Modelling Reference Evapotranspiration using a Novel Regression-based Method: Radial basis M5 Model Tree

Ozgur Kisi
Ilia State University
Behrooz Keshtegar (beh.keshtegar@tdtu.edu.vn)
Ton Duc Thang University
Mohammad Zounemat-Kermani
University of Kerman
Salim Heddam
Agronomy Department, Hydraulics Division, University
Nguyen-Thoi Trung
Ton Duc Thang University

Research Article

Keywords: Reference evapotranspiration, Modelling, Radial basis M5 model tree, Response surface method

Posted Date: February 18th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-208205/v1

Modelling reference evapotranspiration using a novel regression-based method: Radial basis M5 model tree

Ozgur Kisi[1,2], Behrooz Keshtegar[3*], Mohammad Zounemat-Kermani[4], Salim Heddam[5], Nguyen-Thoi Trung[6,7]

[1] Department of Civil Engineering, School of Technology, Ilia State University, Tbilisi, Georgia
[2] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
[3]Department of Civil Engineering, Faculty of Engineering, University of Zabol, P.B. 9861335856, Zabol, Iran
[4]Department of Water Engineering, Shahid Bahonar University of Kerman, Kerman, Iran
[5]Faculty of Science, Agronomy Department, Hydraulics Division, University 20 Août 1955, Route El Hadaik, BP 26, Skikda, Algeria.
[6]Division of Computational Mathematics and Engineering, Institute for Computational Science, Ton Duc Thang University, Ho Chi Minh City, Vietnam
[7]Faculty of Civil Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam

[*] Corresponding author; **Email: bkeshtegar@uoz.ac.ir**

Abstract

In the current study, an ability of a novel regression-based method is evaluated in modelling daily reference evapotranspiration (ET_0), which is an important issue in water resources management plans and helps farmers in irrigation planning. The method was developed by hybridizing radial basis function and M5 model tree and called as radial basis M5 model tree (RM5Tree). The radial-based kernel function was used to control the input variables in modelling process of M5 model tree. The new model results were compared with traditional M5 model tree (M5Tree), response surface method (RSM) and two neural networks (multi-layer perceptron neural networks, MLPNN & radial basis function neural network, RBFNN) with respect to several statistical indices. Daily climatic data (relative humidity, *RH*, solar radiation, *SR*, wind speed, air temperature, *T*) recorded at three stations in Turkey, Mediterranean Region, were used. The effect of each weather data on ET_0 was also investigated by utilizing three different input scenarios with various combinations of input variables. On the whole, the RM5Tree provided the best results (NSE>0.997) followed by the MLPNN (NSE>0.990), and M5Tree (NSE>0.945) in modelling daily ET_0. The *SR* was observed as the most effective input parameter on ET_0 which was followed by the *T* and *RH*. However, the findings of the third modelling scenario revealed that taking into account of all variables would considerably increase models' accuracies for the three stations.

Keywords: Reference evapotranspiration; Modelling; Radial basis M5 model tree; Response surface method

1. Introduction

Evapotranspiration (ET) is one of main components in hydrological cycle and accurate estimation of ET has vital importance in design and management of the irrigation systems, water resources studies, and other similar cases. Knowing the ET rate of a plant can help determine the accurate amount of water required for irrigation, which will subsequently lead to increased productivity. Failure in determining the accurate ET rate may lead to the overestimation of plants' water requirement, which will consequently cause adverse effects such as waterlogged lands, soil nutrients washout, as well as contamination of groundwater resources. On the other hand, underestimation of plants' water requirement will incur moisture tension on them, followed consequently by a reduced crop yield. In this regard, equations such as FAO Penman, FAO Penman-Monteith, Blaney-Criddle, etc. can be used to study the reference evapotranspiration (ET_0) of plants (Feng et al. 2017; Chen et al. 2019). However, despite their acceptable performance in most of the cases, using these methods needs the access to a large amount of input data, which is not of course always possible regarding the conditions within various regions.

In such conditions, using indirect approaches such as soft computing techniques and artificial intelligence based (AI-based) models can be a proper alternative modelling solution (Zounemat-Kermani et al., 2020). Lots of researchers have applied artificial intelligence (AI)-based models including artificial neural networks (e.g., MLPNN & RBFNN), fuzzy logic concepts (e.g. ANFIS), regression and classification tree models (e.g. M5 tree & CHAID) and machine learning approaches (SVM & SVR) for simulating and modelling ET_0 (Kişi and Öztürk, 2007; Shiri et al., 2014; Shiri et al. 2015; Gocić et al., 2015; Yin et al. 2017; Mehdizadeh et al. 2017; Dou and Yang 2018; Zounemat-Kermani et al., 2019; Chia et al., 2020; Chen et al., 2020; Yamaç and Todorovic, 2020; Adnan et al., 2021). In the following paragraphs some of the most recent pertinent studies about modelling ET using AI-based models are presented. The neuro-fuzzy and artificial neural network (ANN) were compared for modelling ET_0 with two input combinations to select suitable input data via Shari et al. (2015) and they reported that the local training can be applied to validate the alternative modelling by using AI. Kisi (2016) used three different models consisting multivariate adaptive regression splines, M5Tree and least square SVR for approximating the ET_0. In general, it was found that the M5Tree performed superior to the other modelling approaches applied. Rahimikhoob (2016) compared the ability of artificial neural network and M5Tree

for estimating ET_0 of an arid area. The study did not report the complete superiority of the utilized models.

Khoshravesh et al. (2017) analysed performances of multivariate fractional polynomial model, Bayesian and robust regressions to estimate ET_0 in arid climates. Outcomes of the study showed that accuracy of multivariate fractional polynomial model is better than other two models. Daily ET predictions were investigated using extreme learning machine (ELM) and generalized regression NN (GRNN) with input data of temperature at 6 stations of China by Feng et al. (2017). It is conducted that ELM is the robust and accurate models compared to GRNN. Three different AI-based models (ANFIS-GP, fuzzy genetic model and M5Tree) were used and executed in modelling monthly ET_0 values by Wang et al. in 2017. It was reported that the fuzzy genetic model performed better in comparison to the ANFIS-GP and M5Tree models. The hybrid genetic algorithm (GA) and SVM as an AI-based model was used for simulating daily ET_0 of semi-arid environment in northwest China by Yin et al. (2017) and it was compared with SVM and ANN based on eight different combinations of climatic input data set. Based on their results, the SVR combined by GA model had superior performances. Antonopoulos and Antonopoulos (2017) compared ANN-based multi-layer feed forward back-propagation with several empirical models for calculation of daily ET_0 in northern Greece and showed that the ANN with a sigmoid transfer function in hidden layer can provide more accurate predictions than the empirical models.

In another study, Keshtegar et al. (2018) explored the ability of subset ANFIS, ANFIS, ANNs and M5Tree. They claimed that the subset ANFIS is superior to other applied methods in modelling daily ET_0. Gavili et al. (2018) compared the ability of ANN, ANFIS and gene expression programming (GEP) in modelling daily ET_0. The results attained from the AI-based models were compared with those of the empirical models. Comparing the results, it was revealed that the AI-based models provided better accuracies compared to empirical models. Dou and Yang (2018) investigated and compared the feasibilities and abilities of four AI-based models using ELM, ANFIS, ANN and SVR for prediction of daily ET_0 for four sites in China. It conducted that the AI-based ANFIS and ELM models can produce better performances compared to the ANN and SVM while the ELM model was considerably reduced computational time in modelling process. Sanikhani et al. (2018) investigated six AI-based methods for modeling ET_0 for Antalya and Isparta stations in Turkey. At Isparta station, the ANFIS-SC and RBNN models had superior results than other AI-based models, while the performance of GEP and GRNN models was conducted better than others at the Antalya station. Keshtegar et al. (2019) compared the ability of RSM and

polynomial chaos expansion (PCE) in modelling ET_0. They reported that the PCE model was more accurate approach to estimate daily ET_0. The abilities of four learning algorithms as multilayer perceptrons-based deep feed-forward ANN, gradient-boosting machine, random forest regression (RF) using M5Tree model and generalized linear model were compared for ET_0 estimations for the Punjab Northern India stations by Saggi and Jain (2019). By extracted results from Saggi and Jain (2019), the deep feed-forward ANN performed better than the other models. Sanikhani et al., (2019) employed six AI-based models namely multilayer perceptron ANN, GRNN, radial basis neural networks, integrated ANFIS with subtractive clustering and grid partitioning, and GEP for modeling ET_0 with small number of input climatic data. In general, it was reported that all the applied models have highly practical and reliable performances for investigated stations. Heddam et al. (2018) applied and compared three evolving connectionist (ECoS) models namely, offline-based dynamic evolving neural-fuzzy inference systems named DENFIS-OF,(*ii*) online-based dynamic evolving neural-fuzzy inference systems named DENFIS-ON, and (*iii*) the evolving fuzzy neural network called (EFuNN), for modelling daily ET_0 in the northern region of the Algeria. According to the obtained results, the best accuracy was obtained using the DENFIS-OF model. Tao et al. (2018) proposed a hybrid model called adaptive neuro-fuzzy inference systems (ANFIS) with firefly algorithm (ANFIS-FA) for predicting daily ET_0 at Burkina Faso and reported that the hybrid ANFIS-FA provides higher accuracy compared to the standard ANFIS. Karbasi (2018) employed the Gaussian process regression (GPR) for forecasting daily ET_0 and demonstrated that the wavelet decomposition significantly improved the performances of the models. Fang et al. (2018) employed the RF, SVM and MLR for predicting monthly ET_0 in China, and demonstrated that the SVM was more accurate. The accuracy of models to predict the ET_0 is one of challenges in hydrology field to manage irrigation systems and water resources.

More recently, Zhu et al. (2020) employed a hybrid extreme learning machine (ELM) with the particle swarm optimization (PSO) model for daily ET_0 prediction. They claimed that the PSO-ELM model offered the best accuracy among other applied models such as ANN and random forest (RF) models. Nagappan et al. (2020) attempted to predict ET_0 for irrigation scheduling using machine learning methods like deep learning neural network (DLNN) and RBNN. It was found that the DLNN model acted better in the prediction process.

In a similar study, Ferreira and Cunha (2020) showed that deep learning performed slightly better than ANN and RF in predicting ET_0. Salam and Islam (2020) compared various data-driven models in ET_0 prediction. They utilized standard SVM model as well as ensemble learning models for the prediction process including the bagging random tree (RT), RF, and

random subspace (FS) models. The findings showed that the RT model performed superior followed by the RF, RS, and SVM.

Generally, the AI-based neural networks or machine learning approaches are used to provide the accurate prediction of ET_0 due to flexible ability for providing the nonlinear relations. However, the AI-based data-driven approaches have some limitations as; i) input data variables highly affect their predictions, ii) some control parameters are required to train the models and iii) training process to provide a model is time consuming. The regression-based models are the efficient modelling approaches with simple regressed process. However, the regression-based data driven approaches have some drawbacks including i) the regressed function is important for accurately predictions, the highly nonlinearity of the input data are neglected in the regression process and iii) the linear cross -correlation between the input-output data is used in the modelling process. Consequently, the efficient and accurate modelling approach using the regression-based models that is free from a complex training process is needed in modelling of ET_0.

The nonlinear mapping with efficient regression region can provide a flexible nonlinear response with efficient modelling process for machine learning models. The input data for training of the M5Tree models can be controlled based on the nonlinear maps using radial basis function. Thus, the nonlinearity of input variables of the response can be considered applying the radial map. Therefore, it can improve the nonlinear functions for accurate ET_0 predictions using M5Tree. In this study, the nonlinear forms of the input data are improved for M5Tree-based regressed model using the radial sample data set. This map controls the input data and it can provide high-ability to provide nonlinear cross-correlation between input-output data set. The proposed model was tested at three climate stations in Turkey (Isparta, Antalya & Adana stations). Afterwards, the performance of the RM5tree model will also be compared with the neural network models, multi-layer perceptron and radial basis function neural network and two regression methods, response surface method and M5Tree. The results showed that the proposed model has the fixable ability for nonlinear response compared to M5Tree models while by increasing the input data and the accuracy of prediction using RM5Tree models are significantly improved compared to other studied models.

2. Materials and methodology

2.1. Case study

The study used daily weather data comprising relative humidity (*RH*), solar radiation (*SR*), air temperature (*T*), and wind speed (*W*) from Adana (longitude 35° 19' E, latitude 37° 00' N with an altitude of 27 m) Antalya (latitude: 36° 42′ N, longitude: 30° 44′ E with an altitude of 47 m) and Isparta (longitude: 30° 34′ E, latitude: 37° 47′ N with an altitude of 997 m) stations with Mediterranean Region, Turkey **(Figure 1)**. Data cover a period from 01 January of 1972 to 31 December of 2002 for Adana station, from 01 January of 1973 to 31 October of 2002 for Antalya station and from 01 September of 1978 to 31 October of 2002 for Isparta station. There are no missing values in the used data. **Table 1** shows the statistic characteristics of the dataset in terms of minimum (Min), maximum (Max), standard deviation (Std.), mean, coefficient of variation (CV) and correlation between the input parameters and the output parameter (ET_0). **Table 1** implies that the Adana and Antalya stations are more similar in terms of temperature ranges. Isparta is the only station that recorded minus temperatures in the dataset used in this study. In addition, the solar radiation is the most correlated parameter followed by air temperature with ET_0 in all of the three stations. In the applications, data splitting rule of 65%-35% was applied to train and test the studied models.

2.2. Modelling approaches

2.2.1. Response surface method

The RSM is commonly implemented for modelling the nonlinear relations as below (Hill and Hunter 1966):

$$\hat{ET_o} = a_0 + \sum_{i=1}^{NV} a_i x_i + \sum_{i=1}^{NV}\sum_{j=i}^{NV} a_{ij} x_i x_j \quad (1)$$

where, $\hat{ET_o}$ is the predicted ET_0, NV is the number of input variables x including the mean temperature i.e. T_{mean} (°C), solar redation i.e. *SR* (langley), relative humidity i.e. RH_{mean} (%), and wind speed i.e. *W* (m/s). a_0, a_i and a_{ij} are unknown coefficients for polynomial terms of Eq. (2). Generally, the unknown coefficients are calibrated based on the ordinary least square estimator as follows (Keshtegar and Kisi 2017):

$$\boldsymbol{a} = [P(X)^T P(X)]^{-1} [P(X)^T ET_o] \quad (2)$$

Where, $P(\boldsymbol{X})$ is the polynomial basic function which is determined based on input data in training stage (65% total of data) that it extracted to give more details from Keshtegar and Heddam (2017); Keshtegar et al. (2021):

2.2.2. Multilayer perceptron artificial neural networks (MLPNN)

Artificial neural networks (ANN) are a black box models possessing the capabilities to produces a suitable response from an external stimulus, and they are composed of two items: the neurons and the weights. The ANN models are constructed in two distinguished phases: the forward and the backward phases, the two, successively achieved during the backpropagation training algorithm. The knowledge information available in the predictor variables is transferred from the input neurons to the hidden neurons via the weight, and then summed to get an estimate of the total stimulus of each hidden unit (Ozonoh et al. 2020). The hidden neurons send the collected information to the output neuron through an activation function, generally the sigmoid (Shahabinejad et al. 2020). Finally, the output neuron provides a response, then, compared to the desired value, and the error expected is calculated. The MLPNN having an input, one hidden and one output layers is the well-known ANN architecture (Figure 2), was employed in the present study, and trained with supervised Levenberg Marquart (LM) learning algorithm. According to Figure 1, the relationship between N possible inputs variables (x_i: climatic variables) and one output variable (ET_0) were created as follow (Zhu et al. 2021):

$$Y = \phi_2\left[\beta_0 + \sum_{j=1}^{n} w_{jk}\left[\phi_1\left(\sum_{j=1}^{n} x_i w_{ij} + \beta_j\right)\right]\right] \quad (3)$$

where x_i is an input variable, w_{ij} is the weight between the input i and the hidden neuron j, β_j is the bias of the hidden neuron j, ϕ_1 the activation sigmoid function, w_{jk} is the weight of connection of neuron j in the hidden layer to unique neuron k in the output layer, β_0 is the bias of the output neuron k (Wang et al. 2020;)

2.2.3. Radial basis function neural network (RBFNN)

Radial basis function neural network (RBFNN) belongs to the category of feedforward neural network (FFNN). Contrary to the well know MLPNN, the RBFNN possess only one hidden layer with large number of neurons, and each one implements a radial basis function, generally the Gaussian function (Tenenbaum et al. 2020). The first input layer transfers the predictor variables to the hidden layer directly and the only output neuron linearly combines the weighted results of all hidden units (Figure 3). The Gaussian activation function can be expressed as follows (Chen et al. 2019):

$$\phi\left(x_k, \mu_i\right) = \exp\left(-\frac{1}{2\sigma_i^2}\sum_{m=1}^{M}\left(x_k - \mu\right)\right) \quad (4)$$

Where $\mu_i = [\mu_{i1}, \mu_{i2}, \ldots, \mu_{iM}]$ is the center of Gaussian function, $x_k = [x_{k1}, x_{k2}, \ldots, x_{kN}]$ is a training sample and σ^2 is the width of the RBFNN neuron, also called the spread width. During the training process, the optimal number of RBFNN neurons, the values of the centers, the weights and biases were determined by minimizing the mean squared errors between observed and modelled values of ET_0 (Zhou et al.2012; Bonanno et al.2012)

2.3.4. M5 tree model

M5 tree model is subset basis data mining and machine learning method. The tree-based methods are indeed a part of data mining methods, the output of which resulted from application of the input and output data will be a model with tree structure (Solomatine and Xue, 2004). The tree models are fundamentally based on the decision and dominance method. Substituting the linear regression equation at the nodes is a method executed in the M5 model, which is capable to predict or estimate the numerical variables. Structure of a decision tree is similar to a tree constituted of the root, branches, nodes, and leaves. A tree model is built up in two steps. Accordingly, in first step, the decision tree is designed by data splitting. The split criterion in M5 model is to maximize the reduction of standard deviation (SDR) of the data at the offspring node. When no reduction of standard deviation of the data at the offspring node is possible, its parent node will not be split and, thus, reach the end node or leaf. The following formula is used to calculate SDR:

$$SDR = sd(T) - \sum_{i=1}^{N} \frac{T_i}{T} sd(T_i) \quad (5)$$

where T represents a set of the samples entering on each node, T_i represents a subset of the samples with the i^{th} result of the potential test, and Sd is standard deviation of the input data, which can be calculated as follows:

$$Sd(T) = \sqrt{\frac{1}{N}(\sum_{i=1}^{N}(y_i)^2 - \frac{1}{N}(\sum_{i=1}^{N} y_i)^2} \quad (6)$$

Yi is a numerical value of the target attribute of sample i and N indicates the number of data. Since the process of branching (classification) at offspring nodes has less standard deviation than the parent nodes, they have more accurate results and are featured with higher homogeneity. Once all the possible classifications are examined, the M5 model selects the one with minimum expected error. However, the second step in designing a tree model is to shrink the overgrown and overfitted tree through pruning the branches and replacing them with linear regression functions (Rahimikhoob, 2016).

3. Radial M5 model tree

To enhance the accuracy of ET_0 estimations basis M5Tree model, radial basis M5Tree (RM5Tree) is introduced. The input data set is controlled by applying the radial input data in feature space. By transferring input data from original to radial map, the radial basis function (RBF) is applied in RM5Tree as follows (Chen et al. 1991; Xiao et al. 2020; Zhang et al. 2020a):

$$K_{ij} = \phi(\left\|\boldsymbol{Z}_i - \boldsymbol{C}_j\right\|, \varepsilon) = \exp(-\varepsilon\left\|\boldsymbol{Z}_i - \boldsymbol{C}_j\right\|^2) \;\; i = 1,..., NV \;\; j = 1,..., RF \tag{7}$$

where, RF is number of radial sets with shape factor of ε, and $\boldsymbol{C}$ represents the center of RBF. Z is normalized map, which is given as follows (Zhang et al. 2019):

$$Z = \frac{X - \mu_x}{\sigma_x} \tag{8}$$

where, μ_x and σ_x are respectively mean and SD of input data *x*. The radial transformations which are given by Eq. (7) with *C*=0 and ε=0.5, are shown in the Fig. 4 that shows a nonlinear map. Thus, there can be utilized a new data set to train a model by transforming data set from *NV* (X-space) to RF (radial-space).

Two parameters of shape and location as center points applied in RBF are selected as ε=0.5 and C=[$\boldsymbol{X}_{min}$ $\boldsymbol{X}_{max}$] which are randomly given from the domain of input dataset with RF as 5, 10, 20 and 50 in this study. The schematic structure of RM5Tree is presented in Fig. 5. This model involves three layers as input, transferring and modelling layers. Using Eq. (7), the input dataset is normalized in the first (input) layer, while the RF-dataset is provided by transferring data in the second layer as follows:

1) Create RF- center point from domain of each input data, randomly.

2) Transfer the input data set in layer 1 by using RBF in Eq. (7) into radial space based on the RF- center point as follows:

$$\boldsymbol{Z} = \begin{bmatrix} z_{1,1} & z_{1,2} & \Lambda & z_{1,NV} \\ z_{2,1} & z_{2,2} & \Lambda & z_{2,NV} \\ \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} \\ z_{N,1} & z_{N,2} & \Lambda & z_{N,NV} \end{bmatrix} \rightarrow \boldsymbol{K} = \begin{bmatrix} K_{1,1} & K_{1,2} & \Lambda & K_{1,RF} \\ K_{2,1} & K_{2,2} & \Lambda & K_{2,RF} \\ \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} \\ K_{N,1} & K_{N,2} & \Lambda & K_{N,RF} \end{bmatrix} \tag{9}$$

where, *N* is number of data in the training stage as 65% of total data, number of input variables and number of radial input data and $\boldsymbol{K}_{ij} \;\; i = 1,2,..., N \;\; j = 1,2,..., RF$. The radial input data is used as input data set in the training phase for M5Tree models. Therefore, the applied

nonlinear map using Gaussian function and the number of canter points improves the accuracy of M5Tree models.

4. Application of the models

4.1. Modelling scenarios

Based on the results of **Table 1** given for the correlation coefficients between the independent variables (*T, RH, SR & W*) and dependent variable (ET_0), three different modelling scenarios for constructing the machine learning methods (M5Tree, Radial M5Tree, MLPNN, RBFNN & RSM) are considered. These scenarios are tabulated in **Table 2**; in the first scenario, just one input parameter is considered for modelling ET_0 including i) T_{mean}; ii) *W*; iii) *SR*; iv) *RH*. The second scenario takes into account the most correlated parameters including *v*) T_{mean}, *SR* and, *vi*) T_{mean}, *SR, RH* and, vii) T_{mean}, *SR, RH*. Finally, the third scenario has all of the independent parameters as *vii*) T_{mean}, *SR, RH, W*.

4.2. Evaluation of the models

The models' accuracies were compared according to the mean absolute error (*MAE*), determination coefficient (R^2), root mean square error (*RMSE*), agreement index (*d*) and Nash and Sutcliffe efficiency (*NES*) statistics (Xiao et al. 2019; Zhang et al. 2020b).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|(ET_0)_i - (ET_p)_i\right| \tag{14}$$

$$RMSE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}[(ET_0)_i - (ET_p)_i]^2} \tag{15}$$

$$d = 1 - \frac{\sum_{i=1}^{N}[(ET_0)_i - (ET_p)_i]^2}{\sum_{i=1}^{N}[\left|(ET_0)_i - \overline{ET}_0\right| + \left|(ET_p)_i - \overline{ET}_0\right|]^2},\ 0 < d \le 1 \tag{16}$$

$$NES = 1 - \frac{\sum_{i=1}^{N}[(ET_0)_i - (ET_p)_i]^2}{\sum_{i=1}^{N}[(ET_0)_i - \overline{ET}_0]^2},\ -\infty < NES \le 1 \tag{17}$$

$$R^2 = \left[\frac{\frac{1}{N}\sum\left((ET_0)_i - \overline{ET}_0\right)\left((ET_p)_i - \overline{ET}_P\right)}{\sqrt{\frac{1}{N}\sum_{i=1}^{n}\left((ET_0)_i - \overline{ET}_0\right)^2}\sqrt{\frac{1}{N}\sum_{i=1}^{n}\left((ET_p)_i - \overline{ET}_p\right)^2}}\right]^2,\ 0 < R^2 \le 1 \tag{18}$$

In which, N is the number of data, ET_0, ET_p, $\overline{ET}_0$ $\overline{ET}_p$ are the FAO-56 PM ET_0, predicted ET_0 mean ET_0, and mean predicted ET_0, respectively.

5. Results and discussion

5.1. Result analysis for the Isparta Station

The final results of the investigated AI-based models (RSM, M5Tree and RM5Tree) in terms of training and testing results for Isparta Station can be seen in **Table 3**. It can be seen that the RM5Tree model performs superior to the M5Tree and RSM models with respect to various criteria in all input combinations (Scenario III). In testing phases, the RMSE is improved (d) as accuracy (tendency) factors using proposed RM5Tree about 42% (6%) and 15% (2%) for Scenario I, 75% (15%) and 60% (3%) for Scenario II and 105% (1%) and 90% (1%) for Scenario III compared to M5tree and RSM models, respectively. Considering Scenario (I) implies that among the single input variables *SR* is the most effective parameter on ET_0 followed by T_{mean} and *RH*, respectively while *W* has the least effect. This result was actually expected according to the calculated correlation coefficients in **Table 3**. In the second scenario (II), including the combination of *SR* (T_{mean}) parameter with T_{mean} (*SR*) considerably improves the models' accuracy. For example, it improved the MAE, RMSE and NES of RM5Tree by 42% (13%), 43% (19%) and 34% (7%), respectively. Adding *RH* parameter to T_{mean} and *SR* inputs also increases the accuracy of the employed models. For example, the values of MAE and RMSE of RM5Tree were decreased from 0.455 and 0.551 mm to 0.215 and 0.32 mm by 52% and 42%, respectively. In scenario III - even though *W* seems to be the least effective parameter from the first four input combinations - adding *W* parameter to other three inputs considerably increases the increments in MAE and RMSE of the models (MAE and RMSE of the RM5Tree by 82% and 84%, respectively). According to the results of scenario (III), the accuracy of M5Tree model with respect to MAE and RMSE was improved by 43% and 51% using RM5Tree, respectively.

5.2. Result analysis for the Antalya Station

Table 4 reports the comparative results of the models in estimating ET_0 of Antalya Station. Similar to Isparta, RM5Tree model outperforms the other models. From the first scenario (categories *i* to *iv*), the effective parameters (from most to least) in modelling ET_0 are *SR*, T_{mean}, *RH* and *W*. The accuracy of the RM5tree with respect to MAE, RMSE and NES is improved up to 10% (15%), 48% (39%) and 50% (27%) by adding the *SR* input, respectively. Similarly, importing *RH* to T_{mean} and *SR* inputs decreases the MAE and RMSE of RM5Tree

by 73% and 42%, respectively. Moreover, including *W* input in T_{mean}, *SR and RH* combination considerably increases the RM5Tree accuracy (MAE and RMSE are decreased by 87% and 89%) in scenario (III).

5.3. Result analysis for the Adana Station

Table 5 compares the training and testing statistics of the three methods for Adana Station. Similar to the Isparta, in this station the RM5Tree model gave the best accuracy in modeling ET_0 with respect to various evaluation statistics. According to the single input combinations in scenario (I), the most effective variables on ET_0 is *SR* followed by T_{mean} and *RH*. Using *SR* (T_{mean}) parameter with T_{mean} (SR) input improves the RM5Tree accuracy with respect to MAE, RMSE and NES by 35% (11%), 40% (12%) and 43% (5%) in the test period, respectively. Including *RH* variable as an input factor to the RM5Tree comprising T_{mean} and *SR* inputs decreases the MAE and RMSE of the model from 0.553 mm and 0.671 mm to 0.167 mm and 0.325 mm by 70% and 50%. Similarly, importing *W* parameter to three inputs (T_{mean}, *SR* and *RH*) in the third scenario considerably increases RM5Tree accuracy; from 0.167 mm to 0.038 (MAE) and from 0.325 mm to 0.122 mm (RMSE), respectively.

5.4. Discussion

The ET_0 estimates of the M5Tree, RSM and RM5Tree models are compared in **Figs. 6-12** for each station and each input combination. The effect of each variable on ET_0 can be seen from the figures visually. Comparison of **Figs. 6-7** indicates that the effective ranks of the variables (from the most to the least) are *SR,* T_{mean}, *RH and W*. Jain et al. (2008) also reported the same trend for the effective parameters (*SR,* T_{mean}, *RH, W* and lastly dew point temperature) by using hourly data of ET_0 for a few stations in the Reynolds Creek Experimental Watershed in South-western Idaho, USA. In addition, the effect of each variable on ET_0 can also be observed from the Figs. 8-9. Comparison of **Figs. 10-11** shows the considerable effect of *W* variable even though this cannot be seen when W is used as input alone. One input model cannot catch the relationship between *W* and ET_0. All these indicate the necessity of this variable in accurately modelling of ET_0. It should be noted that the M5Tree model estimates are not accurate in Adana compared to other stations and methods. The reason of this might be the fact that the relationship between inputs and output is more non-linear in Adana compared to others and the M5tree model having linear nature might not adequately map this

highly non-linear relationship. **Table 6** compares the results of the best RM5tree model with two of the most prevailing AI-based models of MLPNN and RBNN (multi-layer perceptron neural network and radial basis neural network). It can be concluded that all the AI-based models acted better by considering all the input variables considering scenario III (with the exception for the RBNN in Adana Station). Although the MLPNN model gave better results than the RBNN models but it could not surpass the performance of the proposed RM5tree model. Having a better diagnostic analysis of the efficiency of the all AI-based models (M5Tree, RM5Tree, RSM, MLPNN & RBNN), the results of the best input category in scenarios I, II and III in terms of RMSE (mm) are shown in **Fig. 13** using radar charts. Obviously, the smaller size of stars with lower values for RMSE would indicate the better performance of the models. It can be easily seen that involving all the variables (*T, SR, RH, W* in scenario III) would result in lower values of RMSE (with an exception for the RBNN in Adana Station). This major finding is supported by the outcomes of different AI-based model in a similar study done by Kisi (2006). Further evaluation was achieved using the Taylor diagram to check the performances of the models (**Fig. 14**). At all stations RM5Tree performs better than the other models, it is clear evidence that the proposed approach improves the accuracy of the M5Tree model. Finally, to further compare the accuracy of the models all the results using the best input combination for each model has been considered using the Box plot. Box plots corresponding to the test data (**Fig. 15**) clearly shows that the accuracy of the RM5Tree model was higher than the other models.

6. Conclusion

In the presented work, the applicability of a new method which is developed by combining radial basis function and M5Tree methods is investigated in modeling ET_0. The new method was compared with standard M5Tree, RSM, MLPNN and RBNN using daily climatic data from three stations located in Turkey. Various input combinations of available data were tried to see the effect of each input variable on ET_0. The following conclusions were derived from the applications.

i- The comparison of methods revealed that the new proposed method, RM5Tree, provided better ET_0 estimates than the MLPNN, RBNN, M5Tree and RSM. The accuracy of M5Tree models was considerably improved (more than 30% with respect to MAE and RMSE) by using RM5Tree.

ii- The results obtained based on different input combinations indicated that the most effective variable on models' accuracy in estimating ET_0 was solar radiation followed

by the air temperature, relative humidity and wind speed. However, it was also observed that using wind speed together with other three inputs considerably increases models' efficiency (more than 80% with respect to MAE and RMSE).

iii- The study showed that the proposed RM5Tree model could be utilized as a better alternative to the M5Tree model in modeling daily ET_0.

iv- This ability of this method can be compared with other stations or this method can be applied for other hydrological problems in future.

Acknowledgments

This work was supported by University of Zabol under Grant No. UOZ-GR-9618-1 and UOZ-GR-9719-1.

Authorship contribution

Ozgur Kisi: Conceptualization, Investigation, Writing - original draft, Writing - review & editing. Behrooz Keshtegar: Conceptualization, Investigation, Writing - original draft, Writing - review & editing. Mohammad Zounemat-Kermani: Conceptualization, Investigation, Writing - review & editing. Salim Heddam: Conceptualization, Writing - original draft & editing. Nguyen- Thoi Trung: Conceptualization, Writing - review & editing.

Conflict of Interest:

The authors declare that they have no conflict of interest.

Availability of data and martial:

The daily weather data from Adana, Antalya and Isparta stations, Turkey.

References

Antonopoulos, V. Z., & Antonopoulos, A. V. (2017). Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. *Computers and Electronics in Agriculture*, 132, 86-96.

Bonanno, F., Capizzi, G., Graditi, G., Napoli, C., Tina, G.M. (2012). A radial basis function neural network based approach for the electrical characteristics estimation of a photovoltaic module. *Applied Energy*, 97, 956-961.

Chen, Z., Yang, X., Liu, X. (2019). RBFNN-based non-singular fast terminal sliding mode control for robotic manipulators including actuator dynamics. *Neurocomputing*, 362, 72-82.

Chen, S., Cowan, C.F., Grant, P.M. 1991. Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Transactions on Neural Networks*, 2:302-9.

Chen, X., Li, F. W., Wang, Y. X., Feng, P., & Yang, R. Z. (2019). Evolution properties between meteorological, agricultural and hydrological droughts and their related driving factors in the Luanhe River basin, China. *Hydrology Research*. https://doi.org/10.2166/nh.2019.141

462 Dou, X., Yang, Y. (2018). Evapotranspiration estimation using four different machine
463 learning approaches in different terrestrial ecosystems. *Computers and Electronics in*
464 *Agriculture*, 148, 95-106

465 Feng, Y., Gong, D., Mei, X., & Cui, N. (2017). Estimation of maize evapotranspiration using
466 extreme learning machine and generalized regression neural network on the China Loess
467 Plateau. *Hydrology Research*, 48(4), 1156-1168.

468 Gavili, S., Sanikhani, H., Kisi, O., & Mahmoudi, M. H. (2018). Evaluation of several soft
469 computing methods in monthly evapotranspiration modelling. *Meteorological*
470 *Applications*.

471 Gocić, M., Motamedi, S., Shamshirband, S., Petković, D., Ch, S., Hashim, R., & Arif, M.
472 (2015). Soft computing approaches for forecasting reference evapotranspiration.
473 *Computers and Electronics in Agriculture*, *113*, 164-173.

474 Heddam, S., Watts, M.J., Houichi, L., Djemili, L., Sebbar, A. (2018). Evolving connectionist
475 systems (ECoSs): a new approach for modeling daily reference evapotranspiration (ET 0).
476 *Environmental Monitoring and Assessment*, 190(9), 516. https://doi.org/10.1007/s10661-
477 018-6903-0.
478
479 Hill, W.J., Hunter, W.G. 1966. "A Review of Response Surface Methodology: A Literature
480 Survey," *Technometrics*, vol. 8, no. 4, pp. 571-590, 1966.

481 Jain, S. K., Nayak, P. C., & Sudheer, K. P. (2008). Models for estimating evapotranspiration
482 using artificial neural networks, and their physical interpretation. *Hydrological Processes*,
483 22(13), 2225-2234.
484
485 Karbasi M. (2018). Forecasting of Multi-Step Ahead Reference Evapotranspiration Using
486 Wavelet-Gaussian Process Regression Model. *Water Resour Manage* 32(3):1035-1052.
487
488 Keshtegar, B., Heddam, S. 2017. "Modeling daily dissolved oxygen concentration using
489 modified response surface method and artificial neural network: a comparative study,"
490 *Neural Comput. Appl.*, pp. 1-12. https://doi.org/10.1007/s00521-017-2917-8

491 Keshtegar, B., Kisi, O. 2017. "Modified Response-Surface Method: New Approach for
492 Modeling Pan Evaporation," *J. Hydrol. Eng.*, vol. 22, no. 10, pp. 1–14.

493 Keshtegar, B., Kisi, O., & Zounemat-Kermani, M. (2019). Polynomial chaos expansion and
494 response surface method for nonlinear modelling of reference evapotranspiration.
495 *Hydrological Sciences Journal*, 1-11.

496 Keshtegar, B., Kisi, O., Arab, H. G., & Zounemat-Kermani, M. (2018). Subset Modeling
497 Basis ANFIS for Prediction of the Reference Evapotranspiration. *Water Resources*
498 *Management*, 1-16.

499 Keshtegar, B., Bagheri, M., Fei, C.-W., Lu, C., Taylan, O., & Thai, D.-K. (2021). Multi-
500 extremum-modified response basis model for nonlinear response prediction of dynamic
501 turbine blisk. *Engineering with Computers*. doi: 10.1007/s00366-020-01273-8

502 Khoshravesh, M., Sefidkouhi, M. A. G., & Valipour, M. (2017). Estimation of reference
503 evapotranspiration using multivariate fractional polynomial, Bayesian regression, and

504 robust regression models in three arid environments. *Applied Water Science*, *7*(4), 1911-
505 1922.

506 Kisi, O. (2006). Generalized regression neural networks for evapotranspiration modelling.
507 *Hydrological Sciences Journal*, 51(6), 1092-1105.

508 Kisi, O. (2016). Modeling reference evapotranspiration using three different heuristic
509 regression approaches. *Agricultural Water Management*, *169*, 162-172.

510 Kişi, Ö., & Öztürk, Ö. (2007). Adaptive Neurofuzzy computing technique for
511 evapotranspiration estimation. *Journal of Irrigation and Drainage Engineering*, *133*(4),
512 368-379.

513 Feng, Y., Peng, Y., Cui, N., Gong, D., Zhang, K. (2017). Modeling reference
514 evapotranspiration using extreme learning machine and generalized regression neural
515 network only with temperature data. *Computers and Electronics in Agriculture*, 136, 71-
516 78.
517
518 Feng, Y., Peng, Y., Cui, N., Gong, D., & Zhang, K. (2017). Modeling reference
519 evapotranspiration using extreme learning machine and generalized regression neural
520 network only with temperature data. *Computers and Electronics in Agriculture*, 136, 71-
521 78.
522
523 Fang W., Huang S., Huang Q., Huang G., Meng E., Luan J. (2018). Reference
524 evapotranspiration forecasting based on local meteorological and global climate
525 information screened by partial mutual information. *Journal of Hydrology*, 561, 764-779.
526
527 Mehdizadeh, S., Behmanesh, J., Khalili, K. (2017). Using MARS, SVM, GEP and empirical
528 equations for estimation of monthly mean reference evapotranspiration. *Computers and*
529 *Electronics in Agriculture*, 139, 103-114.
530
531 Ozonoh, M., Oboirien, B.O., Higginson, A., Daramola, M.O. (2020). Performance evaluation of
532 gasification system efficiency using artificial neural network. Renewable Energy, 145, 2253-2270.
533 https://doi.org/10.1016/j.renene.2019.07.136.
534
535 Rahimikhoob, A. (2016). Comparison of M5 model tree and artificial neural network's
536 methodologies in modelling daily reference evapotranspiration from NOAA satellite
537 images. *Water Resources Management*, *30*(9), 3063-3075.

538 Saggi, M. K., Jain, S. (2019). Reference evapotranspiration estimation and modeling of the
539 Punjab Northern India using deep learning. *Computers and Electronics in Agriculture*,
540 156, 387-398.

541 Sanikhani, H., Kisi, O., Maroufpoor, E., & Yaseen, Z. M. (2018). Temperature-based
542 modeling of reference evapotranspiration using several artificial intelligence models:
543 application of different modeling scenarios. *Theoretical and Applied Climatology*, 1-14.

544 Sanikhani, H., Kisi, O., Maroufpoor, E., & Yaseen, Z. M. (2019). Temperature-based
545 modeling of reference evapotranspiration using several artificial intelligence models:
546 application of different modeling scenarios. *Theoretical and Applied Climatology*, 135(1-
547 2), 449-462.

548 Shiri, J., Sadraddini, A. A., Nazemi, A. H., Kisi, O., Landeras, G., Fard, A. F., & Marti, P.
549 (2014). Generalizability of gene expression programming-based approaches for estimating
550 daily reference evapotranspiration in coastal stations of Iran. *Journal of Hydrology*, *508*,
551 1-11.

552 Shiri, J., Marti, P., Nazemi, A. H., Sadraddini, A. A., Kisi, O., Landeras, G., & Fakheri Fard,
553 A. (2015). Local vs. external training of neuro-fuzzy and neural networks models for
554 estimating reference evapotranspiration assessed through k-fold testing. *Hydrology*
555 *Research*, 46(1), 72-88.

556 Solomatine, D. P., & Xue, Y. (2004). M5 model trees and neural networks: application to
557 flood forecasting in the upper reach of the Huai River in China. *Journal of Hydrologic*
558 *Engineering*, *9*(6), 491-501.

559 Shahabinejad, H., Vosoughi, N., Saheli, F. (2020). Matrix effects corrections in prompt gamma-ray
560 spectra of a PGNAA online analyzer system using artificial neural network. *Progress in Nuclear*
561 *Energy*, 118, 103146. https://doi.org/10.1016/j.pnucene.2019.103146.
562
563 Tenenbaum, R. A., Taminato, F. O., Melo, V. S. (2020). Fast auralization using radial basis
564 functions type of artificial neural network techniques. *Applied Acoustics*, 157, 106993.
565 https://doi.org/10.1016/j.apacoust.2019.07.041.
566
567 Tao H., Diop L., Bodian A., Djaman K., Ndiaye P.M., Yaseen Z.M. (2018). Reference
568 evapotranspiration prediction using hybridized fuzzy model with firefly algorithm:
569 Regional case study in Burkina Faso. *Agricultural Water Management*, 208, 140-151.
570 https://doi.org/10.1016/j.agwat.2018.06.018.
571
572 Wang, L., Kisi, O., Hu, B., Bilal, M., Zounemat-Kermani, M., & Li, H. (2017). Evaporation
573 modelling using different machine learning techniques. *International Journal of*
574 *Climatology*, *37*(S1), 1076-1092.
575
576 Wang, Y., Liu, H., Yu, Z., Tu, L. (2020). An improved artificial neural network based on
577 human-behaviour particle swarm optimization and cellular automata. *Expert Systems with*
578 *Applications*, 140, 112862. https://doi.org/10.1016/j.eswa.2019.112862.
579
580 Xiao M, Zhang J, Gao L, Lee S, Eshghi AT (2019). An efficient Kriging-based subset
581 simulation method for hybrid reliability analysis under random and interval variables with
582 small failure probability. *Structural and Multidisciplinary Optimization*, 59:2077-2092
583 Xiao M, Zhang J, Gao L (2020). A system active learning Kriging method for system
584 reliability-based design optimization with a multiple response model. *Reliability*
585 *Engineering & System Safety*, 199:106935
586
587 Yin, Z., Wen, X., Feng, Q., He, Z., Zou, S., & Yang, L. (2017). Integrating genetic algorithm
588 and support vector machine for modeling daily reference evapotranspiration in a semi-arid
589 mountain area. *Hydrology Research*, 48(5), 1177-1191.
590
591 Zhang J, Xiao M, Gao L, Chu S (2019). A combined projection-outline-based active learning
592 Kriging and adaptive importance sampling method for hybrid reliability analysis with
593 small failure probabilities. *Computer Methods in Applied Mechanics and Engineering*,
594 344:13-33

595 Zhang J, Gao L, Xiao M (2020a) A new hybrid reliability-based design optimization method
596 under random and interval uncertainties. *International Journal for Numerical Methods in*
597 *Engineering*, 121:4435-4457
598
599 Zhang Y, Gao L, Xiao M (2020b). Maximizing natural frequencies of inhomogeneous cellular
600 structures by Kriging-assisted multiscale topology optimization. *Computers and Structures*,
601 230:106197
602
603 Zounemat-Kermani, M., Kisi, O., Piri, J., & Mahdavi-Meymand, A. (2019). Assessment of
604 Artificial Intelligence–Based Models and Metaheuristic Algorithms in Modeling Evaporation.
605 *Journal of Hydrologic Engineering*, *24*(10), 04019033.
606
607 Zhou, Q., Ren, P., Tan, Y.L. (2012). Soft measurement of paper smoothness based on time-
608 frequency analysis of paper quantization noise. *Measurement*, 45(3), 493-499.
609 doi:10.1016/j.measurement.2011.10.023.
610
611 Zhu S-P, Keshtegar B, Tian K, Trung N-T (2021) Optimization of Load-Carrying
612 Hierarchical Stiffened Shells: Comparative Survey and Applications of Six Hybrid Heuristic
613 Models. *Archives of Computational Methods in Engineering* doi:10.1007/s11831-021-09528-
614 3
615
616 Zounemat-Kermani, M., Matta, E., Cominola, A., Xia, X., Zhang, Q., Liang, Q., &
617 Hinkelmann, R. (2020). Neurocomputing in Surface Water Hydrology and Hydraulics: A
618 Review of Two Decades Retrospective, Current Status and Future Prospects. *Journal of*
619 *Hydrology*, 125085.
620
621 Adnan, R. M., Heddam, S., Yaseen, Z. M., Shahid, S., Kisi, O., & Li, B. (2021). Prediction of
622 Potential Evapotranspiration Using Temperature-Based Heuristic Approaches. *Sustainability*,
623 *13*(1), 297.
624
625 Chen, Z., Zhu, Z., Jiang, H., & Sun, S. (2020). Estimating daily reference evapotranspiration
626 based on limited meteorological data using deep learning and classical machine learning
627 methods. *Journal of Hydrology*, *591*, 125286.
628
629 Zhu, B., Feng, Y., Gong, D., Jiang, S., Zhao, L., & Cui, N. (2020). Hybrid particle swarm
630 optimization with extreme learning machine for daily reference evapotranspiration prediction
631 from limited climatic data. *Computers and Electronics in Agriculture*, *173*, 105430.
632
633 Nagappan, M., Gopalakrishnan, V., & Alagappan, M. (2020). Prediction of reference
634 evapotranspiration for irrigation scheduling using machine learning. *Hydrological Sciences*
635 *Journal*, 1-9.
636
637 Salam, R., & Islam, A. R. M. T. (2020). Potential of RT, Bagging and RS ensemble learning
638 algorithms for reference evapotranspiration prediction using climatic data-limited humid
639 region in Bangladesh. *Journal of Hydrology*, *590*, 125241.
640
641 Ferreira, L. B., & da Cunha, F. F. (2020). Multi-step ahead forecasting of daily reference
642 evapotranspiration using deep learning. *Computers and Electronics in Agriculture*, *178*,
643 105728.
644

Chia, M. Y., Huang, Y. F., & Koo, C. H. (2020). Support vector machine enhanced empirical reference evapotranspiration estimation with limited meteorological parameters. *Computers and Electronics in Agriculture*, *175*, 105577.

Yamaç, S. S., & Todorovic, M. (2020). Estimation of daily potato crop evapotranspiration using three different machine learning algorithms and four scenarios of available meteorological data. *Agricultural Water Management*, *228*, 105875.

655

Figure captions

Fig. 1. The location of the stations in Mediterranean region of Turkey

Fig. 2. Schematic view of RBF (***K***) for C =0 and ε =0.5

Fig. 3. Schematic structure of RM5Tree model

Fig.4. Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of mean temperature (T) in test (35% from all data) period for Isparta, Antalya, and Adana stations

Fig. 5. Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of wind speed (W) in test (35% from all data) period for Isparta, Antalya, and Adana stations

Fig. 6. Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of solar radiation (SR) in test (35% from all data) period for Isparta, Antalya, and Adana stations

Fig. 7. Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of relative humidity (RH) in test (35% from all data) period for Isparta, Antalya, and Adana stations

Fig. 8. Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of mean temperature (T) and solar radiation (SR) in test (35% from all data) period for Isparta, Antalya, and Adana stations

Fig. 9. Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of mean temperature (T), solar radiation (SR) and relative humidity (RH) in test (35% from all data) period for Isparta, Antalya, and Adana stations

Fig. 10. Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of mean temperature (T), solar radiation (SR), relative humidity (RH) and wind speed (W) in test (35% from all data) period for Isparta, Antalya, and Adana stations

656 **Fig. 11.** Radar chart for the best calculated values of RMSE (mm) for the applied models
657 using the three input scenarios.

Figures

Figure 1

The location of the stations in Mediterranean region of Turkey Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

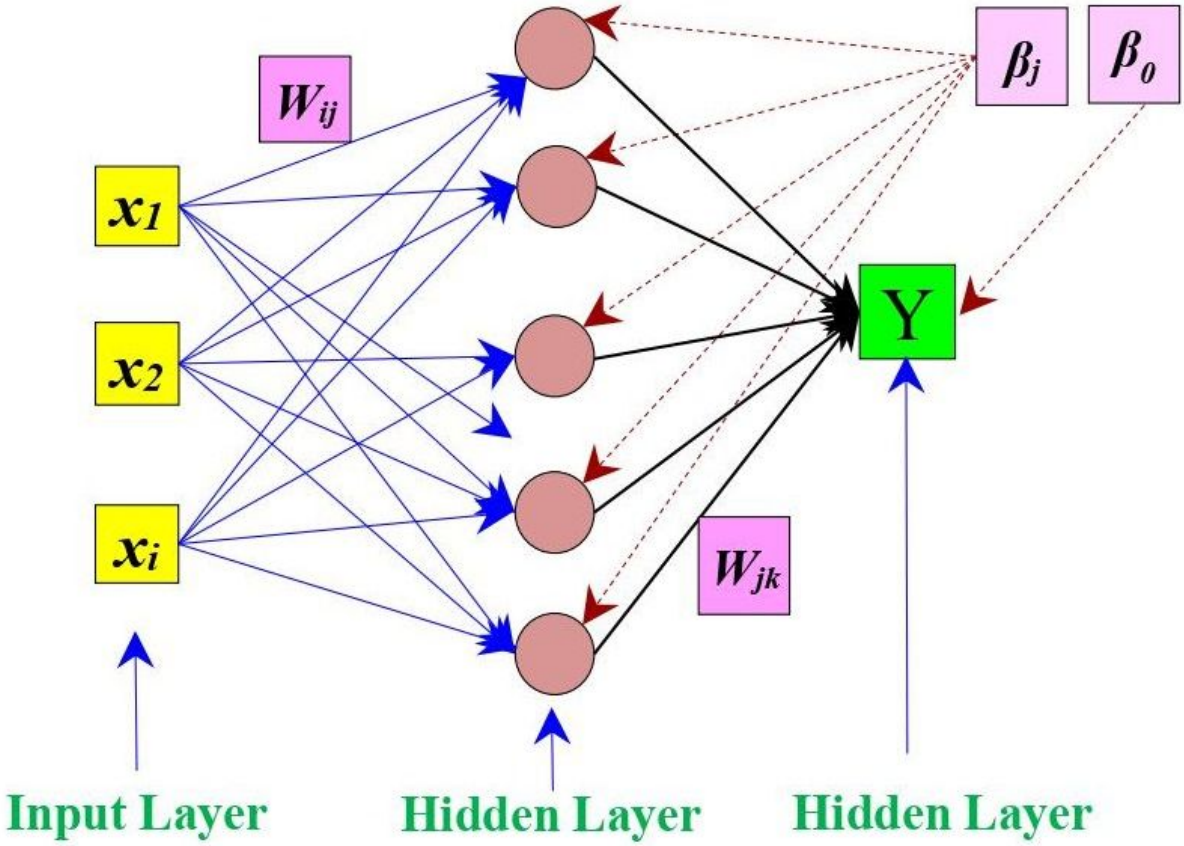

Figure 2

General Structure of the MLPNN model

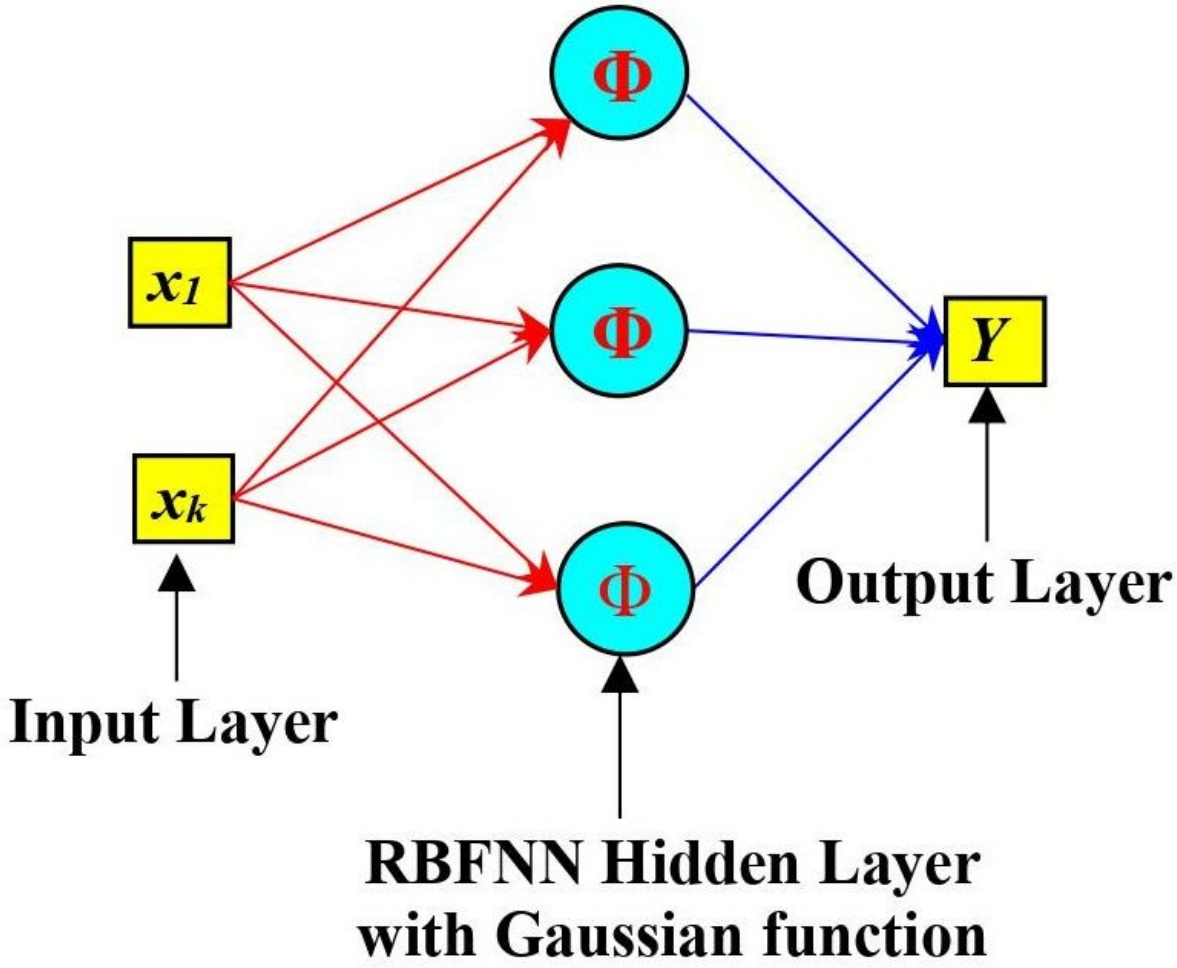

Figure 3

Architecture of the established structure radial basis function neural network (RBFNN)

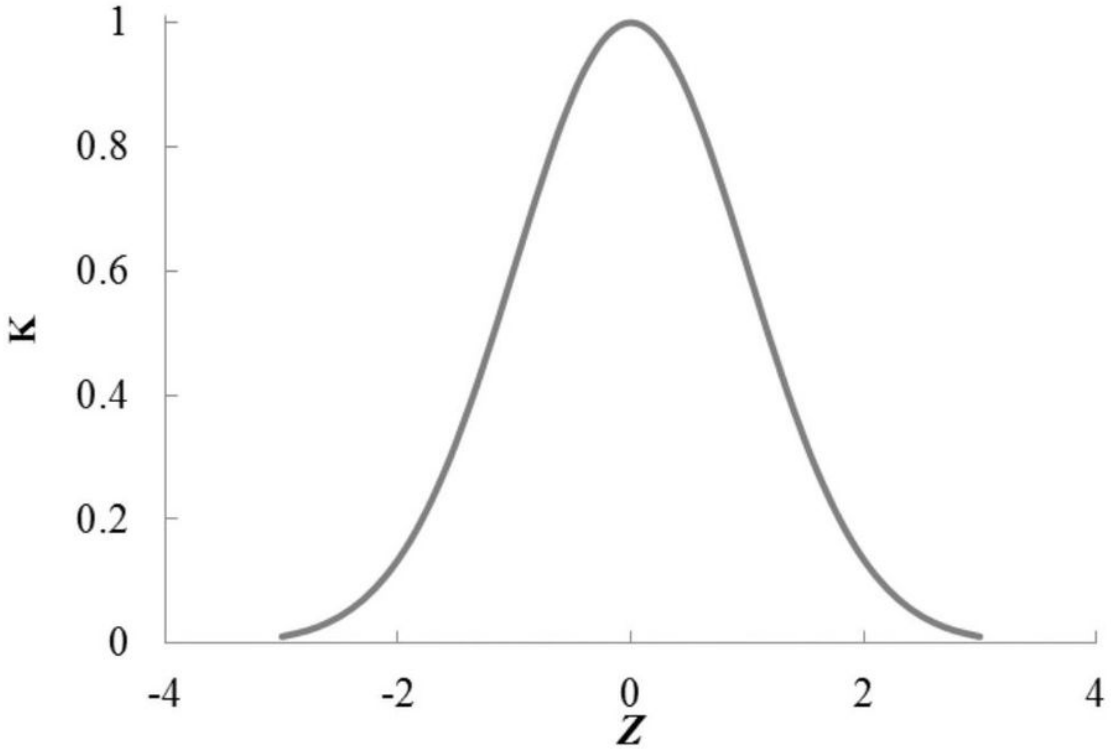

Figure 4

Schematic view of RBF (K) for =0 and =0.5

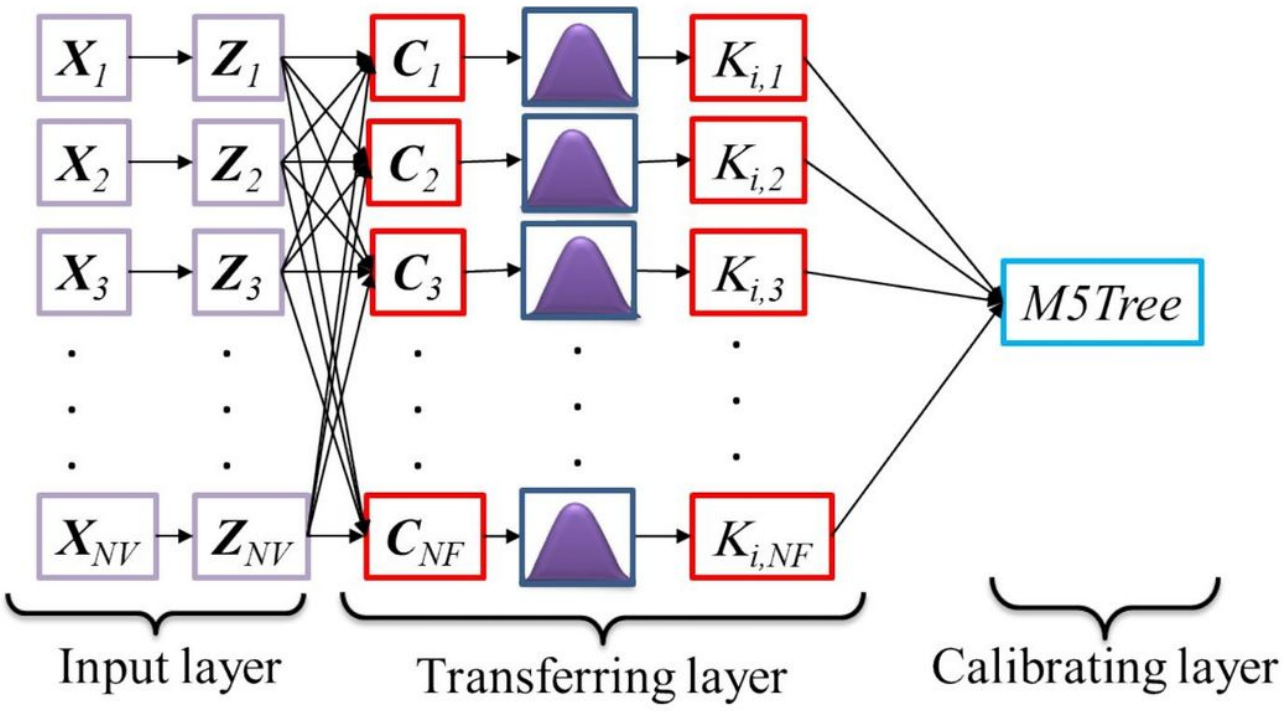

Figure 5

Schematic structure of RM5Tree model

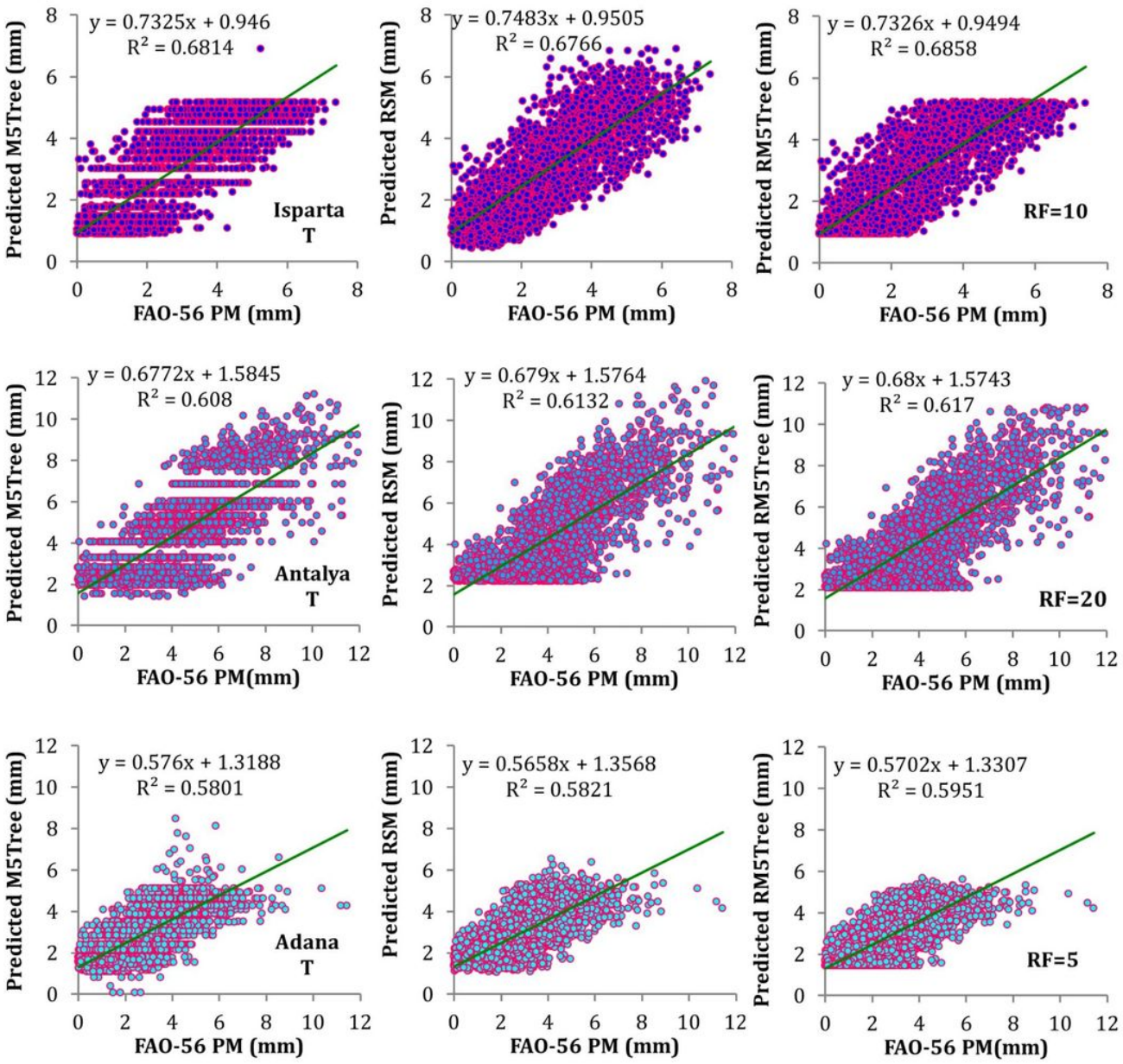

Figure 6

Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of mean temperature (T) in test (35% from all data) period for Isparta, Antalya, and Adana stations (RF: number of radial function)

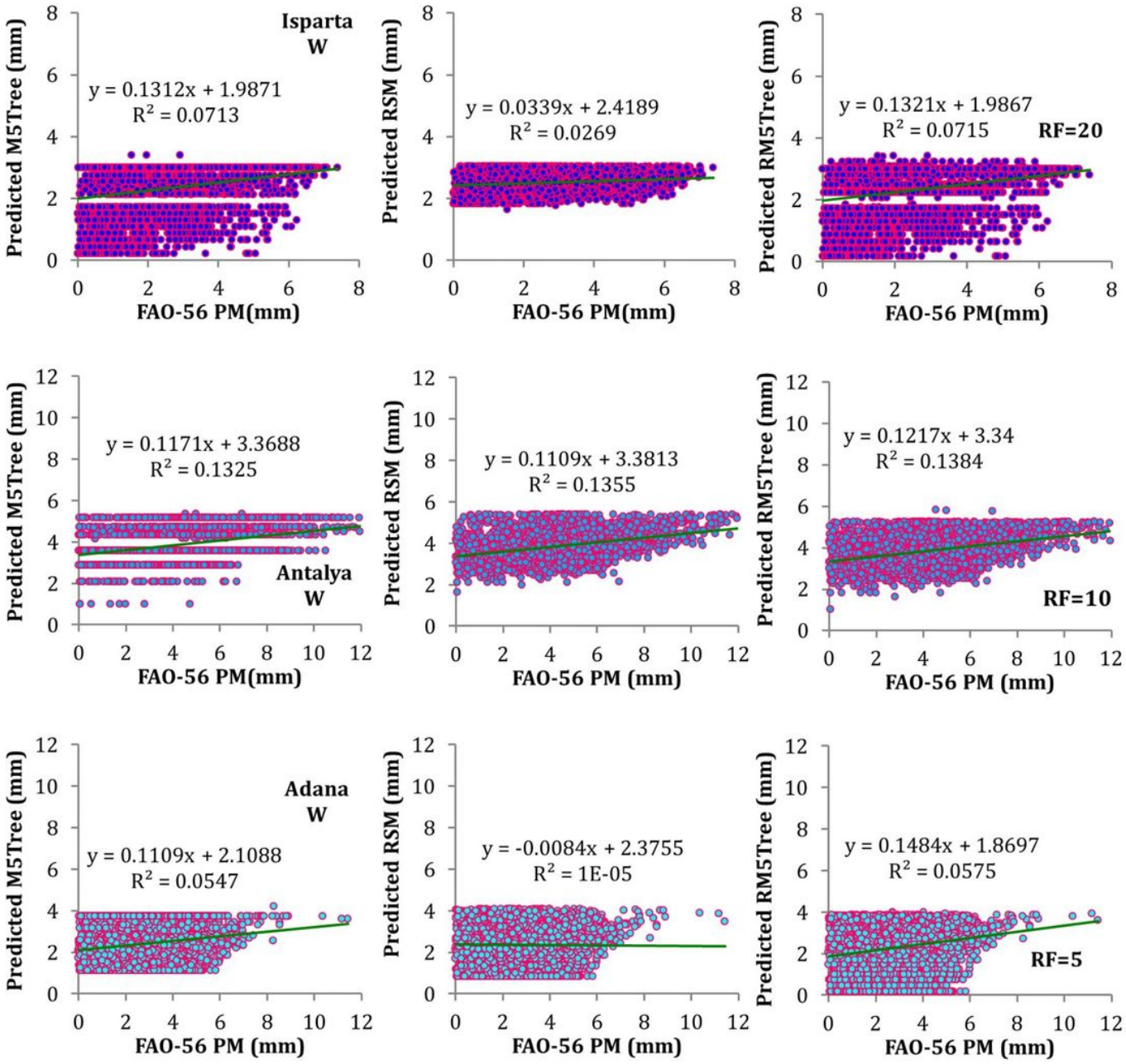

Figure 7

Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of wind speed (W) in test (35% from all data) period for Isparta, Antalya, and Adana stations (RF: number of radial function)

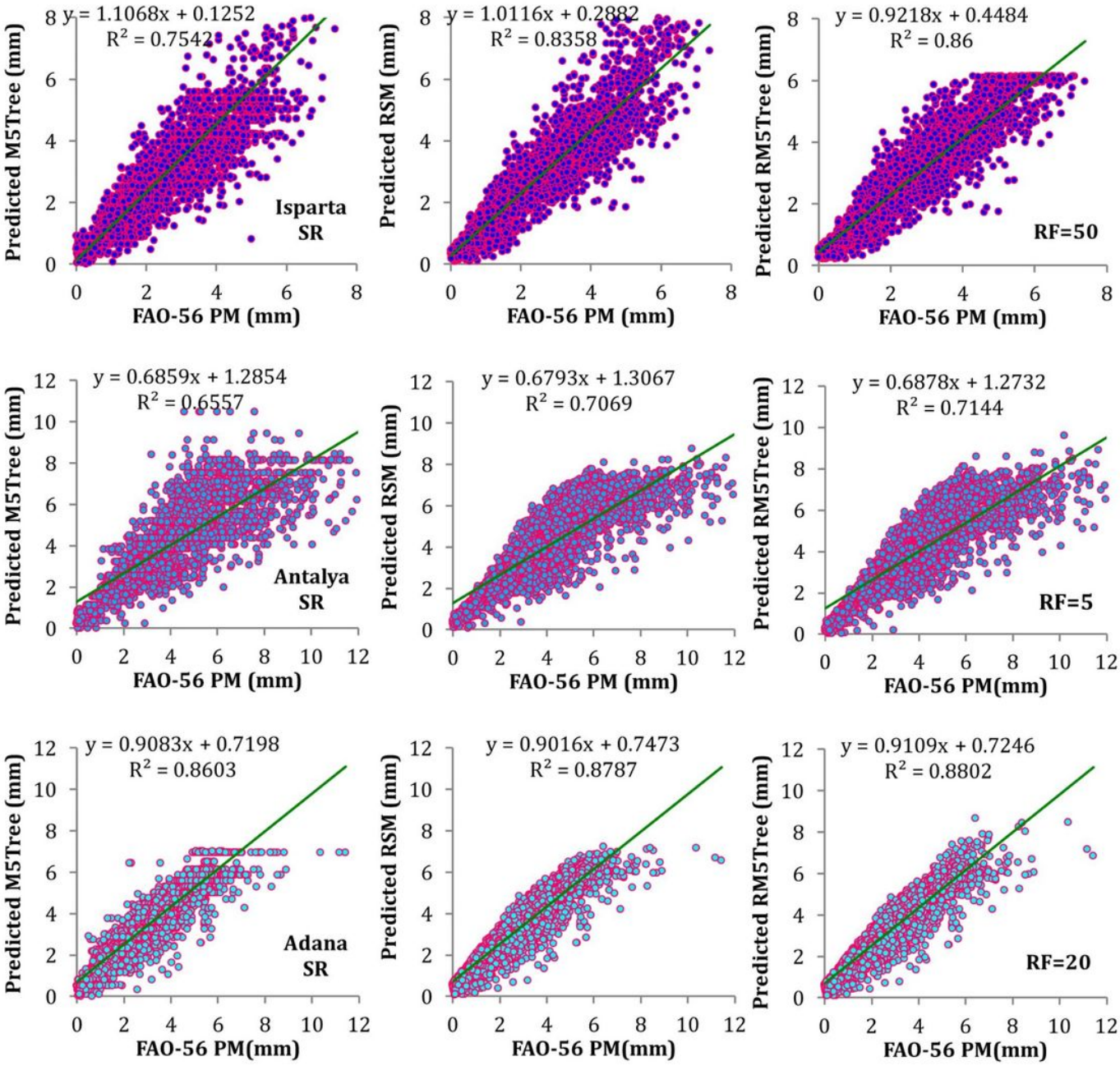

Figure 8

Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of solar radiation (SR) in test (35% from all data) period for Isparta, Antalya, and Adana stations (RF: number of radial function)

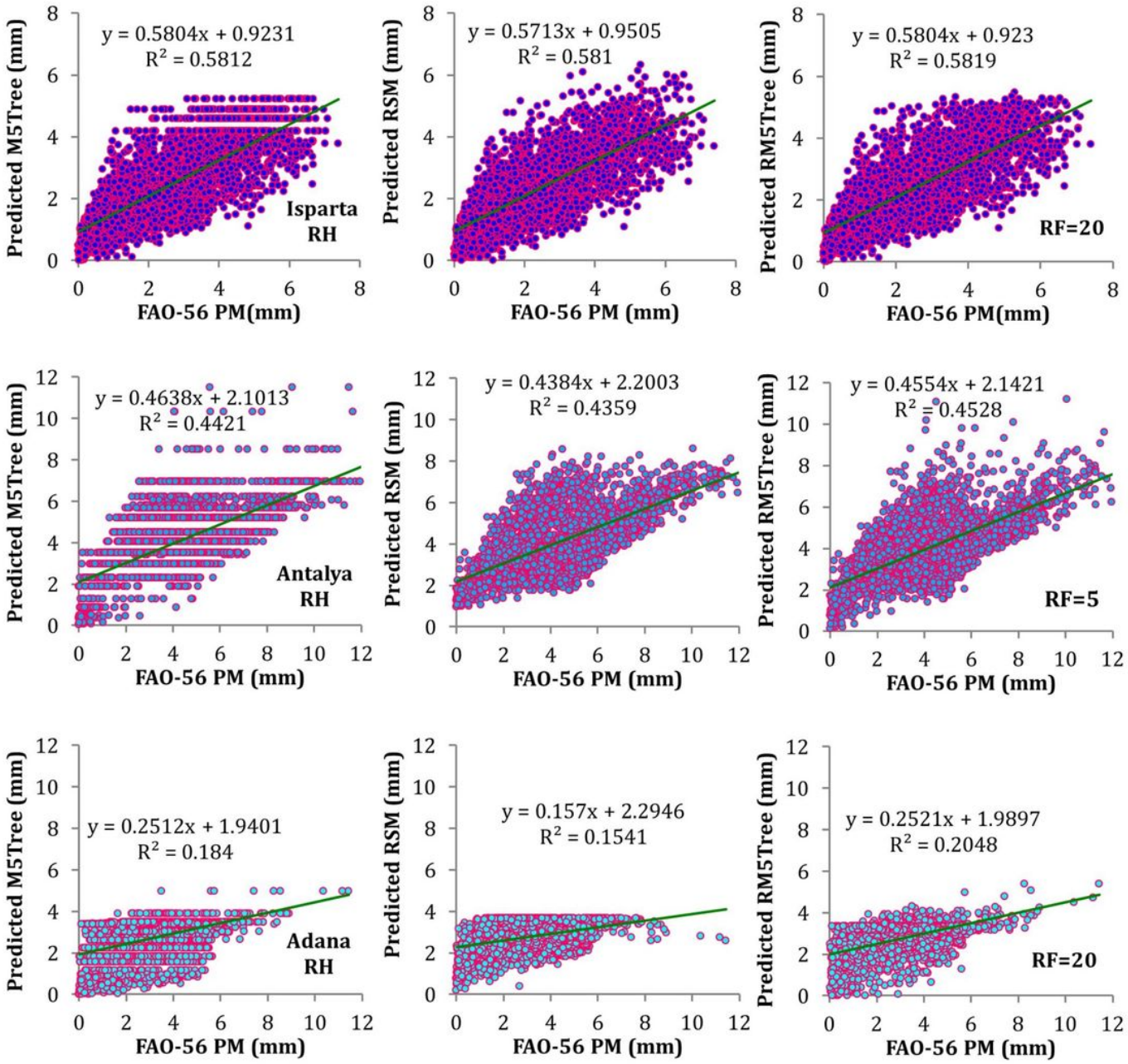

Figure 9

Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of relative humidity (RH) in test (35% from all data) period for Isparta, Antalya, and Adana stations (RF: number of radial function)

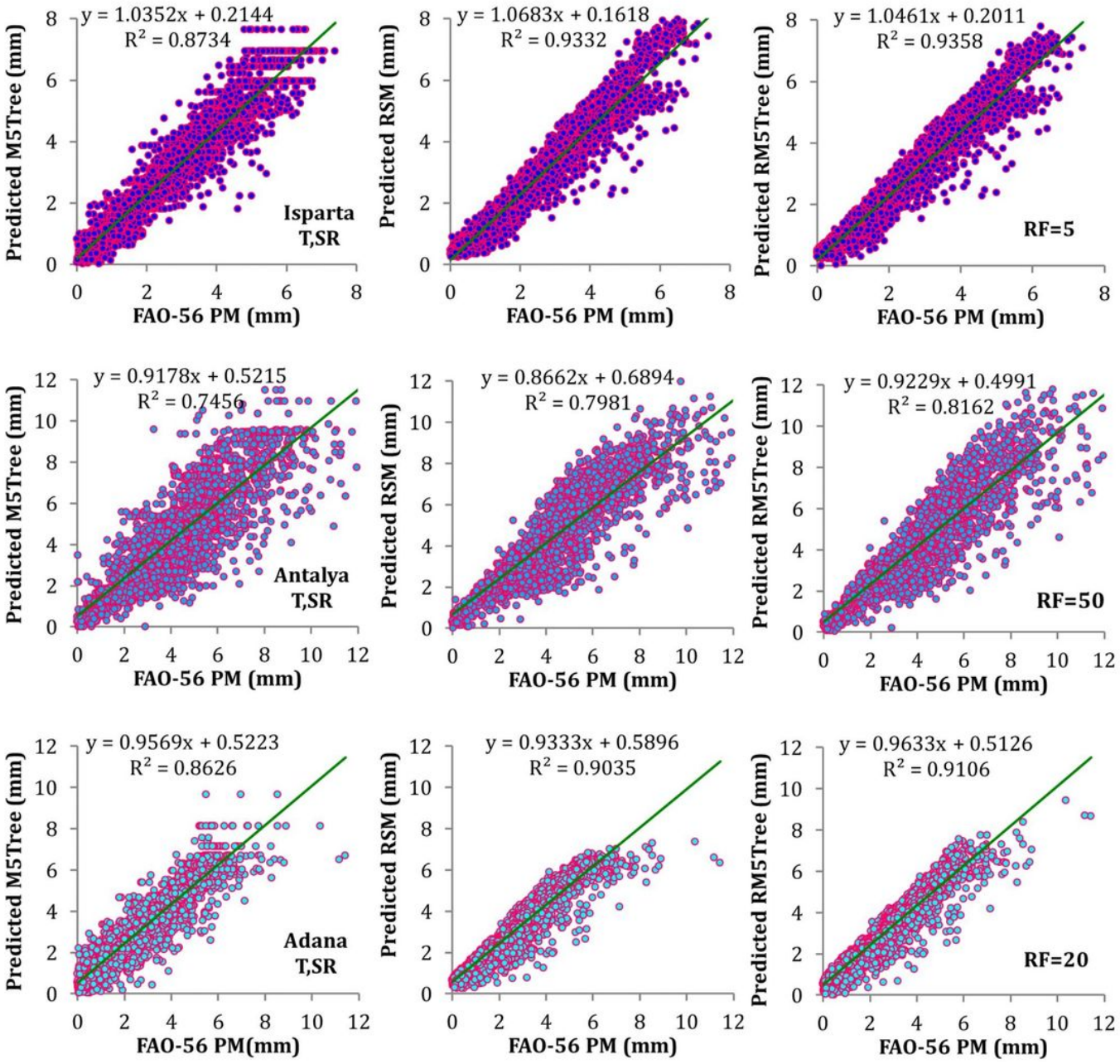

Figure 10

Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of mean temperature (T) and solar radiation (SR) in test (35% from all data) period for Isparta, Antalya, and Adana stations (RF: number of radial function)

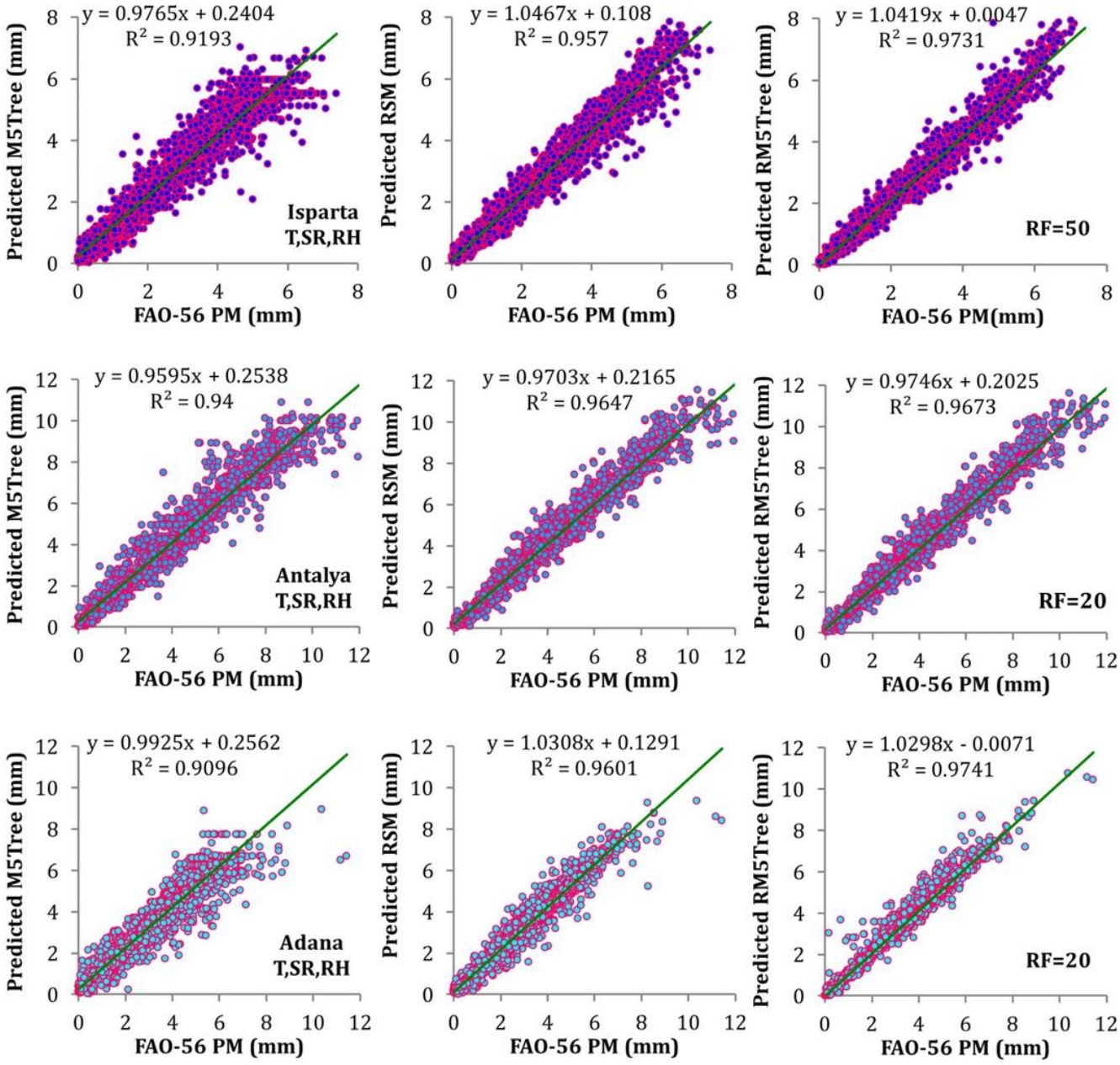

Figure 11

Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of mean temperature (T), solar radiation (SR) and relative humidity (RH) in test (35% from all data) period for Isparta, Antalya, and Adana stations (RF: number of radial function)

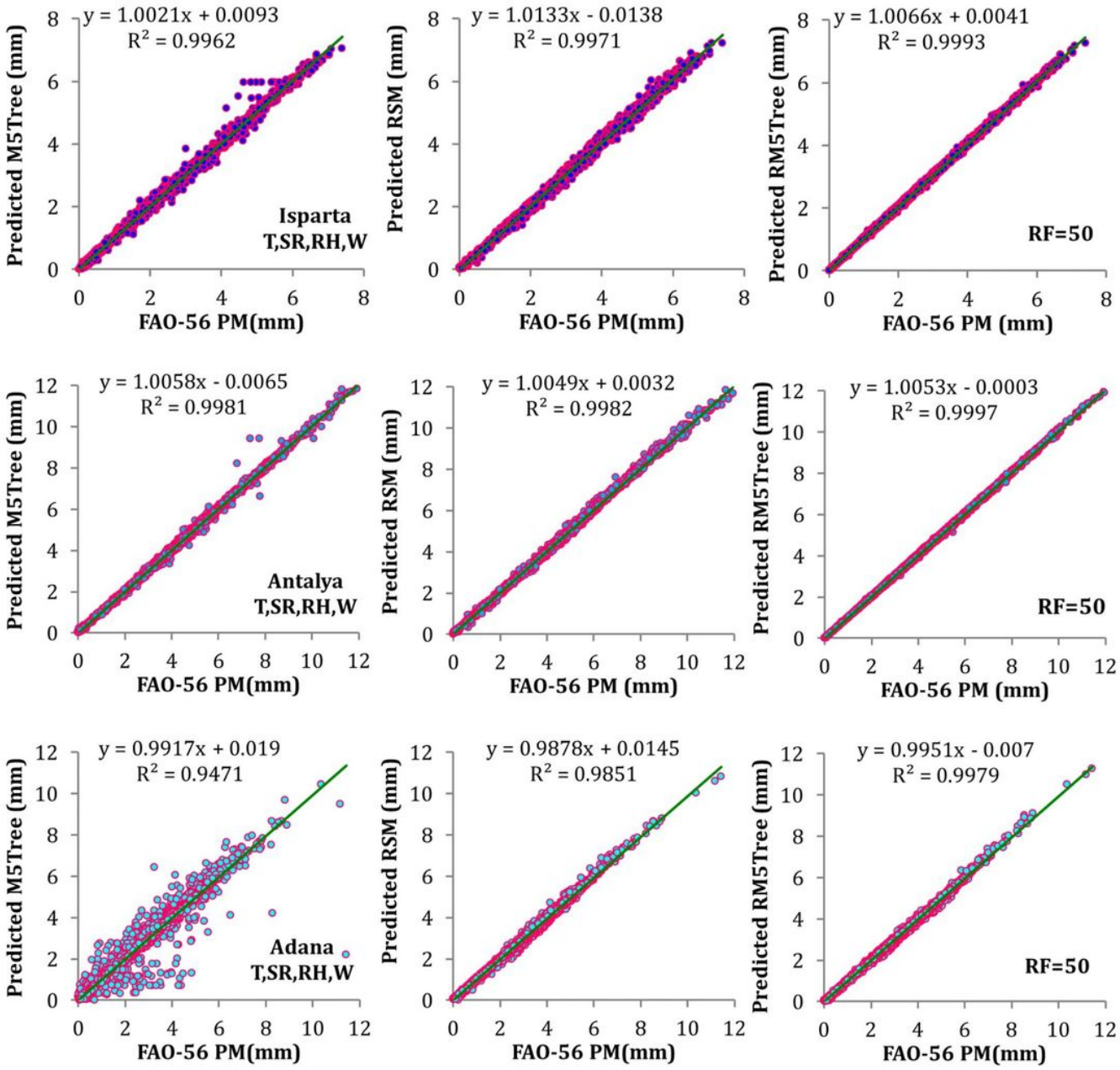

Figure 12

Scatterplot of the M5Tree, RSM, and RM5Tree models based on the input data of mean temperature (T), solar radiation (SR), relative humidity (RH) and wind speed (W) in test (35% from all data) period for Isparta, Antalya, and Adana stations (RF: number of radial function)

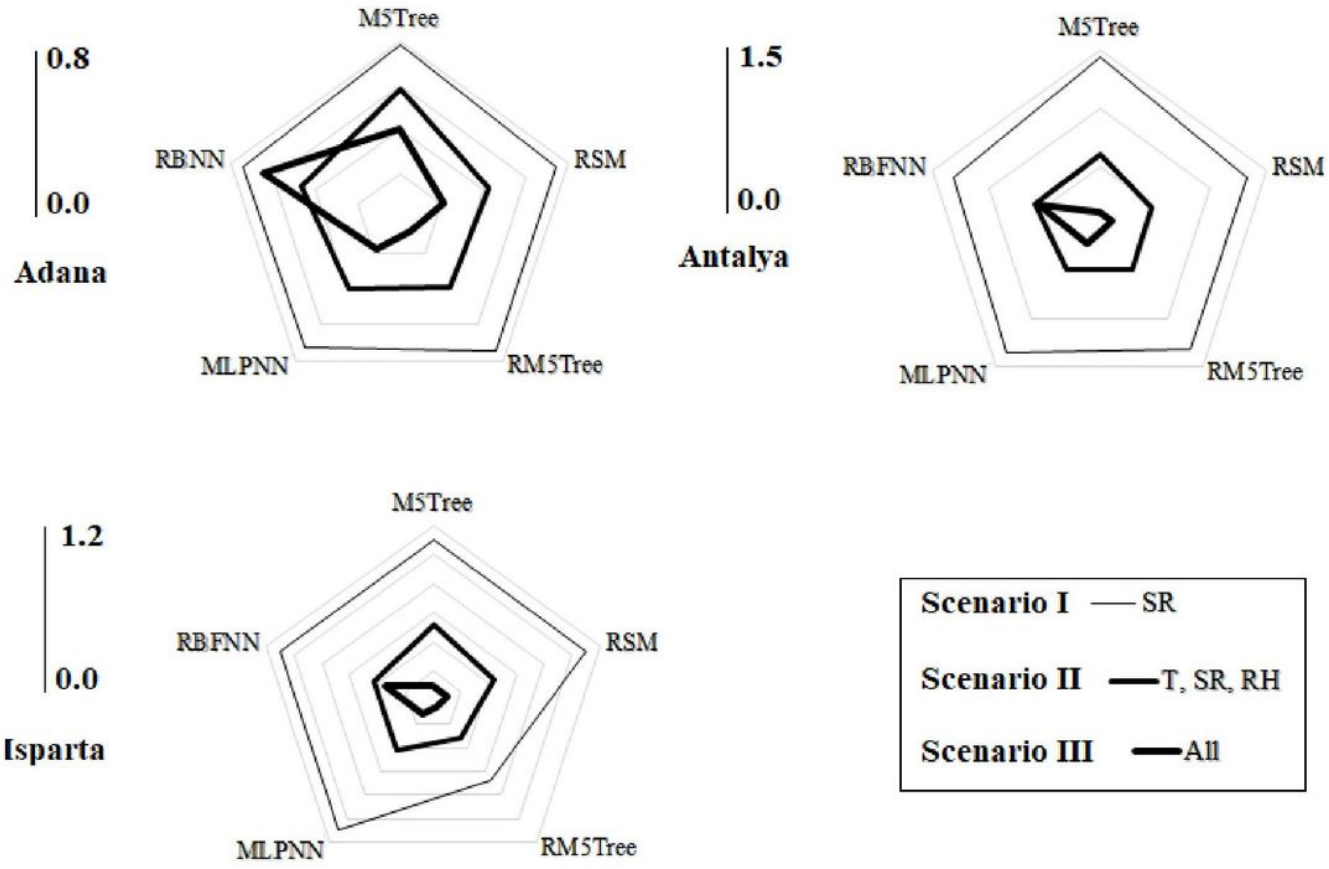

Figure 13

Radar chart for the best calculated values of RMSE (mm) for the applied models using the three input scenarios.

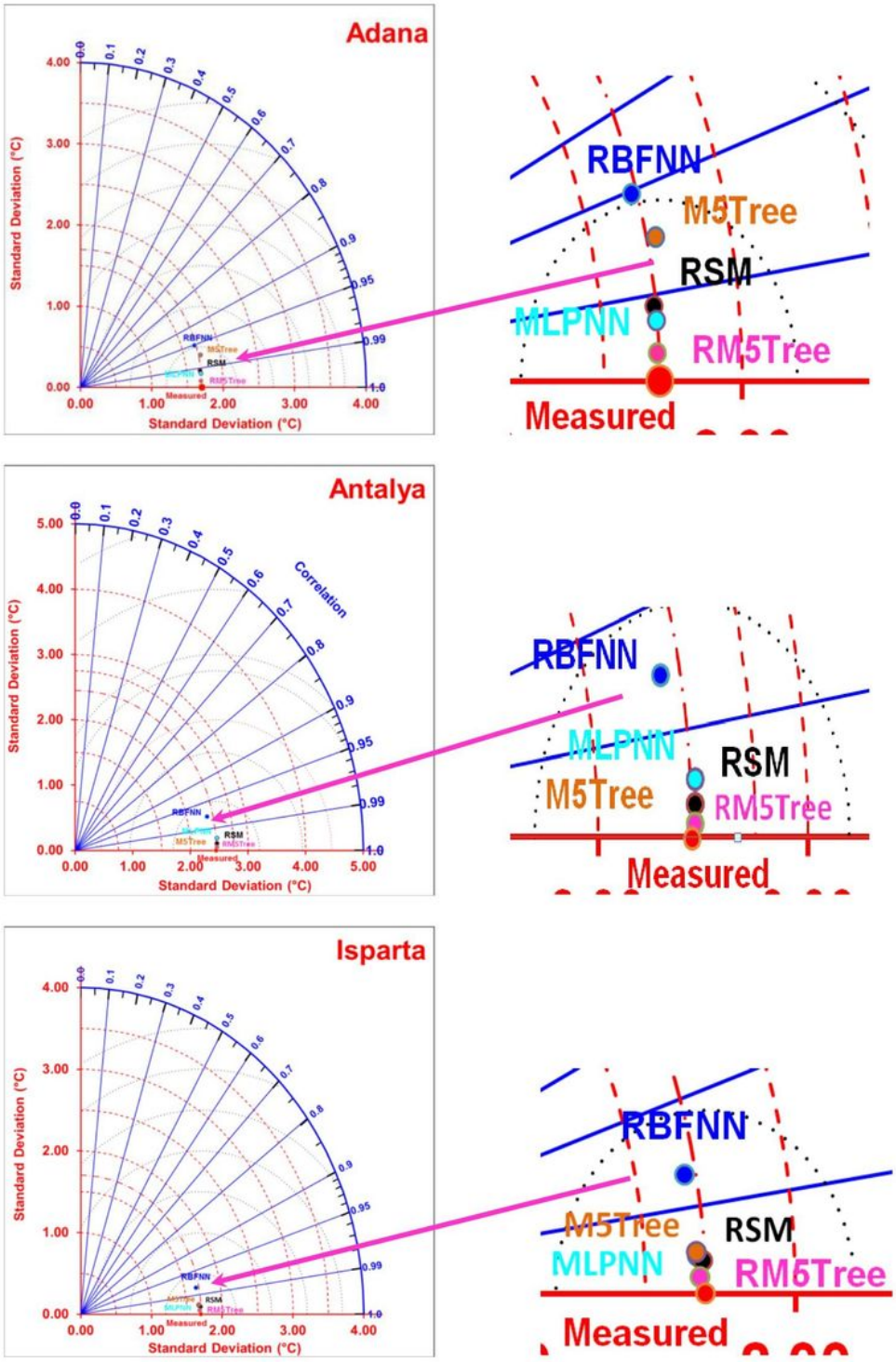

Figure 14

Taylor diagram displaying a statistical comparison of the proposed models with FAO-56 PM (mm). The green circles correspond to circumferences of equal centered normalized root-mean-square (NRMS) difference between measured and calculated ET0, the blue lines correspond to lines of equal correlation coefficients, and doted red circles correspond to circumferences of equal standard deviations.

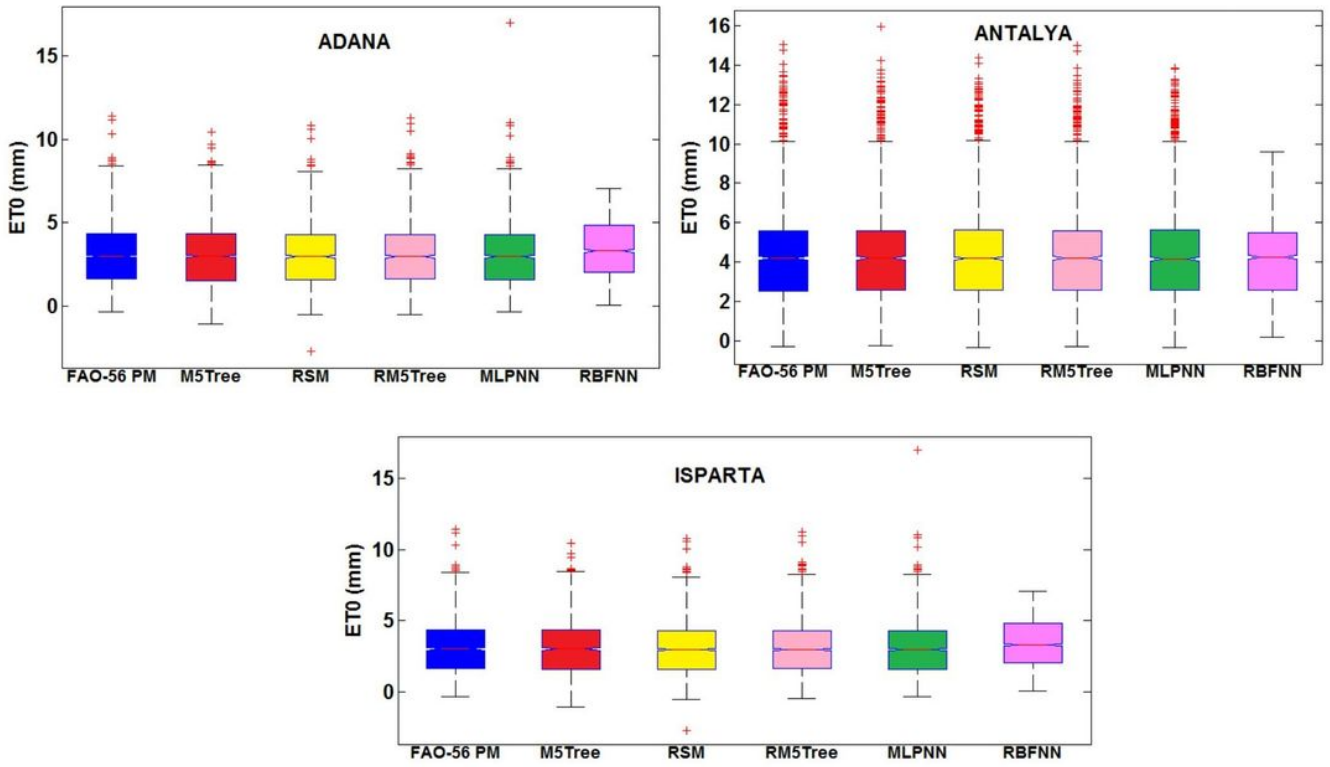

Figure 15

Box plots of FAO-56 PM and calculated values of ET0 in the test phase of all stations. The box stretches from the 25th percentile to the 75th percentile. The median is shown as a red line, and the whiskers correspond to the most extreme data points.