	0.757	0.291	0.558	0.808	0.761	0.293	0.561	<b>0.5%</b>	7.4%
	GRNN2	<b>0.862</b>	<b>0.645</b>	<b>0.224</b>	<b>0.458</b>	0.795	0.767	0.274	0.577	18.9%	5.5%
	SVM2	0.812	0.757	0.263	0.535	0.810	0.762	0.265	0.544	0.7%	5.5%
	KNEA2	0.828	0.718	0.250	0.528	<b>0.820</b>	<b>0.738</b>	<b>0.257</b>	<b>0.543</b>	2.8%	7.4%
	M5Tree2	0.836	0.700	0.243	0.509	0.799	0.781	0.272	0.576	11.6%	7.0%
	XGBoost2	0.858	0.681	0.226	0.485	0.818	0.741	0.258	0.549	8.8%	<b>8.4%</b>
ANFIS2	0.835	0.703	0.244	0.518	0.809	0.751	0.265	0.558	6.8%	6.5%	
MARS2	0.815	0.745	0.259	0.552	0.811	0.756	0.263	0.560	1.5%	5.5%	

<https://doi.org/10.1371/journal.pone.0217520.t006>

the SVM and KNEA models exhibited very close predication accuracy in NEC (on average R<sup>2</sup> = 0.871, RMSE = 0.659 mm day<sup>-1</sup>, NRMSE = 0.287 and MAE = 0.436 mm day<sup>-1</sup>) and NC (on average R<sup>2</sup> = 0.806, RMSE = 0.788 mm day<sup>-1</sup>, NRMSE = 0.273 and MAE = 0.576 mm day<sup>-1</sup>). Much lower statistical errors were observed in QTP (on average R<sup>2</sup> = 0.907, RMSE = 0.489 mm day<sup>-1</sup>, NRMSE = 0.154 and MAE = 0.351 mm day<sup>-1</sup>) compared with those obtained in the other climatic zones (on average R<sup>2</sup> = 0.850, RMSE = 0.685 mm day<sup>-1</sup>, NRMSE = 0.257 and MAE = 0.476 mm day<sup>-1</sup>). The results suggested that these machine learning models gave

Table 7. Statistical values of the eight machine learning models with different input parameters during training and testing at Geermu and Lasa in the Qinghai-Tibetan Plateau.

Station	Inputs / Model	Training				Testing				(RMSE <sub>tes</sub> —RMSE <sub>trn</sub> )%	(RMSEC2—RMSE <sub>C1</sub> ) <sub>tes</sub> %
		R <sup>2</sup>	RMSE	NRMSE	MAE	R <sup>2</sup>	RMSE	NRMSE	MAE		
			(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		
Geermu	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.912	0.544	0.193	0.372	0.916	0.558	0.199	0.386	2.6%	-
	GRNN1	<b>0.935</b>	<b>0.466</b>	<b>0.147</b>	<b>0.305</b>	0.920	0.545	0.173	0.376	17.0%	-
	SVM1	0.919	0.521	0.165	0.346	<b>0.923</b>	<b>0.535</b>	<b>0.170</b>	<b>0.363</b>	2.7%	-
	KNEA1	0.925	0.503	0.159	0.336	0.920	0.541	0.172	0.370	7.6%	-
	M5Tree1	0.927	0.492	0.155	0.328	0.914	0.562	0.178	0.388	14.2%	-
	XGBoost1	0.929	0.485	0.153	0.333	0.921	0.541	0.172	0.373	11.5%	-
	ANFIS1	0.925	0.503	0.159	0.338	0.921	0.537	0.171	0.366	6.8%	-
	MARS1	0.917	0.524	0.166	0.356	0.921	0.542	0.172	0.370	3.4%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.927	0.497	0.176	0.346	0.931	0.520	0.185	0.366	4.6%	6.8%
	GRNN2	<b>0.951</b>	<b>0.436</b>	<b>0.128</b>	<b>0.288</b>	0.923	0.512	0.169	0.368	17.4%	6.1%
	SVM2	0.927	0.485	0.157	0.335	<b>0.937</b>	<b>0.494</b>	<b>0.154</b>	<b>0.344</b>	<b>1.9%</b>	7.7%
	KNEA2	0.934	0.471	0.149	0.323	0.936	0.496	0.157	0.346	5.3%	<b>8.3%</b>
	M5Tree2	0.937	0.46	0.145	0.315	0.922	0.528	0.171	0.377	14.8%	6.0%
	XGBoost2	0.941	0.445	0.141	0.294	0.930	0.503	0.163	0.353	13.0%	7.0%
ANFIS2	0.939	0.471	0.142	0.312	0.886	0.509	0.202	0.368	8.1%	5.2%	
MARS2	0.929	0.485	0.153	0.337	0.932	0.505	0.160	0.354	4.1%	6.8%	
Lasa	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.885	0.468	0.17	0.347	0.884	0.468	0.170	0.347	<b>0.0%</b>	-
	GRNN1	<b>0.905</b>	<b>0.422</b>	<b>0.122</b>	<b>0.308</b>	0.885	0.467	0.135	0.345	10.7%	-
	SVM1	0.891	0.452	0.131	0.331	<b>0.888</b>	<b>0.460</b>	<b>0.133</b>	<b>0.339</b>	1.8%	-
	KNEA1	0.897	0.441	0.127	0.323	0.885	0.465	0.134	0.344	5.4%	-
	M5Tree1	0.902	0.427	0.123	0.311	0.873	0.491	0.142	0.364	15.0%	-
	XGBoost1	0.904	0.424	0.122	0.315	0.887	0.462	0.133	0.342	9.0%	-
	ANFIS1	0.898	0.437	0.126	0.322	0.883	0.472	0.136	0.344	8.0%	-
	MARS1	0.889	0.455	0.132	0.336	0.886	0.463	0.134	0.344	1.8%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.904	0.424	0.157	0.325	0.902	0.430	0.156	0.324	1.4%	8.1%
	GRNN2	<b>0.931</b>	<b>0.378</b>	<b>0.103</b>	<b>0.27</b>	0.895	0.437	0.129	0.332	15.6%	6.4%
	SVM2	0.896	0.411	0.127	0.324	<b>0.909</b>	<b>0.416</b>	<b>0.120</b>	<b>0.311</b>	<b>1.2%</b>	<b>9.6%</b>
	KNEA2	0.911	0.408	0.118	0.304	0.907	0.421	0.122	0.314	3.2%	9.5%
	M5Tree2	0.916	0.396	0.114	0.293	0.894	0.448	0.129	0.336	13.1%	8.8%
	XGBoost2	0.916	0.398	0.115	0.292	0.887	0.433	0.134	0.337	8.8%	6.3%
ANFIS2	0.918	0.394	0.114	0.294	0.892	0.443	0.131	0.325	12.4%	6.1%	
MARS2	0.905	0.422	0.122	0.317	0.902	0.429	0.124	0.322	1.7%	7.3%	

<https://doi.org/10.1371/journal.pone.0217520.t007>

higher accuracy in the Qinghai-Tibet Plateau. The RMSE values obtained by these best-performing models in each climatic zone were generally smaller than or close to those obtained in the corresponding regions by previous studies when using only T<sub>max</sub> and T<sub>min</sub> data, e.g., by the SVM (0.539 mm day<sup>-1</sup>) and ANN (0.561 mm day<sup>-1</sup>) models in Ejina City of China [13], by ELM (0.444–0.498 mm day<sup>-1</sup>) mode, GANN (0.445–0.499 mm day<sup>-1</sup>) and WNN (0.443–0.641 mm day<sup>-1</sup>) models in the humid region of southwest China using T<sub>max</sub> and T<sub>min</sub> [2], by SVM

Table 8. Statistical values of the eight machine learning models with different input parameters during training and testing at Wuhan and Guilin in the humid subtropical central China.

Station	Inputs / Model	Training				Testing				(RMSE <sub>tes</sub> —RMSE <sub>trn</sub> )%	(RMSEC <sub>2</sub> —RMSEC <sub>1</sub> ) <sub>tes</sub> %
		R <sup>2</sup>	RMSE	NRMSE	MAE	R <sup>2</sup>	RMSE	NRMSE	MAE		
			(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		
Wuhan	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.858	0.629	0.286	0.432	0.854	0.645	0.294	0.447	2.5%	-
	GRNN1	<b>0.886</b>	<b>0.557</b>	<b>0.224</b>	<b>0.375</b>	0.853	0.636	0.256	0.443	14.2%	-
	SVM1	0.864	0.609	0.245	0.410	<b>0.859</b>	<b>0.623</b>	<b>0.251</b>	<b>0.425</b>	2.3%	-
	KNEA1	0.871	0.591	0.238	0.401	0.854	0.634	0.255	0.437	7.3%	-
	M5Tree1	0.879	0.573	0.231	0.390	0.844	0.657	0.265	0.457	14.7%	-
	XGBoost1	0.878	0.577	0.232	0.408	0.854	0.636	0.256	0.450	10.2%	-
	ANFIS1	0.870	0.593	0.239	0.405	0.856	0.632	0.254	0.432	6.6%	-
	MARS1	0.862	0.613	0.246	0.422	0.855	0.632	0.254	0.437	3.1%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.883	0.569	0.259	0.397	0.879	0.585	0.262	0.402	2.8%	9.3%
	GRNN2	<b>0.916</b>	<b>0.500</b>	<b>0.193</b>	<b>0.344</b>	0.865	0.580	0.245	0.424	16.0%	8.8%
	SVM2	0.890	0.547	0.220	0.377	<b>0.886</b>	<b>0.562</b>	<b>0.226</b>	<b>0.391</b>	<b>2.7%</b>	9.8%
	KNEA2	0.891	0.543	0.218	0.378	0.883	0.567	0.228	0.396	4.4%	<b>10.6%</b>
	M5Tree2	0.901	0.519	0.209	0.363	0.863	0.594	0.247	0.435	14.5%	9.6%
	XGBoost2	0.901	0.510	0.209	0.349	0.882	0.580	0.229	0.402	13.7%	8.8%
	ANFIS2	0.900	0.521	0.210	0.365	0.875	0.567	0.236	0.406	8.8%	10.3%
	MARS2	0.886	0.557	0.224	0.393	0.880	0.574	0.231	0.405	3.1%	9.2%
Guilin	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.777	0.725	0.301	0.551	0.776	0.733	0.305	0.558	<b>1.1%</b>	-
	GRNN1	<b>0.826</b>	<b>0.634</b>	<b>0.230</b>	<b>0.480</b>	0.778	0.718	0.261	0.548	13.2%	-
	SVM1	0.791	0.695	0.252	0.524	<b>0.787</b>	<b>0.705</b>	<b>0.256</b>	<b>0.533</b>	1.4%	-
	KNEA1	0.796	0.684	0.248	0.516	0.775	0.724	0.263	0.548	5.8%	-
	M5Tree1	0.812	0.657	0.238	0.496	0.750	0.763	0.277	0.580	16.1%	-
	XGBoost1	0.803	0.673	0.244	0.519	0.774	0.724	0.263	0.559	7.6%	-
	ANFIS1	0.801	0.676	0.245	0.510	0.783	0.711	0.258	0.538	5.2%	-
	MARS1	0.781	0.710	0.258	0.546	0.775	0.723	0.262	0.555	1.8%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.813	0.665	0.276	0.509	0.812	0.670	0.278	0.512	<b>0.8%</b>	8.6%
	GRNN2	<b>0.86</b>	<b>0.588</b>	<b>0.206</b>	<b>0.435</b>	0.794	0.651	0.251	0.526	10.7%	9.3%
	SVM2	0.83	0.626	0.227	0.47	<b>0.827</b>	<b>0.635</b>	<b>0.231</b>	<b>0.479</b>	1.4%	9.9%
	KNEA2	0.821	0.641	0.233	0.489	0.814	0.657	0.239	0.502	2.5%	9.3%
	M5Tree2	0.841	0.606	0.22	0.456	0.790	0.680	0.254	0.531	12.2%	10.5%
	XGBoost2	0.847	0.595	0.216	0.45	0.821	0.645	0.234	0.493	8.4%	<b>10.9%</b>
	ANFIS2	0.837	0.612	0.222	0.463	0.808	0.639	0.243	0.504	4.4%	10.1%
	MARS2	0.816	0.651	0.236	0.502	0.813	0.660	0.239	0.508	1.4%	8.7%

<https://doi.org/10.1371/journal.pone.0217520.t008>

(0.530–0.868 mm day<sup>-1</sup>), M5Tree (0.637–0.953 mm day<sup>-1</sup>) and XGBoost (0.532–0.817 mm day<sup>-1</sup>) models in different climatic zones of China [40].

The scatter plots of daily FAO-56 PM ET<sub>0</sub> values and those predicted by the eight machine learning models for the capital city of China (Beijing) over the five-fold cross validation periods under the two input combinations during testing are presented in Figs 3 and 4, respectively. It can be seen that the selected machine learning models had various prediction accuracies over the five periods. Overall, higher statistical errors were attained during the S4

Table 9. Statistical values of the eight machine learning models with different input parameters during training and testing at Guangzhou and Haikou in the humid tropical south China.

Station	Inputs / Model	Training				Testing				(RMSE <sub>tes</sub> —RMSE <sub>trn</sub> )%	(RMSEC <sub>2</sub> —RMSE <sub>C1</sub> ) <sub>tes</sub> %
		R <sup>2</sup>	RMSE	NRMSE	MAE	R <sup>2</sup>	RMSE	NRMSE	MAE		
			(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		(mm d <sup>-1</sup> )		
Guangzhou	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.737	0.650	0.268	0.498	0.741	0.663	0.273	0.509	2.0%	-
	GRNN1	<b>0.784</b>	<b>0.591</b>	<b>0.212</b>	<b>0.444</b>	0.745	0.653	0.234	0.501	10.5%	-
	SVM1	0.755	0.627	0.225	0.469	<b>0.755</b>	<b>0.637</b>	<b>0.228</b>	<b>0.480</b>	<b>1.6%</b>	-
	KNEA1	0.767	0.611	0.219	0.462	0.753	0.639	0.229	0.488	4.6%	-
	M5Tree1	0.780	0.593	0.212	0.448	0.719	0.681	0.244	0.519	14.8%	-
	XGBoost1	0.775	0.600	0.215	0.459	0.748	0.646	0.232	0.497	7.7%	-
	ANFIS1	0.768	0.610	0.218	0.462	0.747	0.649	0.232	0.491	6.4%	-
	MARS1	0.749	0.633	0.227	0.485	0.750	0.645	0.231	0.495	1.9%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.782	0.590	0.243	0.455	0.786	0.592	0.245	0.457	<b>0.3%</b>	10.7%
	GRNN2	<b>0.837</b>	<b>0.520</b>	<b>0.183</b>	<b>0.391</b>	0.771	0.592	0.222	0.477	13.8%	9.3%
	SVM2	0.802	0.564	0.202	0.426	<b>0.801</b>	<b>0.567</b>	<b>0.207</b>	<b>0.440</b>	0.5%	11.0%
	KNEA2	0.796	0.554	0.204	0.437	0.793	0.570	0.210	0.451	2.9%	10.8%
	M5Tree2	0.818	0.539	0.193	0.408	0.766	0.615	0.224	0.480	14.1%	9.7%
	XGBoost2	0.811	0.534	0.198	0.419	0.797	0.580	0.208	0.447	8.6%	10.2%
ANFIS2	0.814	0.545	0.195	0.417	0.779	0.579	0.218	0.464	6.2%	10.8%	
MARS2	0.790	0.568	0.208	0.448	0.789	0.572	0.212	0.460	0.7%	<b>11.3%</b>	
Haikou	<b>C1 (T<sub>max</sub>, T<sub>min</sub>, R<sub>a</sub>)</b>										
	MLP1	0.738	0.786	0.262	0.621	0.751	0.798	0.264	0.629	<b>1.5%</b>	-
	GRNN1	<b>0.782</b>	<b>0.710</b>	<b>0.213</b>	<b>0.551</b>	0.752	0.809	0.242	0.641	13.9%	-
	SVM1	0.750	0.760	0.228	0.595	<b>0.760</b>	<b>0.789</b>	<b>0.239</b>	<b>0.625</b>	3.8%	-
	KNEA1	0.742	0.772	0.231	0.602	0.735	0.825	0.247	0.648	6.9%	-
	M5Tree1	0.768	0.735	0.220	0.579	0.729	0.839	0.251	0.660	14.1%	-
	XGBoost1	0.773	0.724	0.217	0.573	0.756	0.802	0.240	0.636	10.8%	-
	ANFIS1	0.764	0.738	0.221	0.580	0.739	0.826	0.247	0.642	11.9%	-
	MARS1	0.741	0.773	0.231	0.611	0.751	0.812	0.243	0.643	5.0%	-
	<b>C2 (T<sub>max</sub>, T<sub>min</sub>, P<sub>t</sub>, R<sub>a</sub>)</b>										
	MLP2	0.805	0.682	0.224	0.524	0.821	0.705	0.236	0.552	<b>3.4%</b>	11.7%
	GRNN2	<b>0.859</b>	<b>0.592</b>	<b>0.171</b>	<b>0.444</b>	0.789	0.706	0.227	0.597	19.3%	12.7%
	SVM2	0.829	0.632	0.189	0.484	<b>0.837</b>	<b>0.675</b>	<b>0.202</b>	<b>0.524</b>	6.8%	14.4%
	KNEA2	0.807	0.668	0.200	0.520	0.814	0.709	0.212	0.556	6.1%	14.1%
	M5Tree2	0.835	0.618	0.185	0.474	0.801	0.734	0.220	0.572	18.8%	12.5%
	XGBoost2	0.805	0.655	0.202	0.529	0.833	0.678	0.203	0.529	3.5%	<b>15.5%</b>
ANFIS2	0.829	0.628	0.188	0.486	0.779	0.716	0.232	0.566	14.0%	13.3%	
MARS2	0.806	0.669	0.200	0.522	0.819	0.709	0.212	0.556	6.0%	12.7%	

<https://doi.org/10.1371/journal.pone.0217520.t009>

period (2010–2012), and the S5 period (2013–2015) produced higher prediction accuracy. These differences were largely resulted from the time-series changes in climatic variables among the five periods. This confirms the needs to apply the K-fold cross-validation method for accurately estimating daily ET<sub>0</sub> in various climates [27, 59]. Nevertheless, the performances of these machine learning models showed the same tendency at various cross-validation stages, with better daily ET<sub>0</sub> estimates by the SVM, KNEA and MARS models. The dispersion degree of the data points of the SVM, KNEA and MARS models was lower than that of the M5Tree,

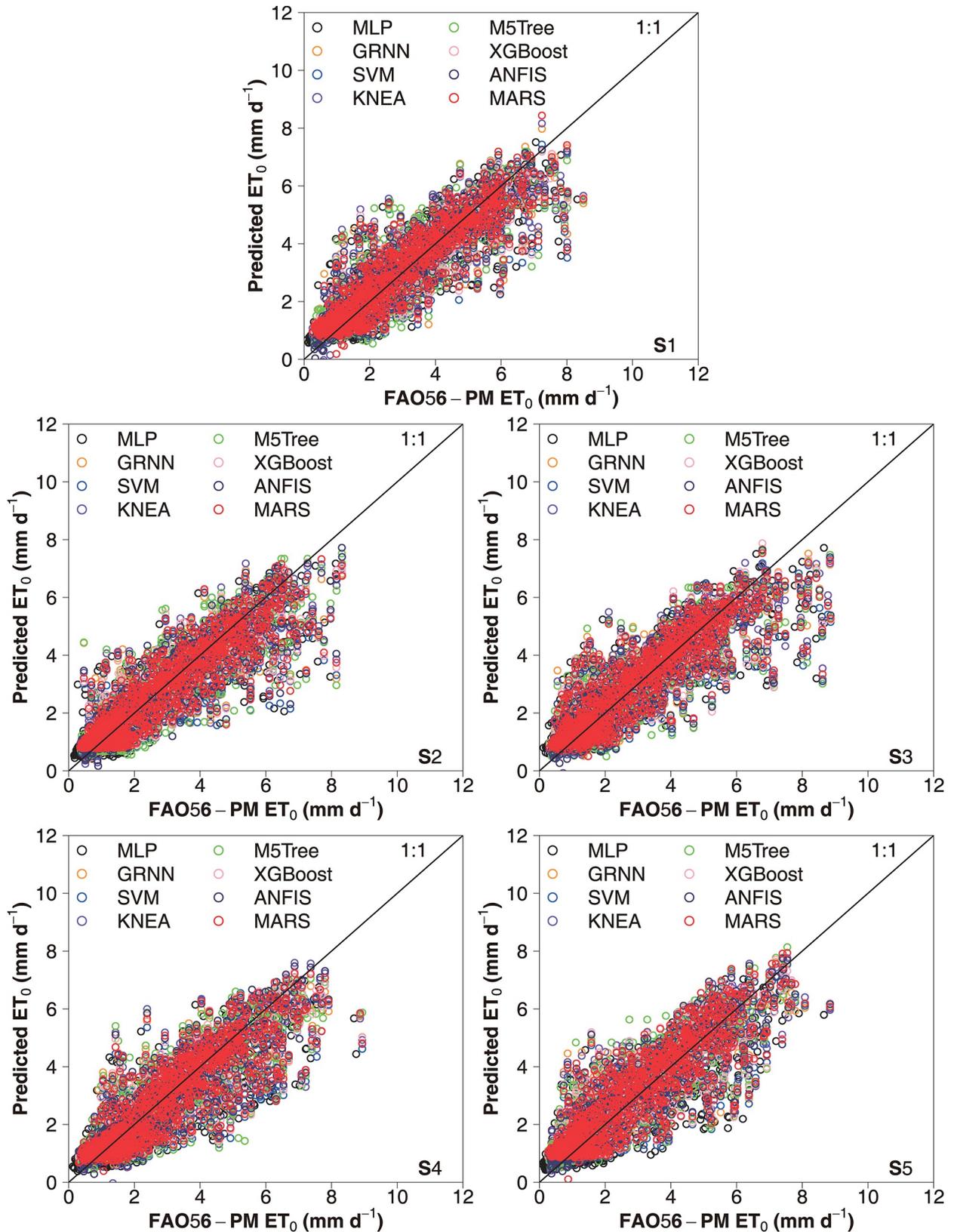


Fig 3. Scatter plots of the  $ET_0$  values calculated by the FAO-56 PM equation for China's capital city of Beijing and the values estimated by the eight machine learning models during five cross-validation stages under the input combination of  $T_{max}$ ,  $T_{min}$  and  $R_a$  in the testing stage.

<https://doi.org/10.1371/journal.pone.0217520.g003>

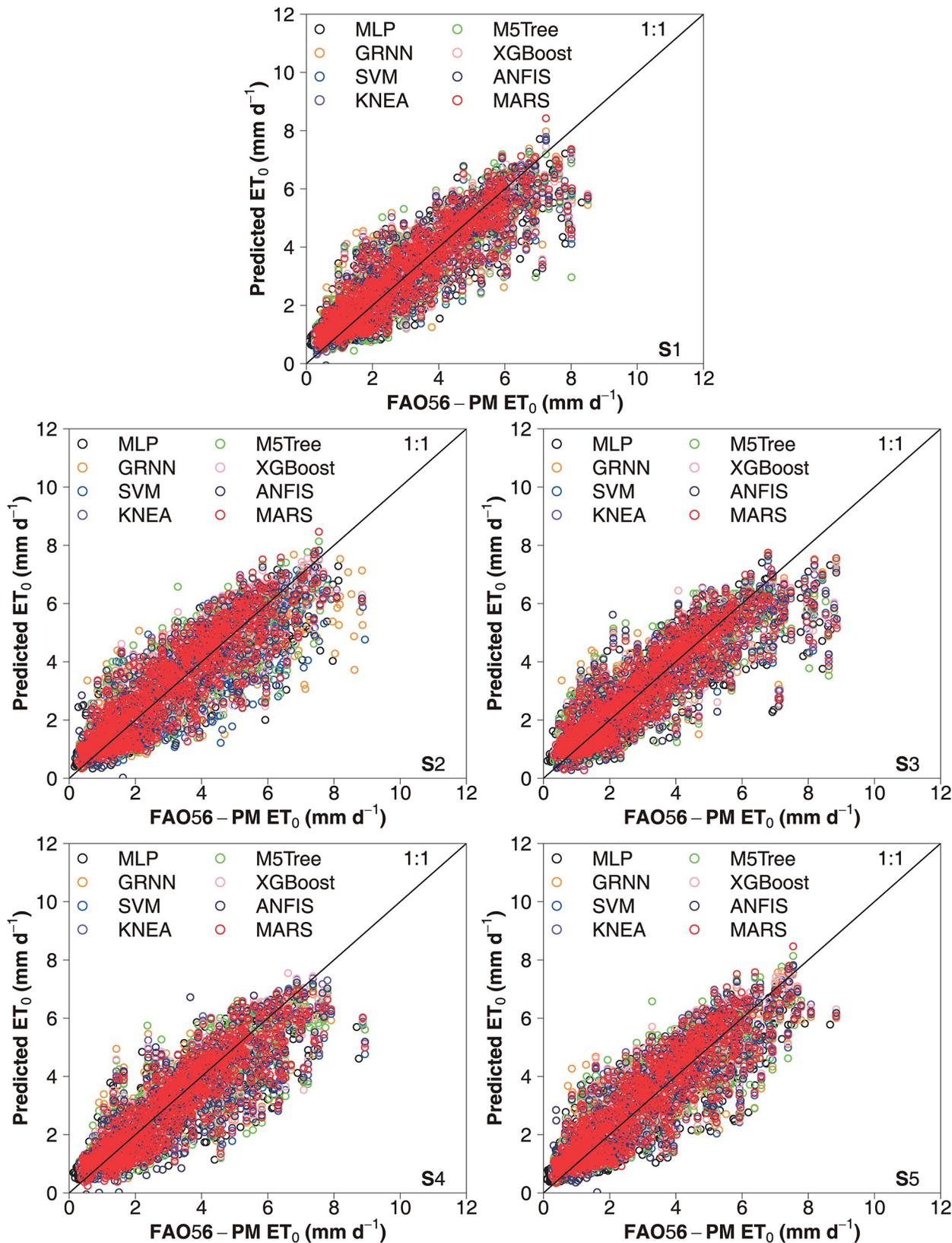


Fig 4. Scatter plots of the  $ET_0$  values calculated by the FAO-56 PM equation for China's capital city of Beijing and the values estimated by the eight machine learning models during five cross-validation stages under the input combination of  $T_{max}$ ,  $T_{min}$ ,  $P_t$  and  $R_a$  in the testing stage.

<https://doi.org/10.1371/journal.pone.0217520.g004>

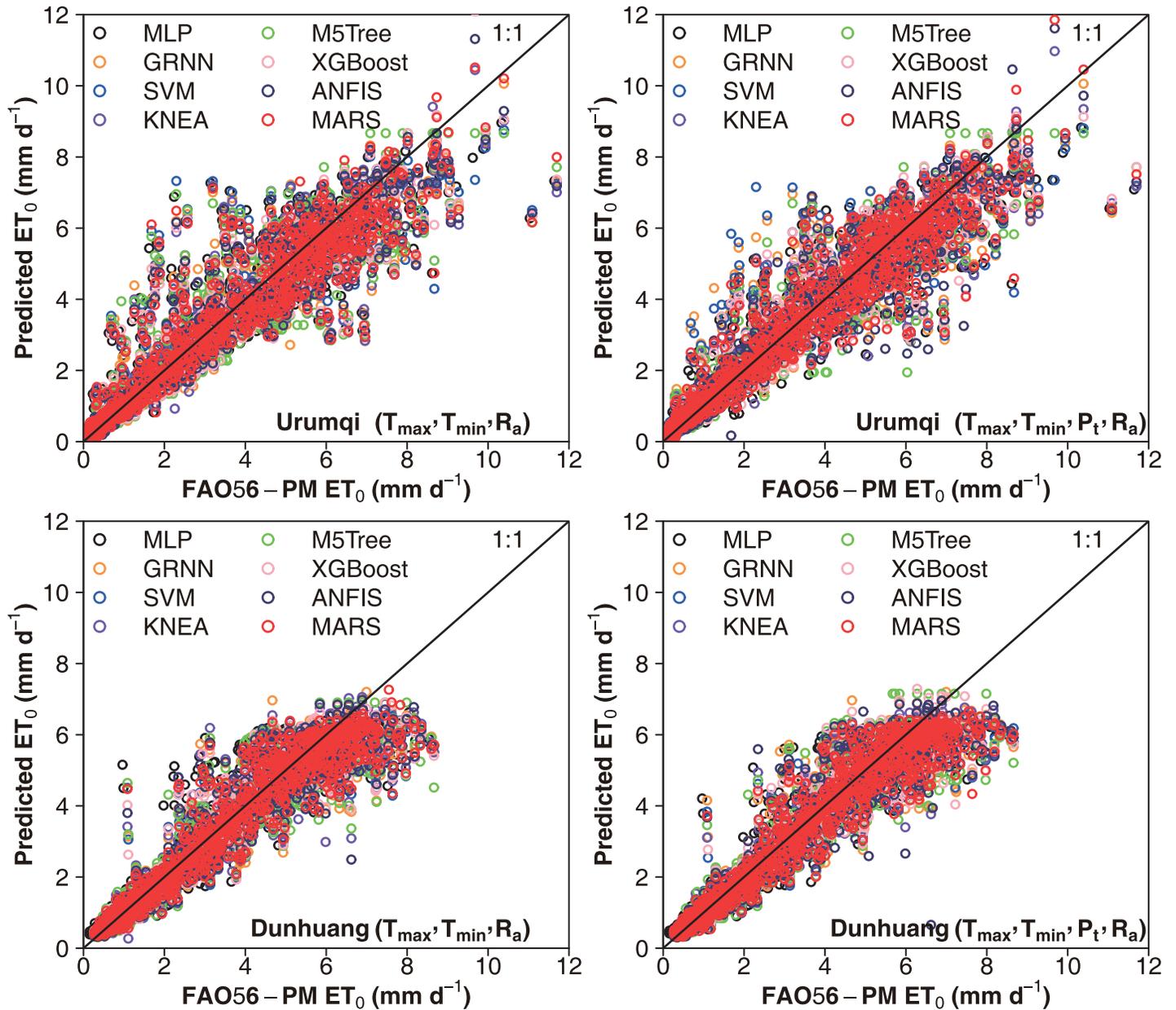


Fig 5. Scatter plots of predicted  $ET_0$  values using the eight machine learning models against their corresponding FAO56-PM values during testing at Urumqi and Dunhuang in the arid desert of northwest China.

<https://doi.org/10.1371/journal.pone.0217520.g005>

MLP and GRNN models. Figs 5 and 6 show the scatter plots of estimated values of daily  $ET_0$  by these machine learning models against the corresponding FAO56-PM values in the arid NWC (Urumqi and Dunhuang) and the humid SC (Guangzhou and Haikou) during the testing period, while the corresponding scatter plots of IM (Yinchuan and Erenhot), NEC (Harbin and Shenyang), NC (Beijing and Zhengzhou), QTP (Geermu and Lasa) and CC (Wuhan and Guilin) can be found in S1–S5 Figs. It is seen that the M5Tree, MLP and GRNN models exhibited more scattered estimates compared with the other models. The daily  $ET_0$  values predicted by the SVM, KNEA and MARS models were more close to the corresponding FAO-56 PM

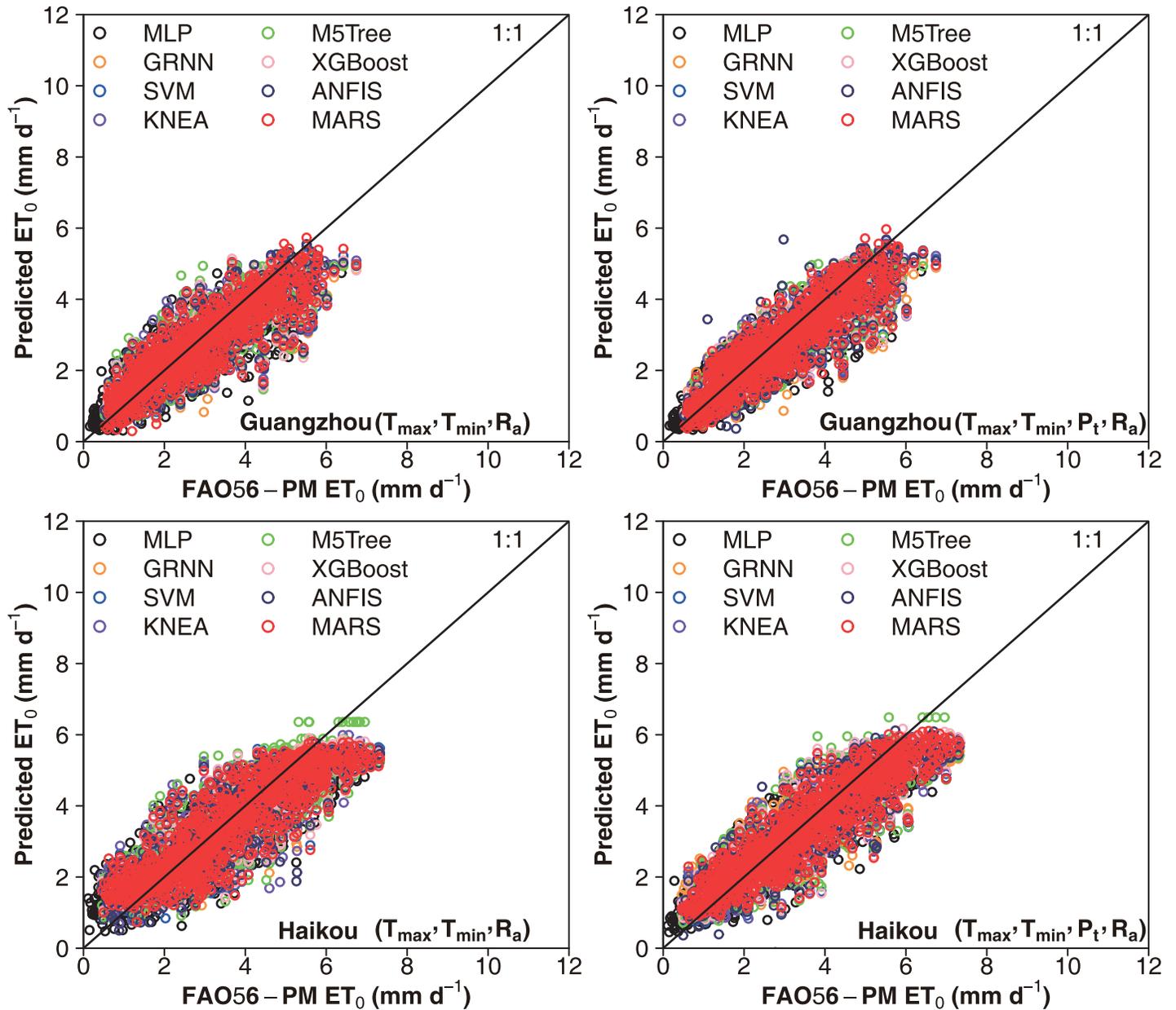


Fig 6. Scatter plots of predicted  $ET_0$  values using the eight machine learning models against their corresponding FAO56-PM values during testing at Guangzhou and Haikou in the humid tropical south China.

<https://doi.org/10.1371/journal.pone.0217520.g006>

values and generally showed the same tendency, which further confirmed the superiority of the SVM, KNEA and MARS for daily  $ET_0$  estimation across China.

### Comparison of prediction accuracy of eight machine learning models under two input combinations

As seen from Tables 3–9, the predicted  $ET_0$  values differed significantly under the two input combinations of climatic variables (C1:  $T_{max}$ ,  $T_{min}$  and  $R_a$ ; C2:  $T_{max}$ ,  $T_{min}$ ,  $P_t$  and  $R_a$ ) during training and testing. Generally, machine learning models with the input combination of  $T_{max}$ ,

$T_{\min}$ ,  $P_t$  and  $R_a$  (on average  $R^2 = 0.879$ ,  $RMSE = 0.613 \text{ mm day}^{-1}$ ,  $NRMSE = 0.214$  and  $MAE = 0.426 \text{ mm day}^{-1}$  during training;  $R^2 = 0.862$ ,  $RMSE = 0.663 \text{ mm day}^{-1}$ ,  $NRMSE = 0.235$  and  $MAE = 0.469 \text{ mm day}^{-1}$  during testing) obtained better prediction accuracy compared with those with the input combination of  $T_{\max}$ ,  $T_{\min}$  and  $R_a$  (on average  $R^2 = 0.844$ ,  $RMSE = 0.666 \text{ mm day}^{-1}$ ,  $NRMSE = 0.232$  and  $MAE = 0.467 \text{ mm day}^{-1}$  during training;  $R^2 = 0.829$ ,  $RMSE = 0.718 \text{ mm day}^{-1}$ ,  $NRMSE = 0.250$  and  $MAE = 0.508 \text{ mm day}^{-1}$  during testing) in the whole China. These statistical results were generally similar to those obtained by various machine learning models with only  $T_{\max}$  and  $T_{\min}$  data in previous studies [2, 13, 40]. However, the incorporation of  $P_t$  as input parameter to the machine learning models introduced an average decrease in RMSE by 7.7% and 7.6% during the training and testing periods, respectively. Specifically, an average decrease in RMSE by 4.8%, 6.0%, 7.4%, 6.5% and 6.5% were obtained during the testing period by considering  $P_t$  in the machine learning models in NWC, IM, QTP, NEC and NC, respectively. However, much higher decreases in RMSE were obtained in CC (by 9.7%) and SC (by 12.4%) by machine learning models with the input combination of  $T_{\max}$ ,  $T_{\min}$ ,  $P_t$  and  $R_a$  compared with those with the input combination of  $T_{\max}$ ,  $T_{\min}$  and  $R_a$ . These results indicated that the incorporation of precipitation information in machine learning models can improve the prediction accuracy of daily  $ET_0$ , particularly in the subtropical and tropical humid regions (Tables 3–9). [59] have found that the prediction accuracy of empirical and machine learning models for estimating daily global solar radiation can be much improved by considering precipitation as an input, because it was a manifestation of cloud cover and could correct the temperature-based models by considering its effects on the radiation reduction. This can also explain why the performance of machine learning models for daily  $ET_0$  prediction from daily minimum/maximum temperature can be enhanced when the information of precipitation was further included.

### Comparison of model stability of eight machine learning models

As illustrated earlier, the GRNN, XGBoost and M5Tree models outperformed the corresponding MLP, SVM, KNEA, ANFIS and MARS models for predicting daily  $ET_0$  in the whole China during the training period in terms of  $R^2$ , RMSE, NRMSE and MAE (Tables 3–9). However, the SVM, KNEA and MARS models produced better daily  $ET_0$  estimates compared with the other machine learning models during the testing period. The percentage increase in RMSE during testing relative to RMSE during training by the eight machine learning models for estimating daily  $ET_0$  in NWC (Urumqi and Dunhuang), IM (Yinchuan and Erenhot), NEC (Harbin and Shenyang), NC (Beijing and Zhengzhou), QTP (Geermu and Lasa), CC (Wuhan and Guilin) and SC (Guangzhou and Haikou) under the two input combinations are also shown in Tables 3–9. The model stability is also an essential factor to consider for predicting more accurate and reliable daily  $ET_0$ . These tables suggested that the SVM, MLP and MARS models were the most stable models with the consistently small percentage increase in RMSE during testing relative to that during training in all the climatic zones of China (on average 1.9%, 2.0% and 2.6%, respectively). The KNEA and ANFIS models also exhibited relatively smaller increase in testing RMSE (on average 6.4% and 7.8%, respectively). However, the GRNN, M5Tree and XGBoost models exhibited the much larger increases in testing RMSE (on average 20.1%, 14.5% and 12.0%, respectively). These increase indicated the instability of the GRNN, M5Tree and XGBoost models as they introduced high decreases in model performances when including new dataset. [40] showed that the kernel-based SVM and ELM models were more stable compared with the tree-based RF, M5Tree and XGBoost models for the estimation of daily  $ET_0$ . [34] also found that the RF and bagging models showed greater increases in RMSE during testing compared with the SVR and gradient models when predicting global solar radiation.

These suggest that the kernel-based machine learning models (e.g., SVM, ELM and KNEA) are generally more stable than the tree-based models (RF, M5Tree and XGBoost).

### Comprehensive evaluation of eight machine learning models

The SVM, KNEA and MARS models outperformed the other machine learning models in daily  $ET_0$  modeling in terms of prediction accuracy during the testing period. Considering the model stability, the SVM, MLP and MARS exhibited very small percentage increase in RMSE during testing ( $< 3.0\%$ ), while the KNEA and ANFIS models showed relatively small increase in testing RMSE ( $< 8.0\%$ ). The SVM model exhibited the best combination of prediction accuracy and model stability among the eight machine learning models, while the KNEA and MARS model also provided satisfactory combination of prediction accuracy and model stability. Comprehensively considering the prediction accuracy and model stability, the SVM, KNEA and MARS models are recommended for estimating daily  $ET_0$  using only temperature and precipitation data across various climatic regions of China and maybe elsewhere in similar climates.

### Conclusions

The performance of eight machine learning models in four categories, e.g. neuron-based (MLP, GRNN, ANFIS), kernel-based (SVM, KNEA), tree-based (M5Tree, XGBoost) and curve-based (MARS) models, for the estimation of daily  $ET_0$  were compared based on only temperature and precipitation data during 2001–2015 obtained from 14 representative stations across various climatic zones of China. The results showed that the machine learning models using only temperature attained satisfactory daily  $ET_0$  estimation. The prediction accuracy was further improved across China when information of precipitation was considered, especially in the (sub)tropical humid regions. This indicates that precipitation is a manifestation of relative humidity to some extent and can correct the temperature-based  $ET_0$  models. The kernel-based SVM, KNEA and curve-based MARS models generally gave more accurate daily  $ET_0$  estimates than the other models for, with the best performance by KNEA in NWC and IM, by SVM in QTP, CC and SC, as well as a similar best performance by them in NEC and NC. The SVM, MLP, MARS and KNEA models showed relatively small percentage increase in RMSE during testing over the training one. Comprehensively considering both prediction accuracy and model stability, SVM is highly suggested, while KNEA and MARS are also alternative models for predicting daily  $ET_0$  in various climatic regions of China. The satisfactory performances of these proposed machine learning models with ambient temperatures and transformed precipitation indicates that it is possible for near-future prediction of daily  $ET_0$  using public weather forecasts, including daily maximum and minimum temperatures and whether there is precipitation or not. Nevertheless, more study is needed to explore the performances of the proposed machine learning models at varying temporal scales or in various climatic regions.

### Supporting information

**S1 Fig. Scatter plots of predicted  $ET_0$  values using the eight machine learning models against their corresponding FAO56-PM values during testing at Yinchuan and Erenhot in the semi-arid steppe of Inner Mongolia.**  
(EPS)

**S2 Fig. Scatter plots of predicted  $ET_0$  values using the eight machine learning models against their corresponding FAO56-PM values during testing at Harbin and Shenyang in**

**the (semi-)humid cold-temperate northeast China.**  
(EPS)

**S3 Fig. Scatter plots of predicted  $ET_0$  values using the eight machine learning models against their corresponding FAO56-PM values during testing at Beijing and Zhengzhou in the semi-humid warm-temperate north China.**

(EPS)

**S4 Fig. Scatter plots of predicted  $ET_0$  values using the eight machine learning models against their corresponding FAO56-PM values during testing at Geermu and Lasa in the Qinghai-Tibetan Plateau.**

(EPS)

**S5 Fig. Scatter plots of predicted  $ET_0$  values using the eight machine learning models against their corresponding FAO56-PM values during testing at Wuhan and Guilin in the humid subtropical central China.**

(EPS)

**S1 Data.**

(RAR)

## Acknowledgments

Thanks to the National Meteorological Information Center of China Meteorological Administration for offering the meteorological data.

## Author Contributions

**Conceptualization:** Junliang Fan.

**Data curation:** Lifeng Wu.

**Formal analysis:** Lifeng Wu.

**Methodology:** Lifeng Wu.

**Supervision:** Junliang Fan.

**Validation:** Junliang Fan.

**Writing – original draft:** Lifeng Wu.

**Writing – review & editing:** Junliang Fan.

## References

1. Fan J, Oestergaard KT, Guyot A, Lockington DA. Estimating groundwater recharge and evapotranspiration from water table fluctuations under three vegetation covers in a coastal sandy aquifer of subtropical Australia. *J Hydrol.* 2014; 519: 1120–1129. <https://doi.org/10.1016/j.jhydrol.2014.08.039>
2. Feng Y, Cui NB, Zhao L, Hu XT, Gong DZ. Comparison of ELM, GANN, WNN and empirical models for estimating reference evapotranspiration in humid region of Southwest China. *J Hydrol.* 2016; 536: 376–383. <https://doi.org/10.1016/j.jhydrol.2016.02.053>
3. Karimi S, Kisi O, Kim S, Nazemi AH, Shiri J. Modelling daily reference evapotranspiration in humid locations of South Korea using local and cross-station data management scenarios. *Int J Climatol.* 2017; 37: 3238–3246. <https://doi.org/10.1002/joc.4911>
4. Wang LC, Kisi O, Hu B, Bilal M, Zounemat-Kermani M, Li H. Evaporation modelling using different machine learning techniques. *Int J Climatol.* 2017; 37: 1076–1092. <https://doi.org/10.1002/joc.5064>

5. Wang X, Fan J, Xing Y, Xu G, Wang H, Deng J, et al. The effects of mulch and nitrogen fertilizer on the soil environment of crop plants. *Adv Agron*. 2019; 153: 121–173. <https://doi.org/10.1016/bs.agron.2018.08.003>
6. Allen RG, Pereira LS, Raes D, Smith M. Crop evapotranspiration—Guidelines for computing crop water requirements—FAO Irrigation and drainage paper 56. 1998:1–15.
7. Kisi O, Zounemat-Kermani M. Comparison of two different adaptive neuro-fuzzy inference systems in modelling daily reference evapotranspiration. *Water Resour Manag*. 2014; 28(9): 2655–2675. <https://doi.org/10.1007/s11269-014-0632-0>
8. Shiri J, Nazemi AH, Sadraddini AA, Landeras G, Kisi O, Fard AF, et al. Comparison of heuristic and empirical approaches for estimating reference evapotranspiration from limited inputs in Iran. *Comput Electron Agric*. 2014; 108: 230–241. <https://doi.org/10.1016/j.compag.2014.08.007>
9. Fan JL, Wu LF, Zhang FC, Xiang YZ, Zheng J. Climate change effects on reference crop evapotranspiration across different climatic zones of China during 1956–2015. *J Hydrol*. 2016; 542: 923–937. <https://doi.org/10.1016/j.jhydrol.2016.09.060>
10. Keshtegar B, Kisi O, Arab H G, Zounemat-Kermani M. Subset modeling basis ANFIS for prediction of the reference evapotranspiration. *Water Resour Manag*. 2018; 32(3): 1101–1116. <https://doi.org/10.1007/s11269-017-1857-5>
11. Karbasi M. Forecasting of multi-step ahead reference evapotranspiration using Wavelet-Gaussian Process Regression model. *Water Resour Manag*. 2018; 32(3): 1035–1052. <https://doi.org/10.1007/s11269-017-1853-9>
12. Tabari H, Grismer ME, Trajkovic S. Comparative analysis of 31 reference evapotranspiration methods under humid conditions. *Irrig Sci*. 2013; 31: 107–117. <https://doi.org/10.1007/s00271-011-0295-z>
13. Wen X, Si J, He Z, Wu J, Shao H, Yu H. Support-vector-machine-based models for modeling daily reference evapotranspiration with limited climatic data in extreme arid regions. *Water Resour Manag*. 2015; 29: 3195–3209. <https://doi.org/10.1007/s11269-015-0990-2>
14. Traore S, Luo Y, Fipps G. Deployment of artificial neural network for short-term forecasting of evapotranspiration using public weather forecast restricted messages. *Agric Water Manag*. 2016; 163: 363–379. <https://doi.org/10.1016/j.agwat.2015.10.009>
15. Antonopoulos VZ., Antonopoulos AV. Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. *Comput Electron Agric*. 2017; 132: 86–96. <https://doi.org/10.1016/j.compag.2016.11.011>
16. Hargreaves GH, Samani ZA. Reference crop evapotranspiration from temperature. *Appl Eng Agric*. 1985; 1: 96–99. <https://doi.org/10.13031/2013.26773>
17. Trabert W. Neue beobachtungen über verdampfungsgeschwindigkeiten. *Meteorol Z*. 1896; 13: 261–263.
18. Priestley CHB, Taylor RJ. On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon Weather Rev*. 1972; 100: 81–92. [https://doi.org/10.1175/1520-0493\(1972\)100<0081:OTAOSH>2.3.CO;2](https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2)
19. Kisi O. Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree. *J Hydrol*. 2015; 528: 312–320. <https://doi.org/10.1016/j.jhydrol.2015.06.052>
20. Wang L, Kisi O, Zounemat-Kermani M, Li H. Pan evaporation modeling using six different heuristic computing methods in different climates of China. *J Hydrol*. 2017; 544: 407–427. <https://doi.org/10.1016/j.jhydrol.2016.11.059>
21. Jovic S, Nedeljkovic B, Golubovic Z, Kostic N. Evolutionary algorithm for reference evapotranspiration analysis. *Comput Electron Agric*. 2018; 150: 1–4. <https://doi.org/10.1016/j.compag.2018.04.003>
22. Landeras G, Ortiz-Barredo A, López JJ. Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). *Agric Water Manag*. 2008; 95: 553–565. <https://doi.org/10.1016/j.agwat.2007.12.011>
23. Kisi O. The potential of different ANN techniques in evapotranspiration modelling. *Hydrol Process* 2008; 22(14): 2449–2460. <https://doi.org/10.1002/hyp.6837>
24. Pour-Ali Baba A, Shiri J, Kisi O, Fard AF, Kim S, Amini R. Estimating daily reference evapotranspiration using available and estimated climatic data by adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN). *Hydrol Res*. 2013; 44(1): 131–146. <https://doi.org/10.2166/nh.2012.074>
25. Sanikhani H, Kisi O, Maroufpoor E, Yaseen Z M. Temperature-based modeling of reference evapotranspiration using several artificial intelligence models: application of different modeling scenarios. *Theor Appl Climatol*. 2019; 135(1–2): 449–462. <https://doi.org/10.1007/s00704-018-2390-z>
26. Tabari H, Kisi O, Ezani A, Talaei PH. SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. *J. Hydrol*. 2012; 444: 78–89. <https://doi.org/10.1016/j.jhydrol.2012.04.007>

27. Kisi O. Modeling reference evapotranspiration using three different heuristic regression approaches. *Agric Water Manag.* 2016; 169: 162–172. <https://doi.org/10.1016/j.agwat.2016.02.026>
28. Abdullah SS, Malek MA, Abdullah NS, Kisi O, Yap KS. Extreme learning machines: a new approach for prediction of reference evapotranspiration. *J Hydrol.* 2015; 527: 184–195. <https://doi.org/10.1016/j.jhydrol.2015.04.073>
29. Feng Y, Peng Y, Cui N, Gong D, Zhang K. Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. *Comput Electron Agric.* 2017; 136: 71–78. <https://doi.org/10.1016/j.compag.2017.01.027>
30. Gocić M, Motamedi S, Shamsirband S, Petković D, Ch S, Hashim R, et al. Soft computing approaches for forecasting reference evapotranspiration. *Comput Electron Agric.* 2015; 113: 164–173. <https://doi.org/10.1016/j.compag.2015.02.010>
31. Yin Z, Wen X, Feng Q, He Z, Zou S, Yang L. Integrating genetic algorithm and support vector machine for modeling daily reference evapotranspiration in a semi-arid mountain area. *Hydrol Res.* 2017; 48: 1177–1191. <https://doi.org/10.2166/nh.2016.205>
32. Ma X, Liu ZB. Predicting the oil production using the novel multivariate nonlinear model based on Arps decline model and kernel method. *Neural Comput Appl.* 2018; 29: 579–591. <https://doi.org/10.1007/s00521-016-2721-x>
33. Wu L, Huang G, Fan J, Zhang F, Wang X, Zeng W. Potential of kernel-based nonlinear extension of Arps decline model and gradient boosting with categorical features support for predicting daily global solar radiation in humid regions. *Energy Convers Manag.* 2019; 183: 280–295. <https://doi.org/10.1016/j.enconman.2018.12.103>
34. Hassan MA, Khalil A, Kaseb S, Kassem MA. Exploring the potential of tree-based ensemble methods in solar radiation modeling. *Appl Energy* 2017; 203: 897–916. <https://doi.org/10.1016/j.apenergy.2017.06.104>
35. Pal M, Deswal S. M5 model tree based modelling of reference evapotranspiration. *Hydrol Process* 2009; 23: 1437–1443. <https://doi.org/10.1002/hyp.7266>
36. Rahimikhoob A. Comparison between M5 model tree and neural networks for estimating reference evapotranspiration in an arid environment. *Water Resour Manag.* 2014; 28: 657–669. <https://doi.org/10.1007/s11269-013-0506-x>
37. Kisi O, Kilic Y. An investigation on generalization ability of artificial neural networks and M5 model tree in modeling reference evapotranspiration. *Theor Appl Climatol.* 2016; 126: 413–425. <https://doi.org/10.1007/s00704-015-1582-z>
38. Feng Y, Cui N, Gong D, Zhang Q, Zhao L. Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. *Agric Water Manag.* 2017; 193: 163–173. <https://doi.org/10.1016/j.agwat.2017.08.003>
39. Shiri J. Improving the performance of the mass transfer-based reference evapotranspiration estimation approaches through a coupled wavelet-random forest methodology. *J Hydrol.* 2018; 561: 737–750. <https://doi.org/10.1016/j.jhydrol.2018.04.042>
40. Fan J, Yue W, Wu L, Zhang F, Cai H, Wang X, et al. Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China. *Agric For Meteorol.* 2018; 263: 225–241. <https://doi.org/10.1016/j.agrformet.2018.08.019>
41. Mehdizadeh S, Behmanesh J, Khalili K. Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. *Comput Electron Agric* 2017; 139: 103–114. <https://doi.org/10.1016/j.compag.2017.05.002>
42. Shiri J. Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in hyper-arid regions of Iran. *Agric Water Manag.* 2017; 188: 101–114. <https://doi.org/10.1016/j.agwat.2017.04.009>
43. Zhao SQ. A new scheme for comprehensive physical regionalization in China. *Acta Geogr Sin.* 1983; 38: 1–10. <https://doi.org/10.11821/xb198301001>
44. Yao N, Li Y, Lei T, Peng L. Drought evolution, severity and trends in mainland China over 1961–2013. *Sci Total Environ.* 2018; 616: 73–89. <https://doi.org/10.1016/j.scitotenv.2017.10.327> PMID: 29107781
45. Specht DF. A general regression neural network. *IEEE Trans Neural Netw.* 1991; 2(6): 568–576. <https://doi.org/10.1109/72.97934> PMID: 18282872
46. Wang L, Kisi O, Zounemat-Kermani M, Salazar GA, Zhu Z, Gong W. Solar radiation prediction using different techniques: model evaluation and comparison. *Renew Sustain Energy Rev* 2016; 61: 384–397. <https://doi.org/10.1016/j.rser.2016.04.024>
47. Jang JS. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern.* 1993; 23: 665–685. <https://doi.org/10.1109/21.256541>

48. Quinlan JR. Learning with continuous classes. In 5th Australian joint conference on artificial intelligence, 1992; 92: 343–348.
49. Wang Y, Witten IH. Induction of model trees for predicting continuous classes. In: Proceedings of the Poster Papers of the European Conference on Machine Learning, University of Economics, Faculty of Informatics and Statistics, Prague, 1997.
50. Chen T, He T, Benesty M, Khotilovich V, Tang Y. Xgboost: extreme gradient boosting. R Packag. version 0.4–2 1–4, 2015.
51. Vapnik V. The nature of statistical learning theory. Springer science & business media, 2003.
52. Ma X, Xie M, Wu W, Zeng B, Wang Y, Wu X. The novel fractional discrete multivariate grey system model and its applications. *Appl Math Model*. 2019; 70: 402–424. <https://doi.org/10.1016/j.apm.2019.01.039>
53. Ma X, Liu Z. The GMC (1, n) model with optimized parameters and its application. *J grey Syst*. 2017; 29(4): 122–138.
54. Ma X. A brief introduction to the Grey Machine Learning; 2018. Preprint. Available from: arXiv1805.01745.
55. Friedman JH. Multivariate adaptive regression splines. *Ann Stat*. 1991; 19(1): 1–67. <https://doi.org/10.1002/9780470057339.vaa008>
56. Fan J, Wu L, Zhang F, Cai H, Ma X, Bai H. Evaluation and development of empirical models for estimating daily and monthly mean daily diffuse horizontal solar radiation for different climatic regions of China. *Renew Sustain Energy Rev*. 2019; 105: 168–186. <https://doi.org/10.1016/j.rser.2019.01.040>
57. Fan J, Wu L, Zhang F, Cai H, Zeng W, Wang X, et al. Empirical and machine learning models for predicting daily global solar radiation from sunshine duration: A review and case study in China. *Renew Sustain Energy Rev*. 2019; 100: 186–212. <https://doi.org/10.1016/j.rser.2018.10.018>
58. Guo X, Sun X, Ma J. Prediction of daily crop reference evapotranspiration (ET<sub>0</sub>) values through a least-squares support vector machine model. *Hydrol Res*. 2011; 42: 268–274. <https://doi.org/10.2166/nh.2011.072>
59. Fan J, Wang X, Wu L, Zhou H, Zhang F, Yu X, et al. Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China. *Energy Convers Manag*. 2018; 164: 102–111. <https://doi.org/10.1016/j.enconman.2018.02.087>