Estimation of Reference Evapotranspiration Using Climatic Data

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
CIVIL AND ENVIRONMENTAL ENGINEERING
April 2017

By
Margaret K. Lum

Thesis Committee:

Sayed Bateni, Chairperson
Roger Babcock
Oceana Francis

ACKNOWLEDGEMENTS

I would like to thank Dr. Bateni, Dr. Babcock, and Dr. Francis for being on my thesis committee and for providing feedback. I would also like to thank Dr. Jalal Shiri and Emily Stack for their contributions.

ABSTRACT

In this study, daily reference evapotranspiration (ET_0) was estimated from climatic data using Multivariate Adaptive Regression Splines (MARS), M5 Model Tree (M5MT), and Gene Expression Programming (GEP). These approaches were trained with climatic data from eight weather stations in Iran for years 2000-2007. Thereafter, they were tested with data from the same eight weather stations in Iran for year 2008 and fourteen weather stations in California for year 2015. Four data combinations were evaluated: daily mean air temperature, daily mean wind speed, daily mean relative humidity, and solar radiation (configuration 1); daily mean air temperature and solar radiation (configuration 2); daily mean air temperature and daily mean relative humidity (configuration 3); daily maximum, minimum, and mean air temperature, and extraterrestrial radiation (configuration 4). In the first part of the study, MARS, M5MT, and GEP models were tested with data from the same Iran stations they were trained with. In the second part of the study, these approaches were tested with data from stations in California. The performance of MARS, M5MT, and GEP models were evaluated using mean absolute error (*MAE*), root mean square error (*RMSE*), and the coefficient of determination (R^2). For all approaches, configuration 1 produced the most accurate results. Configuration 4 was found to be region dependent and is suggested when region specific data is limited (i.e., only temperature data is available). Results indicated MARS, M5MT and GEP could successfully predict ET_0 from climatic data. Comparison of these approaches showed that GEP yielded the most accurate results. Also, it was found that MARS outperformed M5MT.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Evapotranspiration (ET) is an important component of the hydrologic cycle and water resource management. ET involves the loss of water into the atmosphere through evaporation from the soil surface, and transpiration from the stomata of leaves. This process is dependent upon numerous climatologic variables and is responsible for the water stored in the atmosphere (Traore et al., 2010). About 90% of water in the atmosphere is due to evaporation from the land surface (i.e., water bodies and soil), while transpiration of plants is responsible for the remaining 10% (USGS, 2016). On a global scale, roughly 80% of water allotted for crop production and irrigation, is lost through ET (Xu et al., 2016; Lui et al. 2009).

The demand for water has intensified due to global climate change and the significant increase of Earth's population. In the last century, the United Nations reported water use was two times greater than the world's population growth. At this rate, 1.8 billion people will be residing in areas with insufficient water by the year 2025 (National Geographic Society, 2016). In response to droughts and the exhaustion of surface water availability, roughly 10% of the world's food is grown by over pumping groundwater sources. In addition, many water intensive crops are cultivated in arid conditions further exhausting water resources (Postel, 2015). Due to the aforementioned reasons, improving water use efficiency is imperative. An accurate estimation of ET enables proper computation of water budgeting, water planning and water allocation.

Several *in situ* methods have been used to measure ET. Although these methods can provide an accurate estimate of ET over homogeneous areas, limitations and inadequacies exist. For instance, the lysimeter technique utilizes water mass balance to estimate ET. Intricate planning and groundwork are required to obtain field measurements. These measurements only pertain to a small field area and may differ from surrounding areas due to the land surface heterogeneity (Verstraeten et al., 2008). Aside from the arduous installation and maintenance, it is also unsuitable to utilize the lysimeter technique in large areas with mixed vegetation because of the inability to obtain a representative sample (Garcia-Navarro et al., 2004).

Due to the high cost of *in situ* methods and the challenges of determining the relationship between multiple factors affecting evaporative demand and ET, other approaches were developed. The Penman-Monteith (PM) equation is an accepted mathematical expression to estimate reference ET (ET_0) (described in Section 2.4). However, this method requires a large amount of climatic data.

To further improve ET_0 estimation, artificial intelligence (AI) systems were used to estimate ET_0. A number of studies showed that AI approaches (e.g., Gene Expression Programming (GEP), Multivariate Adaptive Regression Splines (MARS), and M5 Model Tree (M5MT)) could estimate ET_0 accurately (Pour Ali Baba et al., 2013; Rahimikhoob, 2008; Shiri and Kisi, 2011; Shiri et al., 2014).

Guven et al. (2008) used GEP to model ET_0 with the following hydrologic inputs: humidity, solar energy, wind speed, and average air temperature. Results indicated the GEP model could perform better than the PM equation. Building upon this study, Shiri et al. (2011) evaluated the feasibility of GEP to estimate ET_0, in which GEP estimates were compared to those obtained from the Adaptive Neuro-Fuzzy Inference System (ANFIS), Hargreaves-Samani, Priestley-Taylor and the PM equation. Both studies found that GEP provided the most accurate estimate of ET_0.

Multivariate Adaptive Regression Splines (MARS) exhibit comparable advantages to GEP. Adamowski et al. (2012) used MARS, Artificial Neural Networks (ANN), and Wavelet Artificial Neural Network (WANN) to estimate mountainous watershed runoff and compared the performance of the models. Results showed MARS and WANN performed similarly in estimating runoff, and both produced better results than ANN. Additionally, Yang et al. (2004) found MARS estimated soil temperature at varying depths more accurately than ANN.

Kisi and Parmar (2016) modeled river water pollution utilizing MARS, M5MT, and Least Square Support Vector Machine (LSSVM). Results indicated MARS and LSSVM obtained similar results and both outperformed M5MT. Using inputs of solar radiation, relative humidity, air temperature and wind speed, Kisi (2015) estimated pan evaporation via MARS, LSSVM, and M5MT. It was concluded MARS performed better than LSSVM and M5MT.

Pal and Deswal (2009) found ET_0 estimates from M5MT to be comparable to those from the PM and Hargreaves-Semani equations. Although ANN surpassed M5MT in Sattari et al. (2013), M5MT delivered simple linear relationship between the inputs and output. In predicting wave height, the M5MT proved to be advantageous over ANN because it generates slightly better estimates and provides rules that can be easily understood by the user (Etemad-Shahidi and Mahjoobi, 2009).

The first objective of this study is to estimate ET_0 from climatic data using MARS, M5MT, and GEP approaches. The second goal is to compare the performance of models created from these three approaches using four different data combinations. The third purpose is to assess which combination has the most significant contribution towards the estimation of ET_0. Lastly, we will investigate whether these models can be applied to another region.

2. METHODOLOGY

2.1 Multivariate Adaptive Regression Splines (MARS)

MARS performs linear regression in steps, and has the ability to model non-linearities (Friedman, 1991; Bansal and Salling, 2013). MARS is efficient in producing models dealing with high dimensional input data sets as well as establishing the connection between dependent and independent variables. The models created are adaptive and are modified through multiple iterations to form a polynomial expression.

The process of constructing the MARS model involves splitting the data into multiple sections. For each section, containing p input variables X_j (j = 1,..., p), linear regression functions (i.e., basis functions) are generated describing the data (Sharda et al., 2006). The basis functions consist of the following reflected pair functions:

$$(X_j - t)_+ = \max\,(0, X_j - t) = \begin{cases} X_j - t, & X_j > t \\ 0, & X_j \le t \end{cases} \tag{1a}$$

$$(t - X_j)_+ = \max\,(0, t - X_j) = \begin{cases} t - X_j, & X_j < t \\ 0, & X_j \ge t \end{cases} \tag{1b}$$

where t is the junction formed between two basis functions (i.e., the knot) at an observed value. MARS then takes the reflected pairs (Equations (1a) and (1b)) and creates a collection of basis functions, C:

$$C = \left\{(X_j - t)_+, (t - X_j)_+\right\};\ t \in \left\{x_{1j}, x_{2j}, x_{3j}, \dots\ x_{kj}\right\};k = 1, 2, \dots, n\ ;\ j = 1,2, \dots, p \tag{2}$$

where x_{kj} is the n^{th} observed value of input variable X_j. The user then specifies the constraints, which involve the number of variables inputted into the model, the maximum number of basis functions, and the maximum permissible degree of interaction (described Section 4.1). Once specified, MARS implements two stepwise evaluations (Kisi, 2015). The first process (i.e., forward step) involves a robust model containing a large amount of knots where the MARS algorithm is trialing functions in C to determine which basis functions reduce inaccuracy. Immediately following is the backward step, in which the unnecessary surplus of basis functions is eliminated to correct the over-fitted model. Through this piecewise method, the function created by MARS (f) is generally in the form of:

$$f(x) = \beta_0 + \sum_{i=1}^{A} \beta_i h_i(x) \tag{3}$$

where $h_i(x)$ is the function created from the selected basis functions in C, β_0 is the bias, β_i is the coefficient estimated from the least square method, and A is the number of terms in the final model after implementation of forward and backward steps (Adamowski et al., 2012; Kisi, 2015).

2.2 M5 Model Tree (M5MT)

M5MT is an extension of a regression tree and provides the user with multiple linear functions (Witten and Frank, 2005; Pal and Deswal, 2009). Its main objective is to generate a model that is capable of determining the relationship between the independent variables and target values (i.e., dependent variables) (Bhattacharya and Solomatine, 2005). The performance of the final model is examined by comparing its estimates with the target values.

The input data is subjected to two different stages to generate a final model. In the first stage, the input space is divided into a number of regions (Figure 1a), which corresponds to the nodes in a tree-like structure. The amount of error for each of these nodes corresponds to the standard deviation of each region. For each value that propagates to a specific node, the expected error reduction is calculated, which is referred to as the standard deviation reduction (SDR) (Pal and Deswal, 2009). The formula utilized to calculate the SDR is as follows:

$$SDR = sd(T) - \sum_{i=1}^{A} \frac{|T_i|}{|T|} sd(T_i) \tag{4}$$

where T is the set of instances that reach a node, T_i is the subset of instances exhibiting the i^{th} outcome of a potential set, A is the final amount of instances in set T, and sd is the standard deviation. After the SDRs are calculated, the model begins to execute computational iterations. This involves the dividing process (Figure 1b), which produces subsequent nodes with a reduced standard deviation from the previous nodes. This dividing process is continued where all possible splits are considered, and terminates when the maximum expected error reduction is achieved (Rahimikhoob et al., 2013). At the end of the first stage, the model tree exhibits a large structure with region-specific linear regression models (Bhattacharya and Solomatine, 2005). This stage is prone to over-fitting (or poor generalization) and leads to the second stage of pruning the overgrown model tree.

Pruning will occur when the estimated error at a specific node is less than or equal to the error of the all nodes branched below (Atiaa and Ghalib, 2008; Pal and Deswal, 2009). The pruned nodes are then replaced with linear equations (i.e., linear models (LM)) (Figure 1c). This process simplifies and increases the overall accuracy of the model tree (Wang and Witten, 1997; Etemad-Shahidi and Mahjoobi, 2009). See Figure 1d for an example of a final version of a model tree with an input space divided into six regions corresponding to six linear models.

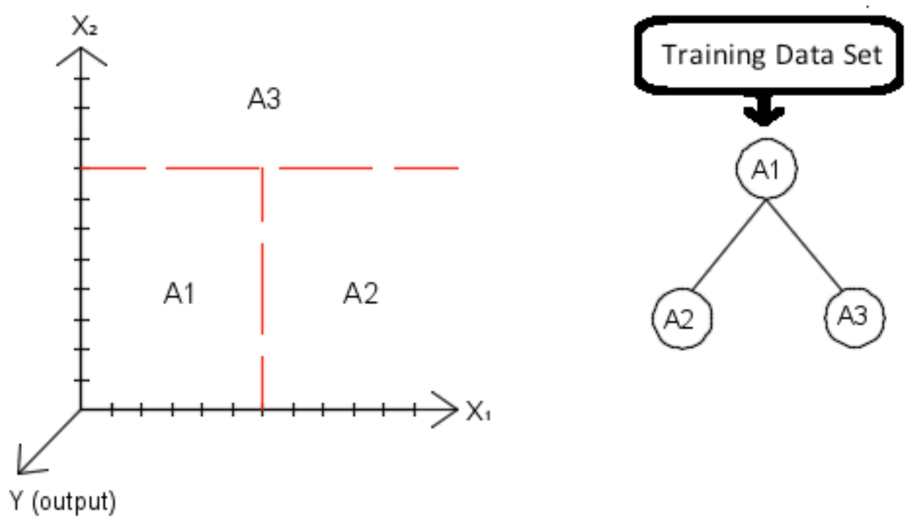


Figure 1a. M5MT separates training data set into regions.

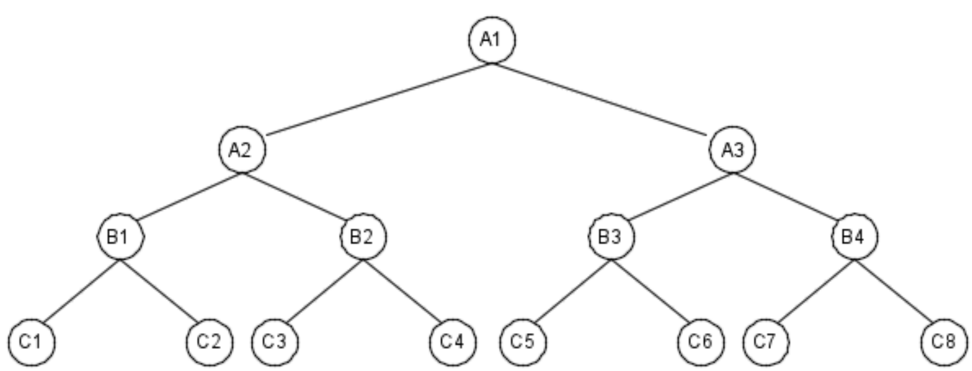


Figure 1b. Dividing process of M5MT.

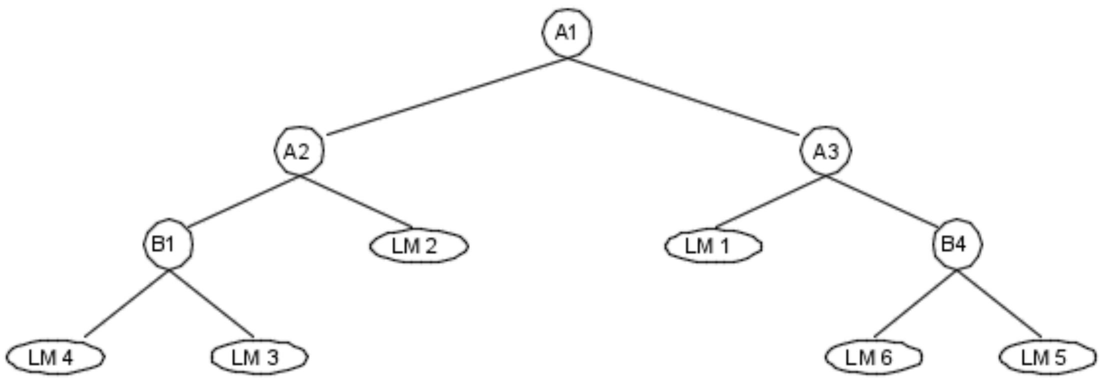


Figure 1c. M5MT pruned nodes are replaced with linear models.

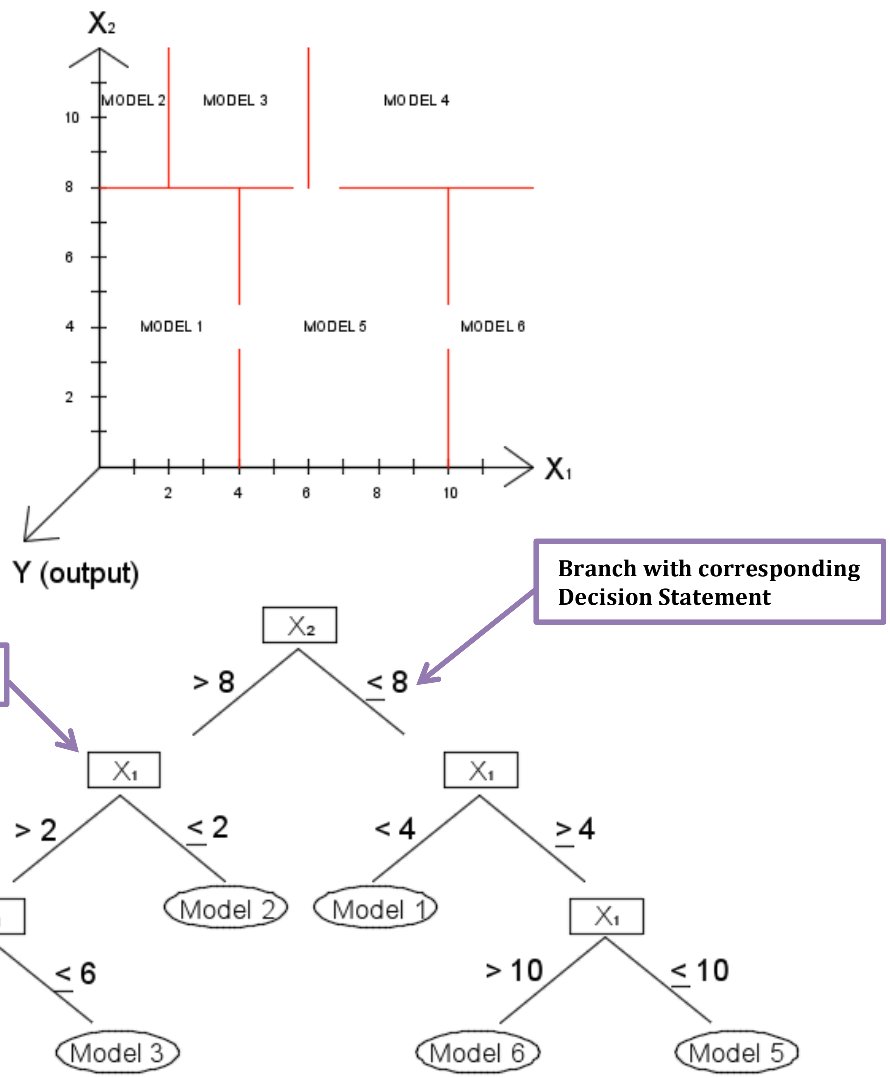


Figure 1d. Example of final M5MT and corresponding input space.

2.3 Gene Expression Programming (GEP)

GEP is a progressive algorithm that continuously adapts to determine the relationship between a given set of inputs and output(s) (Ferreira, 2001). GEP utilizes evolving computer programs (expression trees) of different sizes that are encoded with a finite linear string of input data (chromosomes). A final expression tree is produced consisting of mathematical expressions and polynomials (Guven and Gunal, 2008).

GEP begins with a random generation of chromosomes for a specific program. These chromosomes are made up of genes (finite linear string containing input data) that are linked by arithmetic operators (e.g., +, -, ×, /) (Oltean and Groson, 2003). GEP genes are made up of two parts: the head and the tail. The head consists of both terminal symbols (independent variables) and functional symbols (e.g., Arctan). The tail only consists of terminal symbols and is calculated as follows:

$$t = (n - 1)h + 1 \tag{5}$$

where t is the tail length, n is the amount of arguments in the functions, and h is the head length chosen by the user.

After the chromosomes are generated and used in the program, the accuracy of the program is evaluated using a fitness function (e.g., Root Mean Square Error (*RMSE*)) and fitness cases (i.e., the training data). The fitness function compares the dependent variable (e.g., ET_0) from the given training dataset to the values created by the program. The program will either be selected or rejected based upon its ability to simulate the dependent variable of the dataset. If the program is selected, then it is replicated and further improved through chromosome modifications.

The chromosomes are modified through replication, mutation, transposition, and crossover (i.e., recombination) (Ferreira, 2001). Replication copies an existing chromosome and uses it in the next generation. Mutation of a chromosome is when symbols in the head of a gene changes into other functional or terminal symbols, and symbols in the tail changes into other terminal symbols (Figure 2a).

Transposition is the random selection of segments within a chromosome and moves it to another position. Three types of transposition can occur: insertion sequence (IS), root insertion sequence (RIS), and gene transposition. In IS, random segments with functional or terminal symbols move and replace existing segments (Figure 2b). In RIS, random segments starting with a functional symbol move to the beginning of the gene. The whole head then shifts downstream (i.e., to the right), while losing the last symbols of the head (Figure 2c). Gene transposition moves a randomly selected gene to the beginning of the chromosome (Figure 2d).

Recombination affects two random chromosomes that exchange terminal and functional symbols at arbitrary points. There are three types of recombination: one-point recombination, two-point recombination, and gene recombination. One-point recombination splits the two chromosomes at the same arbitrary point and all symbols after that point will be exchanged between the chromosomes (Figure 2e). Two-point recombination separates the chromosomes at two points and the symbols between the points are exchanged (Figure 2f). Gene recombination involves the entire gene of a chromosome to be swapped with the other chromosome (Figure 2g).

After the aforementioned modifications, the newly improved chromosomes are used in the next iteration. This process is repeated for a programmed set of generations (iterations) or until a solution is reached.

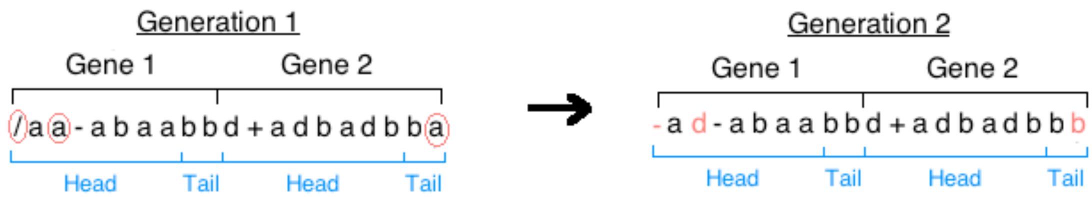


Figure 2a. Example of GEP chromosome mutation.

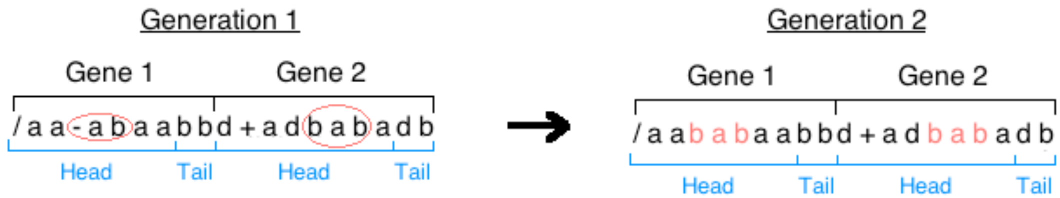


Figure 2b. Example of GEP chromosome IS transposition.

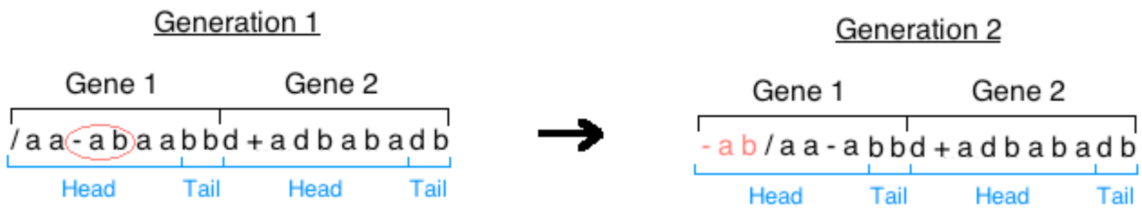


Figure 2c. Example of GEP chromosome RIS transposition.

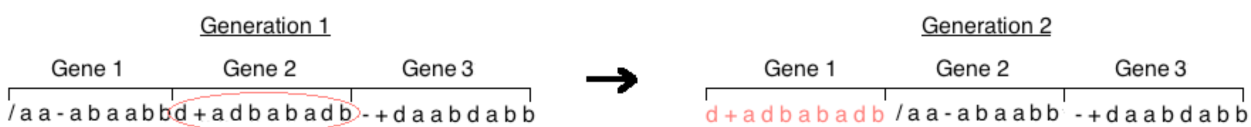


Figure 2d. Example of GEP chromosome gene transposition.

Figure 2e. Example of GEP one-point recombination with two chromosomes.

Figure 2f. Example of GEP two-point recombination with two chromosomes.

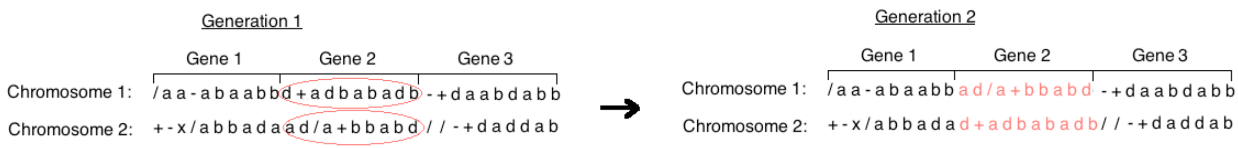


Figure 2g. Example of GEP gene recombination with two chromosomes.

2.4 Penman-Monteith (PM) Equation

The PM equation has been accepted as a reference equation for ET_0 estimation (Shahidian et al., 2012). ET_0 is defined as evapotranspiration that occurs when using a theoretical grass surface with an assumed height of 0.12 meter, a surface resistance (i.e., resistance of vapor leaving the surface) of 70 seconds per meter and a surface albedo (i.e., reflectivity of surface) of 0.23 (Allen et al, 1998). Other ET_0 equations have been calibrated based upon the PM equation (Allen et al., 1998; Shiri et al., 2013). The PM equation is given by:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma\frac{900}{T_{mean}+273}W_s(e_s - e_a)}{\Delta + \gamma(1 + 0.34W_s)} \tag{6}$$

where ET_0 is the reference evapotranspiration (mm/d), Δ is the slope of saturation vapor pressure function (kPa/°C), R_n is the net radiation (MJ/m^{2}day), G is the soil heat flux density (MJ/m^{2}day), γ is the psychrometric constant (kPa/°C), T_{mean} is the mean air temperature (°C), W_s is the daily mean wind speed at a height of 2 m (m/s), e_s is the saturation vapor pressure (kPa), and e_a is the actual vapor pressure (kPa).

3. DATA

3.1 Iran

Daily climatic data from January 1, 2000 to December 31, 2008 was collected from 8 weather stations in the coastal regions of Iran. The recorded data were daily average relative humidity, daily mean wind speed at a reference height of 2 meters, daily maximum, minimum and mean air temperatures, and incoming solar radiation. The hydrological data and the corresponding calculated ET_0 (from Equation (6)) were used to train and test the models (see Section 3.3). The locations, altitudes, longitudes and latitudes of the weather stations are shown in Table 1 and Figure 3.

Table 1. Locations, altitudes, latitudes, and longitudes of weather stations in Iran.

Station	Altitude, m	Latitude, °	Longitude, °
Abadan	6.6	30.2	48.2
Ahwaz	22.5	31.2	48.4
Bandar-e-Abbas	9.8	27.1	56.2
Bandar-e-Lenge	22.7	26.3	54.2
Bushehr	9	28.6	50.5
Gorgan	13.3	36.5	54.1
Rasht	-8.6	37.1	49.4
Sari	23	36.3	53.0

Figure 3. Location of weather stations in Iran.

3.2 California

Daily climatic data from January 1, 2015 to December 31, 2015 was collected from 14 California Irrigation Management Information System (CIMIS) weather stations. The collected data were daily average relative humidity, daily mean wind speed at a reference height of 2 meters, daily maximum, minimum and mean air temperature, and incoming solar radiation. The hydrological data and the corresponding calculated ET_0 (from Equation (6)) were used to test the models (see Section 3.3). The locations, altitudes, longitudes, and latitudes of the CIMIS weather stations are shown in Table 2 and Figure 4.

Table 2. Locations, altitudes, latitudes, and longitudes of weather stations in California.

Station	Altitude, m	Latitude, °	Longitude, °
Bishop	1271	37.4	-118.4
King City Oasis Rd.	164.6	36.1	-121.1
Blythe NE	83.8	33.7	-114.6
Atascadero	269.8	35.5	-120.6
Delano	91.4	35.8	-119.3
Gilroy	56.4	37.0	-121.5
Arleta	298.7	34.3	-118.4
Gerber South	75	40.0	-122.2
Woodland	25	38.7	-121.8
Diamond Springs	624.8	38.6	-120.8
Lompoc	16.8	34.7	-120.5
Santa Maria II	65.5	34.9	-120.5
Macdoel II	1300	41.8	-122.0
Moreno Valley	488.9	33.9	-117.2

Figure 4. Location of CIMIS weather stations.

3.3 Training and Testing Data

The training dataset contained the climatic data from the 8 stations located in Iran (Table 1) and the corresponding ET_0 estimates from the PM equation. The dataset was limited to the years 2000-2007, while the data from 2008 was reserved for testing. MARS, M5MT, and GEP used the training data to generate models (i.e., equations), which described the relationship between the input data (e.g., mean air temperature, solar radiation, etc.) from the 8 stations in Iran and the corresponding ET_0 values.

The testing dataset contained the climatic data from the same 8 stations in Iran (year 2008 only) and 14 CIMIS stations in California (year 2015) (Section 3.2). The testing dataset was used to assess the capabilities of MARS, M5MT, and GEP models to predict ET_0. These predicted ET_0 values were compared to ET_0 estimates from the PM equation (6).

The MARS, M5MT and GEP training and testing data is outlined in Table 3. In the first part of this study, MARS, M5MT, and GEP models were tested with data from same Iran stations it was trained with. In the second part of this study, MARS, M5MT, and GEP models were tested with data from 14 CIMIS stations in California.

Table 3. Training and testing data.

	Approach	**Training Data**	**Testing Data**
1st Part of Study	MARS M5MT GEP	Iran (2000-2007)	Iran (2008)
2nd Part of Study	M5MT MARS GEP	Iran (2000-2007)	California (2015)

3.3.1 Data Combinations

The PM equation requires several hydrological variables (i.e., net radiation, soil heat flux, air temperature, relative humidity, and wind speed), which are typically unavailable. As a result, different combinations of hydrological data were investigated, which were based upon equations requiring fewer climatic data to estimate ET_0 (e.g., Makkink, Romanenko, and Hargreaves-Semani).

The Makkink equation utilizes solar radiation (Makkink, 1957):

$$ET_0 = 0.61 \frac{\Delta R_s}{(\Delta + \gamma)\lambda} - 0.12 \tag{7}$$

where R_s is the incoming solar radiation (MJ/m^{2}day) and λ is the latent heat of evaporation (MJ/kg).

The Romanenko equation uses the concept water mass transfer, and requires relative humidity and air temperature (Romanenko, 1961):

$$ET_0 = 0.0018(T_{mean} + 25)^2(100 - RH) \tag{8}$$

where RH is relative humidity (%).

The Hargreaves-Samani method uses air temperature and extraterrestrial radiation (Hargreaves and Samani, 1985):

$$ET_0 = 0.0023 \frac{R_a}{\lambda}(T_{mean} + 17.8)\sqrt{T_{max} - T_{min}} \tag{9}$$

where R_a is extraterrestrial radiation (mm/d), T_{max} is the daily maximum air temperature (°C) and T_{min} is the daily minimum air temperature (°C).

Based on the required inputs in equations (6) – (9) and the study conducted by Shiri et al. (2014), the data combinations used to train and test MARS, M5MT, and GEP are:

i. Configuration 1: W_s, RH_{mean}, T_{mean}, R_s [MARS1, M5MT1, GEP1]
ii. Configuration 2: T_{mean}, R_s [MARS2, M5MT2, GEP2]
iii. Configuration 3: T_{mean}, RH_{mean} [MARS3, M5MT3, GEP3]
iv. Configuration 4: T_{mean}, T_{max}, T_{min}, R_a [MARS4, M5MT4, GEP4]

3.4 Statistical Analysis

To evaluate the performance of MARS, M5MT, and GEP models, statistical metrics such as mean absolute error (*MAE*), root mean square error (*RMSE*), and the coefficient of determination (R^2) were used and are defined as follows:

$$MAE = \frac{\sum_{i=1}^{n} |O_i - P_i|}{n} \quad (10)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (O_i - P_i)^2}{n}} \quad (11)$$

$$R^2 = \left[\frac{\sum_{i=1}^{n} (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^{n} (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^{n} (P_i - \bar{P})^2}} \right]^2 \quad (12)$$

where n is the number of observations, O_i and P_i are the i^{th} observed and estimated ET_0, respectively, and $\bar{O}$ and $\bar{P}$ are the mean of observed and simulated ET_0 values.

4. RESULTS

4.1 MARS

In this study, the MATLAB ARESLab toolbox was used. It consisted of multiple functions enabling the creation of piecewise linear and cubic regression MARS models. Parameters in ARESLab that were optimized based on trial and error were '*maxInteractions*', '*c*', and '*maxFuncs*'. The parameter '*maxInteractions*' controlled the largest amount of interactions that can occur between the data variables, the '*c*' parameter dictated smoothness of the model (the larger the value, the fewer the knots), and the '*maxFuncs*' determined the highest number of basis functions to be included in the model during the forward phase (see Equations (1a) and (1b) in Section 2.1).

Figure 5 shows the variations in *RMSE* and R^2 of ET_0 estimates from MARS1–MARS4 versus the number of basis functions. As shown, MARS1 and MARS2 generated the lowest *RMSE* and the highest R^2 with 15 basis functions. To obtain optimal values of R^2 and *RMSE,* MARS3 and MARS4 utilized 6 and 3 basis functions, respectively. Due to the large amount of training data, Friedman (1991) recommended a '*c*' value of 0, which will optimize knot production. Maximum interactivity (value of -1) between variables was allowed because it was shown to yield better results (Emamgolizadeh et al., 2015). The parameters used in each MARS model are presented in Table 4.

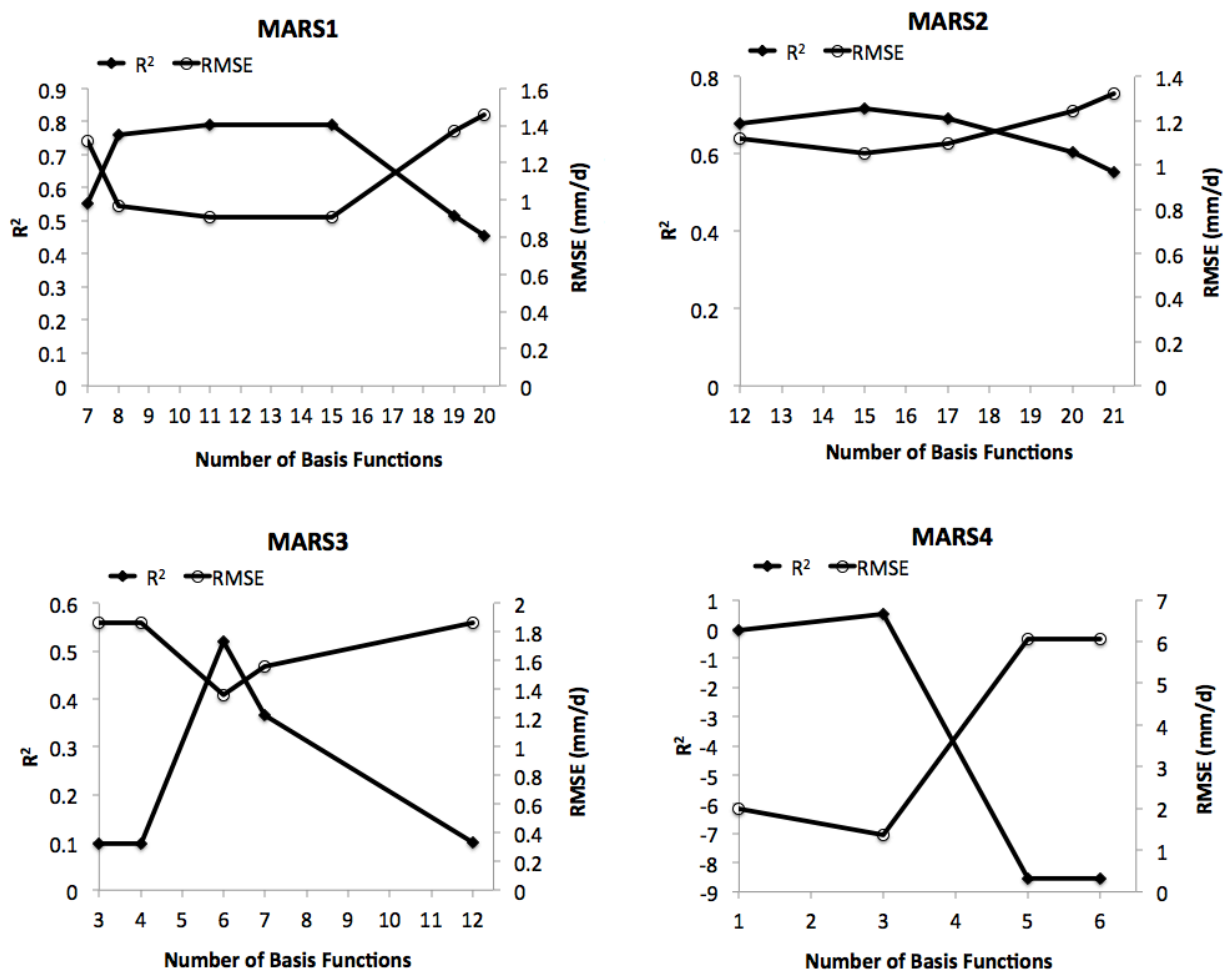


Figure 5. Variation in *RMSE* and R^2 of ET_0 estimates from MARS1–MARS4 versus number of basis functions.

Table 4. Parameter values for each MARS model.

Model	Parameters	Value
MARS1	*maxInteractions*	-1
	c	0
	maxFuncs	15
MARS2	*maxInteractions*	-1
	c	0
	maxFuncs	15
MARS3	*maxInteractions*	-1
	c	0
	maxFuncs	6
MARS4	*maxInteractions*	-1
	c	0
	maxFuncs	3

Summarized in Table 5 are the coefficients and basis functions generated by the four MARS models. The coefficients (β_i) represent the weight (i.e., importance) of the variable in terms of contributing towards the estimation of ET_0 (see Equation (3)). It is noteworthy that MARS1-MARS3 models effectively established a relationship between all input variables and ET_0. In contrast, three (i.e., T_{mean}, T_{min}, and R_a) of the four inputs in the MARS4 model were eliminated.

Figures 6-9 compare ET_0 estimates from the four MARS models with those from the PM equation for the different stations located in California and Iran. MARS1 and MARS2 yielded ET_0 values more concentrated within the vicinity of the "best fit" line compared to MARS3 and MARS4. Figures 10-13 show the time series of predicted ET_0 values from the four MARS models. Calculated ET_0 values from the PM equation were also plotted on the same figures for comparison. The general trend in Figures 10-13 suggests an underestimation of ET_0 in California. This is due to measurement errors and also the fact that none of the stations in California were used for training the MARS models. Nevertheless, the ET_0 estimates from the MARS models were able to capture the daily oscillations of the observed ET_0 values.

The performance of the MARS models was compared in Figure 14 using *MAE, RMSE* and R^2. In all stations, MARS1 provided the best results, and MARS2 provided better estimates than MARS3 and MARS4. For California stations, average *MAE* of ET_0 estimates from MARS2, MARS3, and MARS4 were 18%, 50%, and 52% greater than those of MARS1, respectively. Also, their average *RMSE* values were respectively 19%, 58%, and 57% greater than those of MARS1. When applied to the Iran stations, MARS1 resulted in the lowest *MAE* and *RMSE* values (0.44 mm/d and 0.603 mm/d, respectively) and the highest R^2 value (0.94). MARS2 tested with Iran stations resulted in better ET_0 estimates compared to MARS3 and MARS4, with lower *MAE* (0.66 mm/d) and *RMSE* (0.93 mm/d) and higher R^2 (0.84).

The statistical metrics of ET_0 estimates from MARS1-MARS4 are given in Tables 6 and 7. High variability in *MAE*, *RMSE*, and R^2 can be observed for all California and Iran stations. Compared to MARS3 and MARS4, MARS1 and MARS2 models were able to more robustly learn the relationship between the input data and ET_0 (see Tables 6 and 7). For all stations, the average *MAE* of MARS1 was 20%, 53%, and 56% less than those of MARS2, MARS3, and MARS4, respectively. Similarly, the average *RMSE* of all stations tested with MARS1 was reduced by 20%, 59%, and 61% when compared with those of MARS2, MARS3, and MARS4, respectively.

On average, MARS1 and MARS2 achieved better results than MARS3 and MARS4. The outcomes from MARS1 and MARS2 tested with California stations provided validation that MARS is a feasible approach to estimate ET_0. Furthermore, statistics from MARS2 indicated the ability of MARS to generate an equation to estimate ET_0 from limited climatic data, which can be useful for areas with limited resources.

Table 5. Basis functions and corresponding coefficients in MARS1-MARS4.

Model	**Basis Functions, $h_i(x)$**	**Coefficients, β_i**
MARS1	Intercept (β_0)	11.9
	$(R_s - 10.8)_+$	0.249
	$(10.8 - R_s)_+$	-7.63E-02
	$(6.3 - W_s)_+ \times (RH_{mean} - 31.9)_+$	1.62E-02
	$(6.3 - W_s)_+ \times (31.9 - RH_{mean})_+$	3.30E-02
	$(W_s - 6.3)_+$	0.109
	$(6.3 - W_s)_+$	-0.964
	$(RH_{mean} - 30.8)_+$	-0.126
	$(30.8 - RH_{mean})_+$	-0.242
	$(T_{mean} - 42.6)_+$	-4.19E-03
	$(42.6 - T_{mean})_+$	-0.185
	$(R_s - 10.8)_+ \times (T_{mean} - 42.6)_+$	-1.49E-05
	$(R_s - 10.8)_+ \times (42.6 - T_{mean})_+$	-5.28E-03
	$(RH_{mean} - 30.8)_+ \times (T_{mean} - 31)_+$	8.33E-04
	$(RH_{mean} - 30.8)_+ \times (31 - T_{mean})_+$	2.48E-03
MARS2	Intercept (β_0)	4.08
	$(10.8 - R_s)_+$	-5.07E-02
	$(R_s - 10.8)_+ \times (T_{mean} - 42.7)_+$	-2.46E-02
	$(R_s - 10.8)_+ \times (42.7 - T_{mean})_+$	1.11E-02
	$(42.7 - T_{mean})_+$	-0.321
	$(33.5 - T_{mean})_+$	0.233
	$(T_{mean} - 1006.2)_+$	-7.76E-02
	$(1006.2 - T_{mean})_+$	2.43E-03
	$(R_s - 10.8)_+ \times (T_{mean} - 27.2)_+$	2.45E-02
	$(R_s - 10.8)_+ \times (27.2 - T_{mean})_+$	-1.52E-02
	$(T_{mean} - 42.7)_+ \times (17.1 - R_s)_+$	-1.37E-04
MARS3	Intercept (β_0)	10.8
	$(T_{mean} - 42)_+$	-3.22E-03
	$(42 - T_{mean})_+$	-0.208
	$(RH_{mean} - 31.6)_+$	-6.67E-02
	$(31.6 - RH_{mean})_+$	-0.253
MARS4	Intercept (β_0)	0.449
	$(T_{max} - 12.2)_+$	0.221
	$(12.2 - T_{max})_+$	0.330

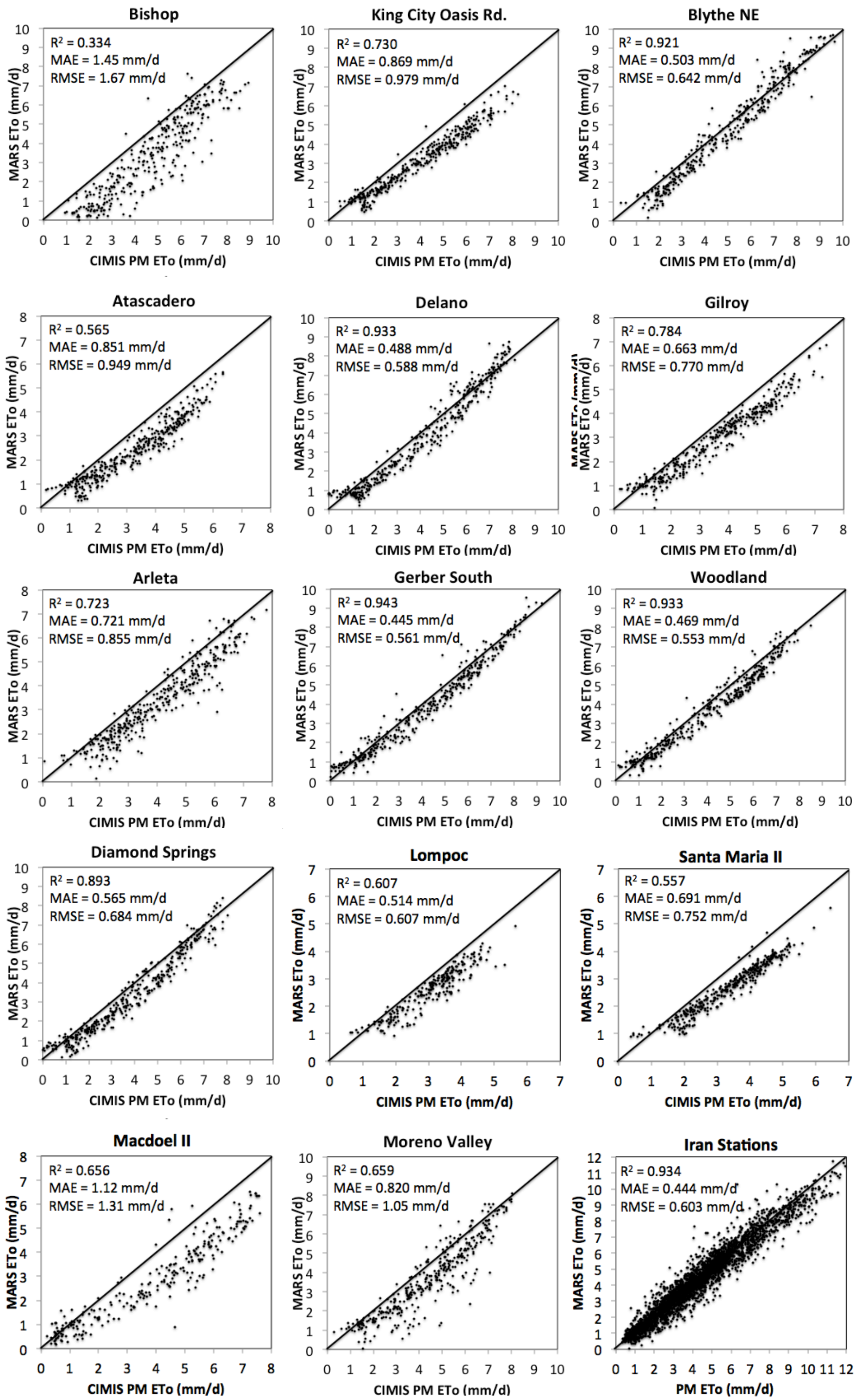


Figure 6. Estimated ET_0 values from MARS1 versus observations.

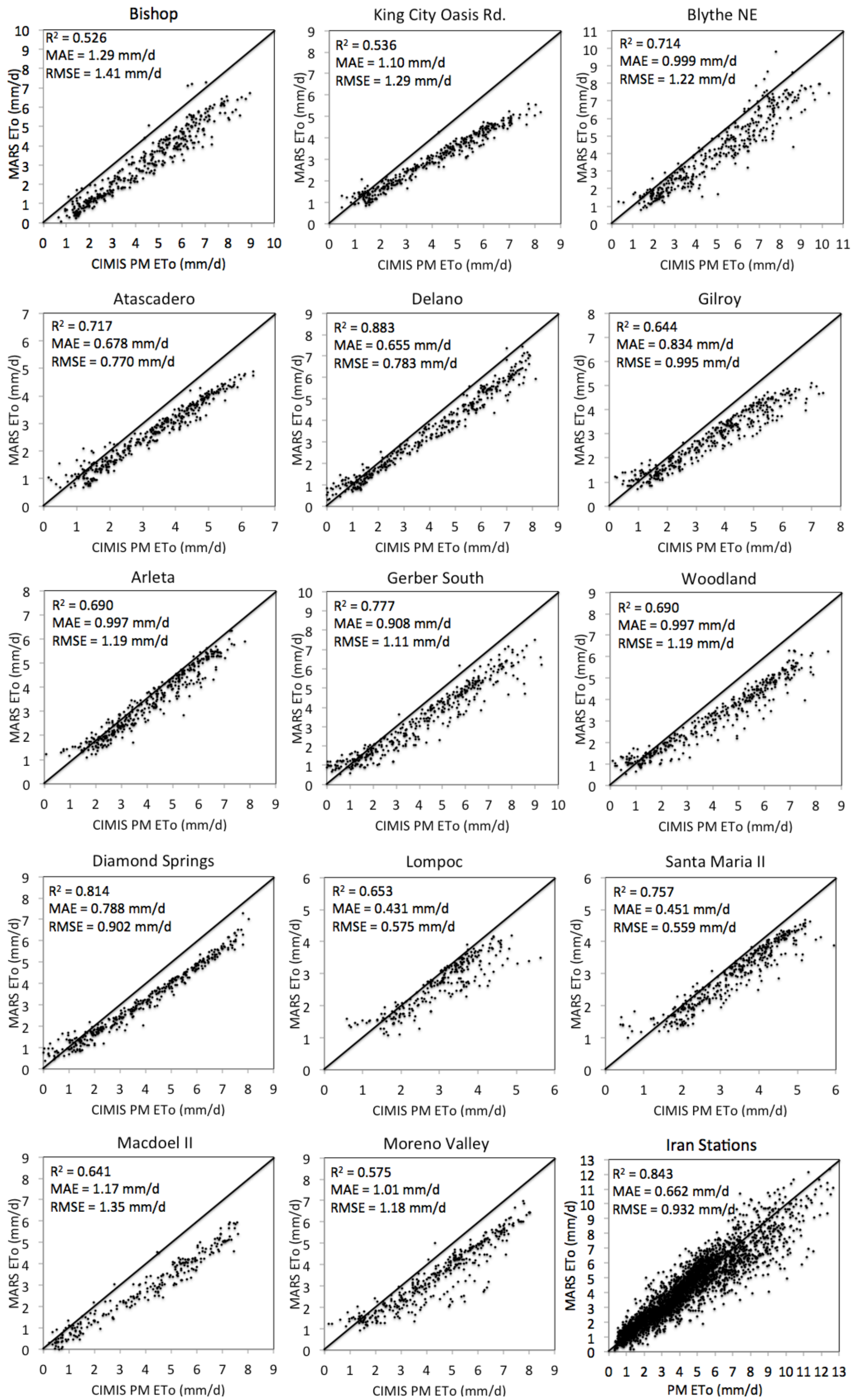


Figure 7. Estimated ET_0 values from MARS2 versus observations.

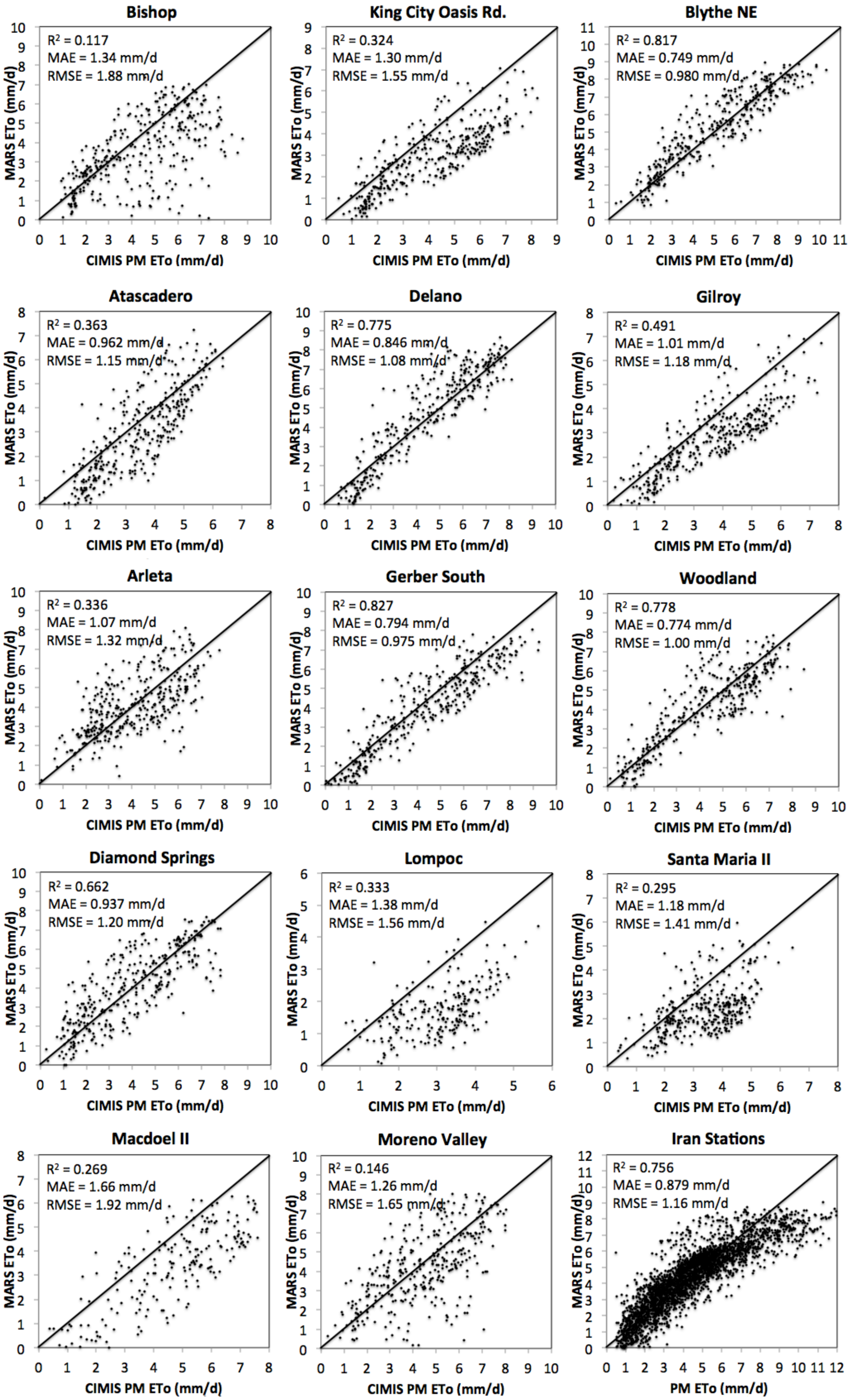


Figure 8. Estimated ET_0 values from MARS3 versus observations.

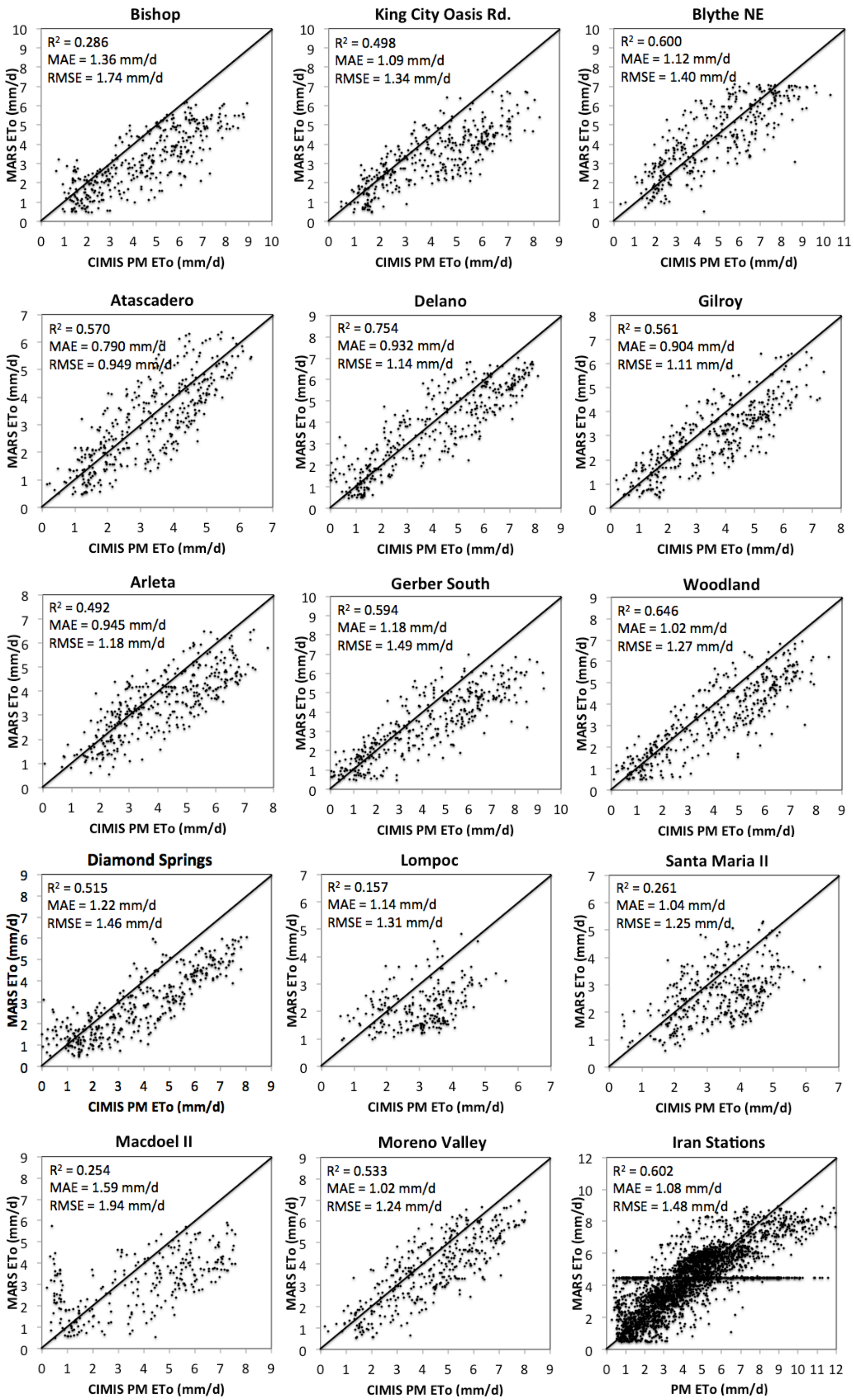
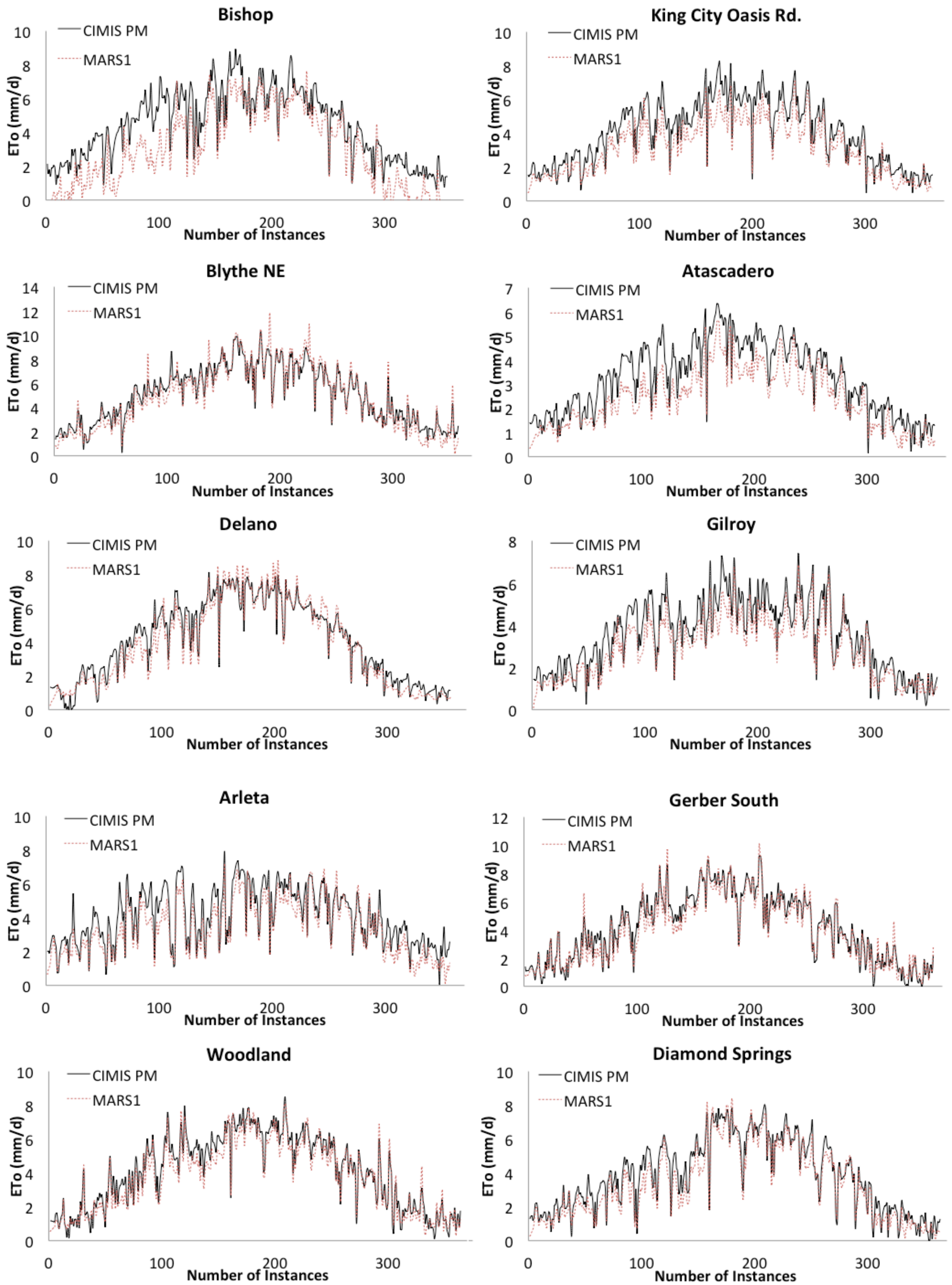


Figure 9. Estimated ET_0 values from MARS4 versus observations.


Bishop
CIMIS PM
MARS1
ETo (mm/d)
Number of Instances
King City Oasis Rd.
CIMIS PM
MARS1
ETo (mm/d)
Number of Instances
Blythe NE
CIMIS PM
MARS1
ETo (mm/d)
Number of Instances
Atascadero
CIMIS PM
MARS1
ETo (mm/d)
Number of Instances
Delano
CIMIS PM
MARS1
ETo (mm/d)
Number of Instances
Gilroy
CIMIS PM
MARS1
ETo (mm/d)
Number of Instances
Arleta
CIMIS PM
MARS1
ETo (mm/d)
Number of Instances
Gerber South
CIMIS PM
MARS1
ETo (mm/d)
Number of Instances
Woodland
CIMIS PM
MARS1
ETo (mm/d)
Number of Instances
Diamond Springs
CIMIS PM
MARS1
ETo (mm/d)
Number of Instances

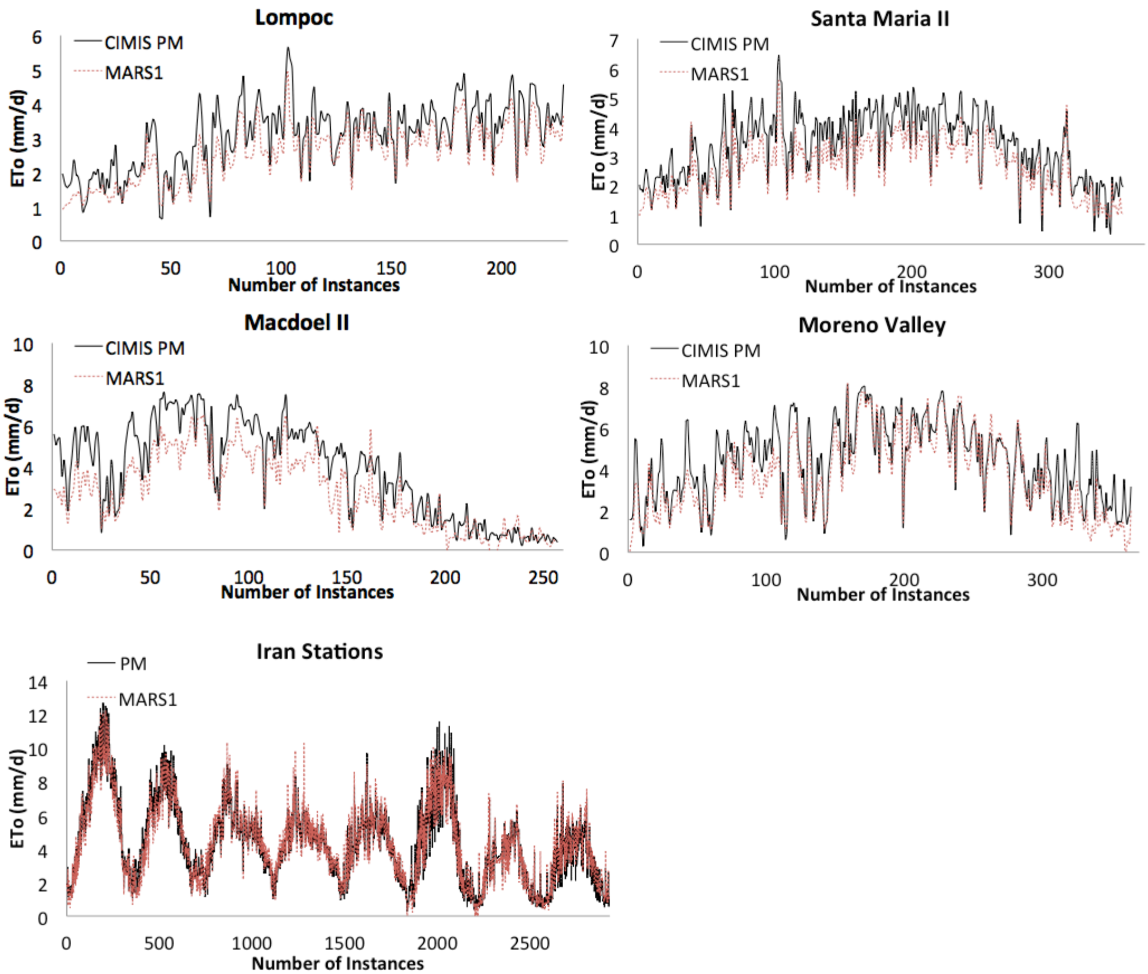
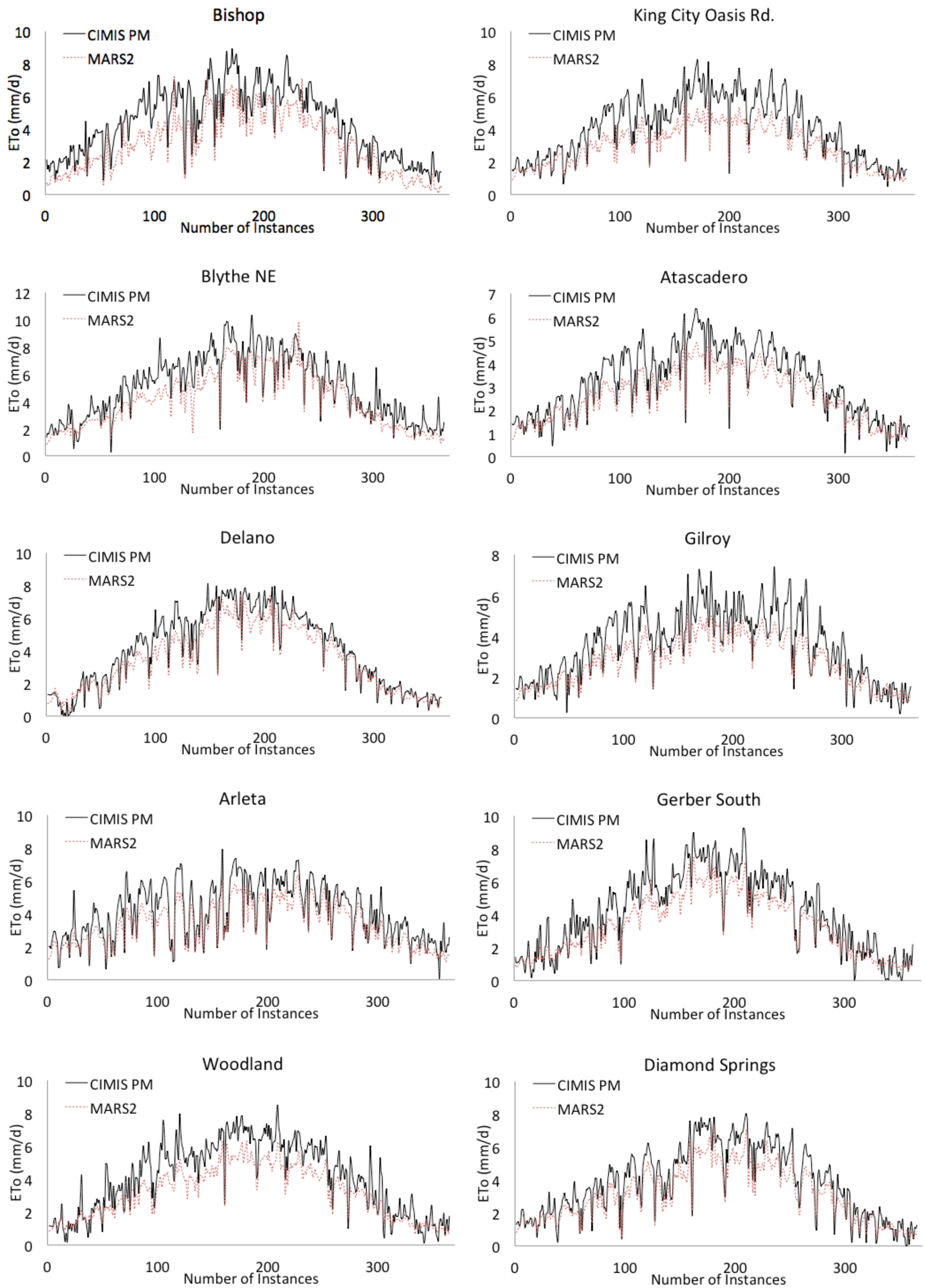


Figure 10. Time series of observed and estimated ET_0 values from MARS1.


Bishop
King City Oasis Rd.
Blythe NE
Atascadero
Delano
Gilroy
Arleta
Gerber South
Woodland
Diamond Springs
CIMIS PM
MARS2
ETo (mm/d)
Number of Instances

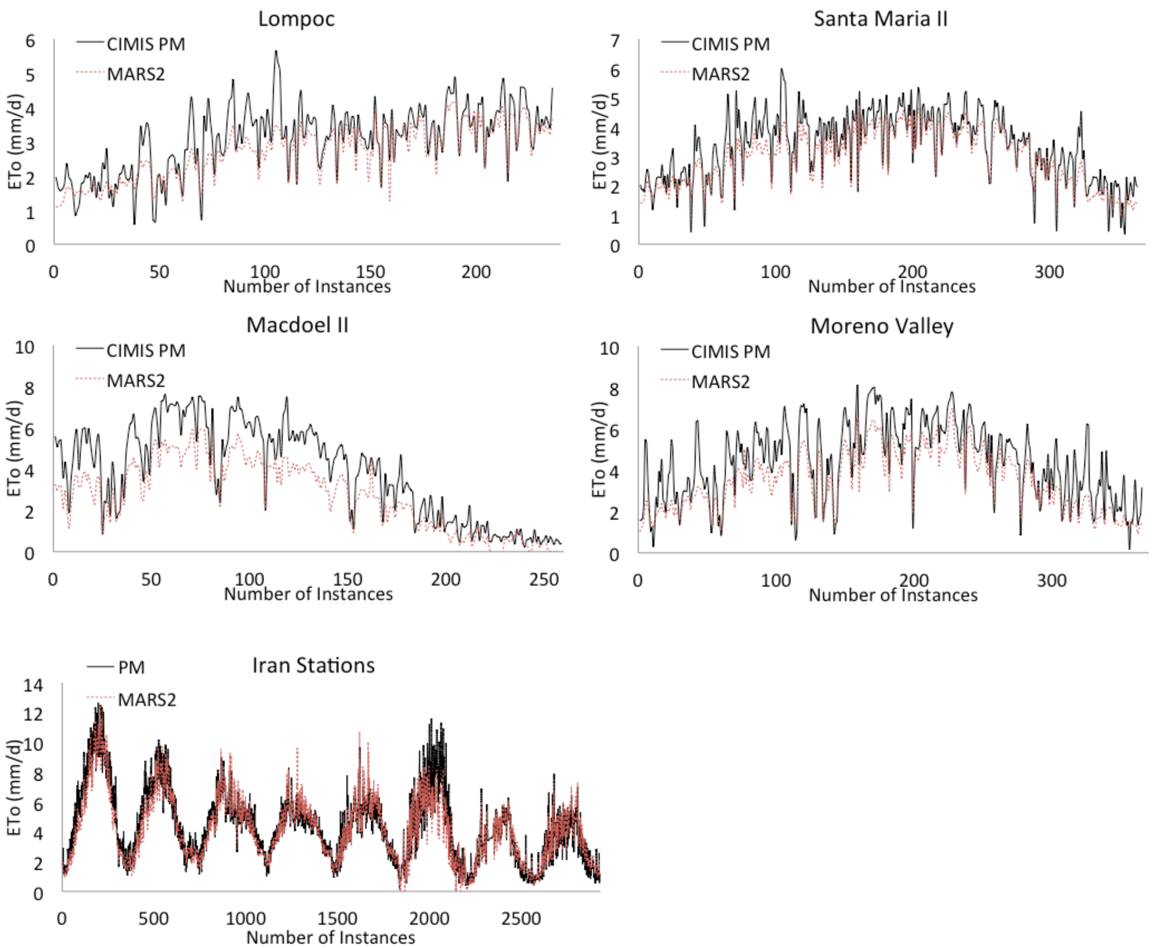
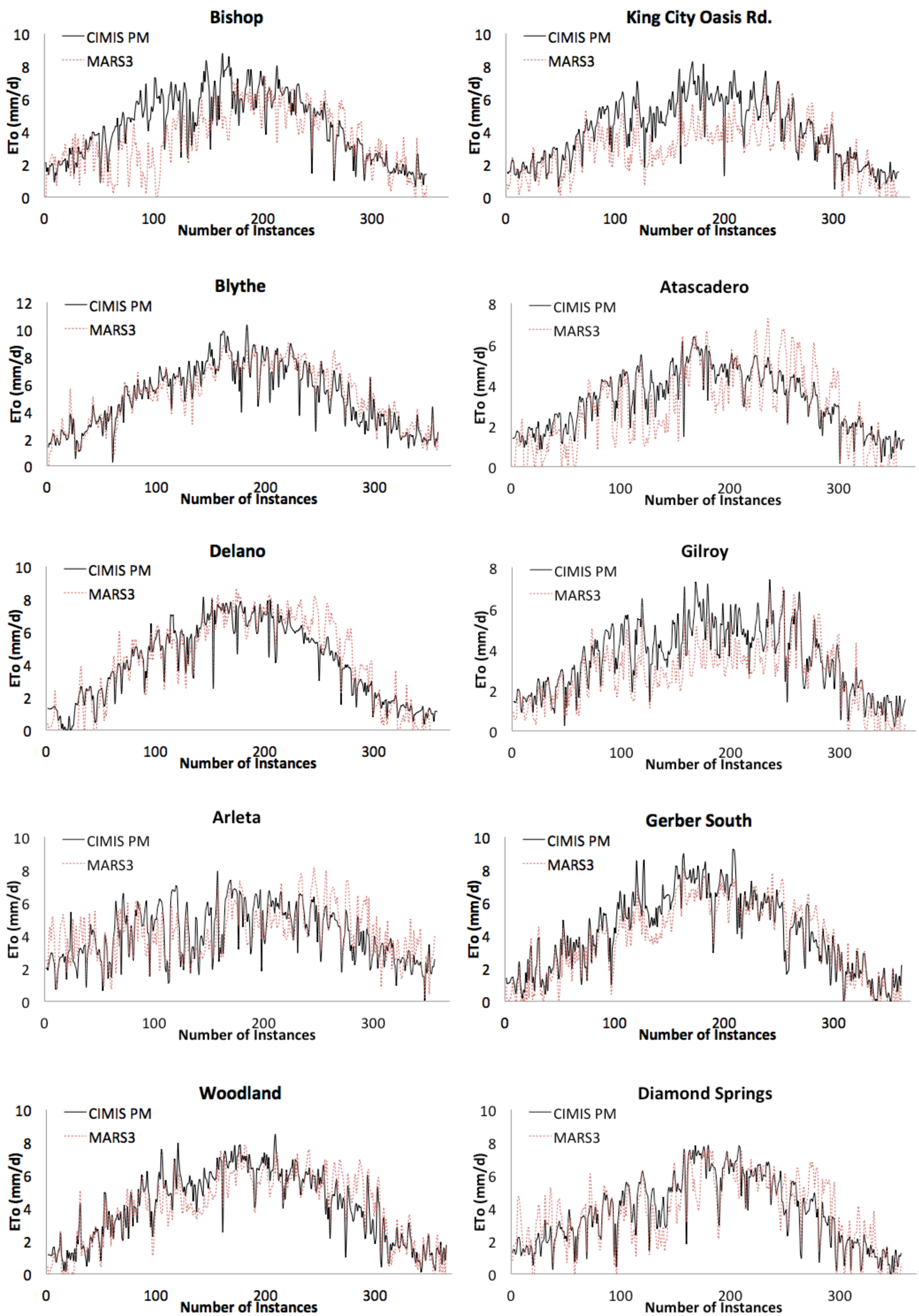


Figure 11. Time series of observed and estimated ET_0 values from MARS2.


Bishop
King City Oasis Rd.
Blythe
Atascadero
Delano
Gilroy
Arleta
Gerber South
Woodland
Diamond Springs
CIMIS PM
MARS3
ETo (mm/d)
Number of Instances

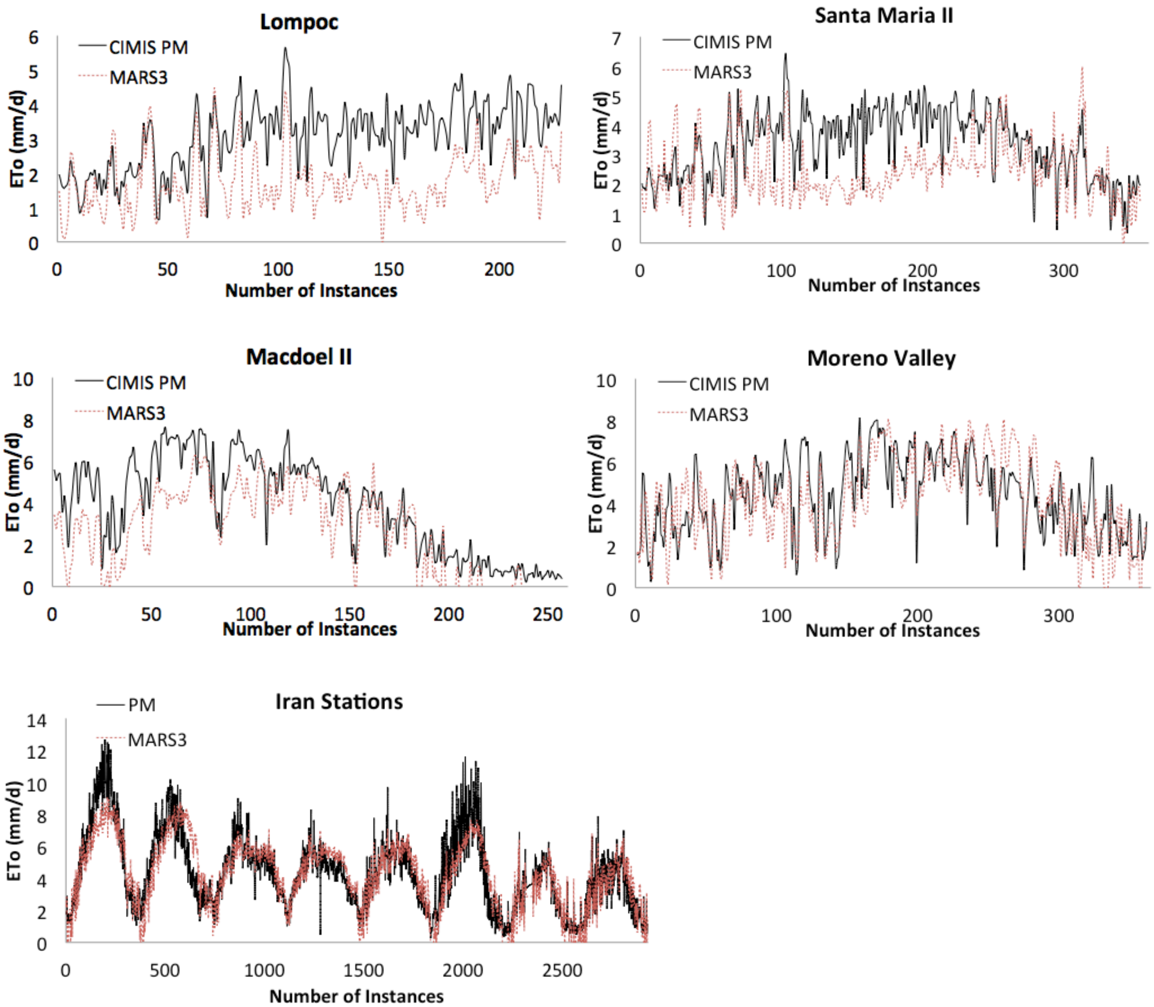
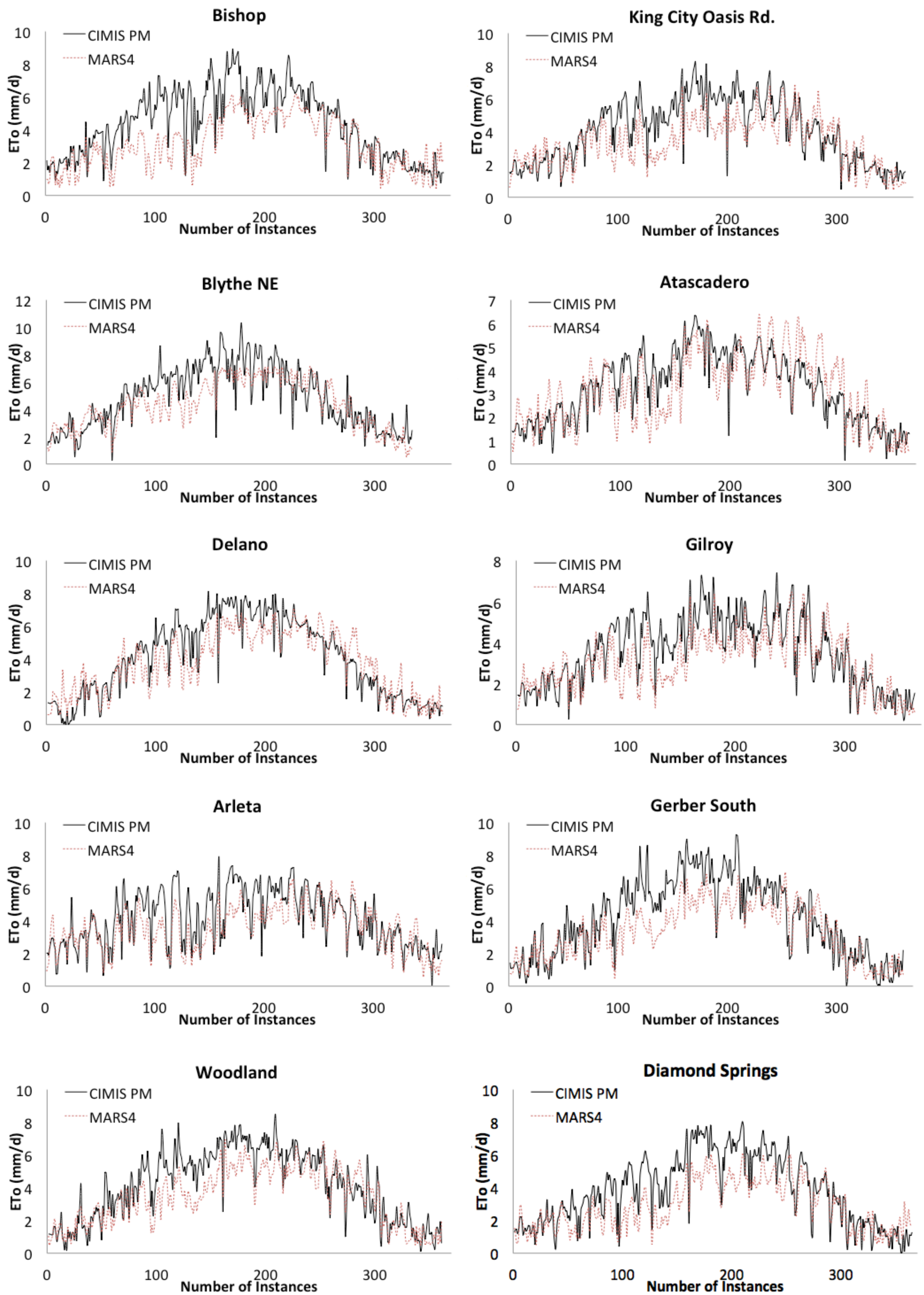


Figure 12. Time series of observed and estimated ET_0 values from MARS3.


Bishop
King City Oasis Rd.
Blythe NE
Atascadero
Delano
Gilroy
Arleta
Gerber South
Woodland
Diamond Springs
CIMIS PM
MARS4
ETo (mm/d)
Number of Instances

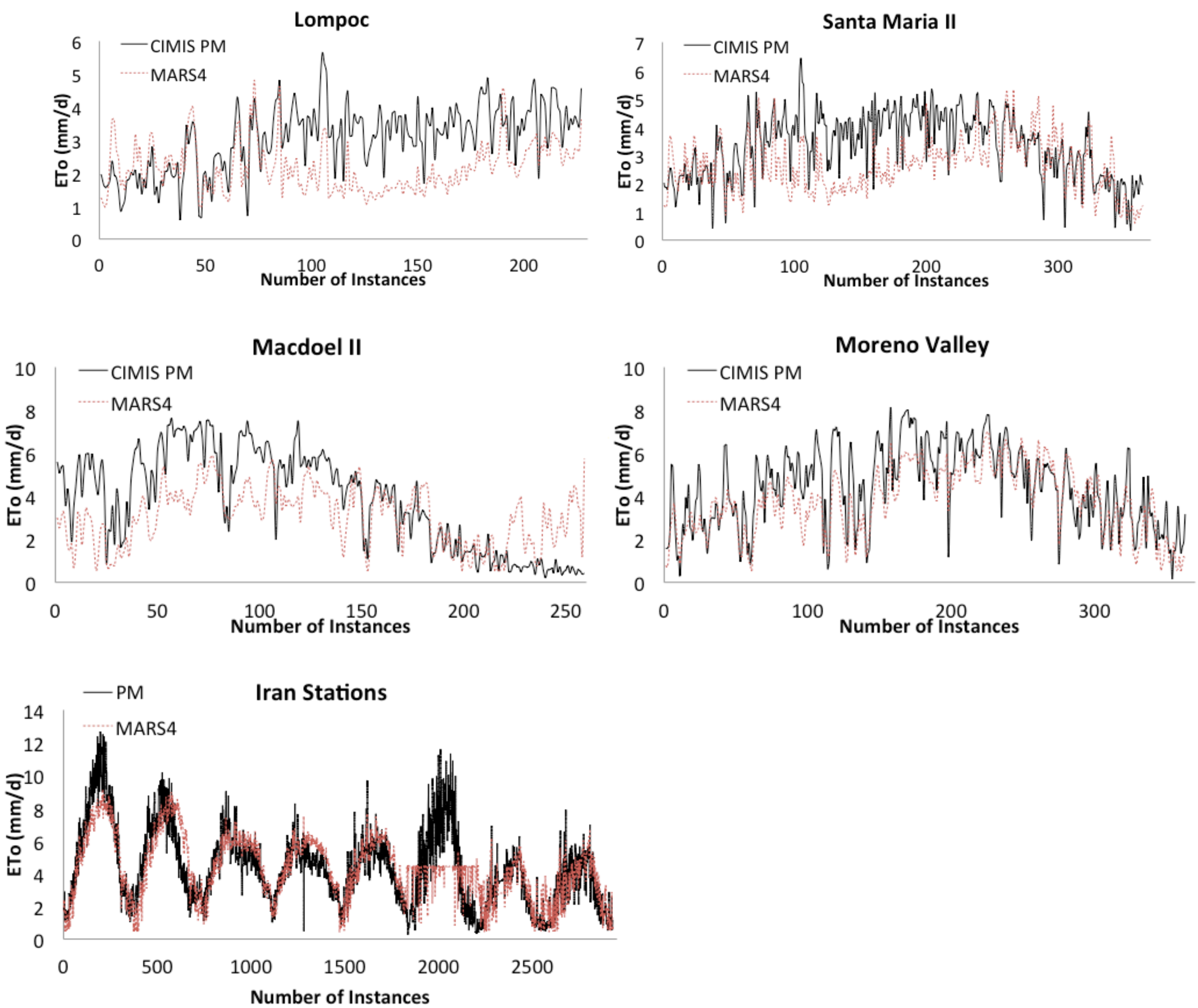


Figure 13. Time series of observed and estimated ET_0 values from MARS4.

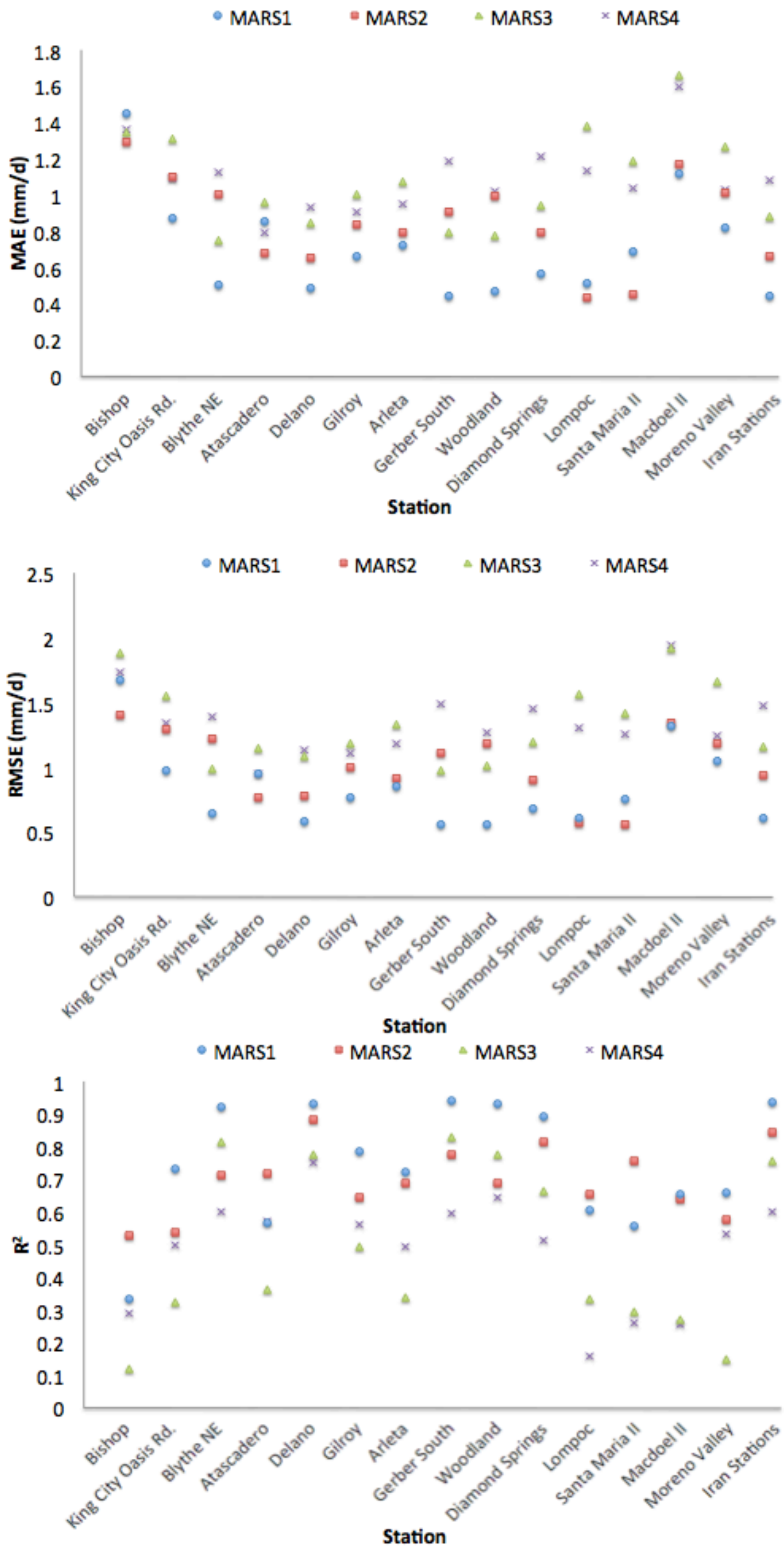


Figure 14. Comparing performance of MARS models in different stations.

Table 6. Statistical metrics of ET_0 estimates from MARS1 and MARS2.

Model	Station	MAE (mm/d)	RMSE (mm/d)	R^2
MARS1	Bishop	1.45	1.67	0.33
	King City Oasis Rd.	0.87	0.98	0.73
	Blythe NE	0.50	0.64	0.92
	Atascadero	0.85	0.95	0.57
	Delano	0.49	0.59	0.93
	Gilroy	0.66	0.77	0.78
	Arleta	0.72	0.86	0.72
	Gerber South	0.44	0.56	0.94
	Woodland	0.47	0.55	0.93
	Diamond Springs	0.56	0.68	0.89
	Lompoc	0.51	0.61	0.61
	Santa Maria II	0.69	0.75	0.56
	Macdoel II	1.12	1.31	0.66
	Moreno Valley	0.82	1.05	0.66
	All Iran Stations	0.44	0.60	0.93
MARS2	Bishop	1.29	1.41	0.53
	King City Oasis Rd.	1.10	1.29	0.54
	Blythe NE	1.00	1.22	0.71
	Atascadero	0.68	0.77	0.72
	Delano	0.65	0.78	0.88
	Gilroy	0.83	0.99	0.64
	Arleta	0.79	0.92	0.69
	Gerber South	0.91	1.11	0.78
	Woodland	1.00	1.19	0.69
	Diamond Springs	0.79	0.90	0.81
	Lompoc	0.43	0.57	0.65
	Santa Maria II	0.45	0.56	0.76
	Macdoel II	1.17	1.35	0.64
	Moreno Valley	1.01	1.18	0.58
	All Iran Stations	0.66	0.93	0.84

Table 7. Statistical metrics of ET_0 estimates from MARS3 and MARS4.

Model	Station	MAE (mm/d)	RMSE (mm/d)	R^2
MARS3	Bishop	1.34	1.88	0.12
	King City Oasis Rd.	1.30	1.55	0.32
	Blythe NE	0.75	0.98	0.82
	Atascadero	0.96	1.15	0.36
	Delano	0.85	1.08	0.78
	Gilroy	1.01	1.18	0.49
	Arleta	1.07	1.32	0.34
	Gerber South	0.79	0.97	0.83
	Woodland	0.77	1.00	0.78
	Diamond Springs	0.94	1.20	0.66
	Lompoc	1.38	1.56	0.33
	Santa Maria II	1.18	1.41	0.29
	Macdoel II	1.66	1.92	0.27
	Moreno Valley	1.26	1.65	0.15
	All Iran Stations	0.88	1.16	0.76
MARS4	Bishop	1.36	1.74	0.29
	King City Oasis Rd.	1.09	1.34	0.50
	Blythe NE	1.12	1.40	0.60
	Atascadero	0.79	0.95	0.57
	Delano	0.93	1.14	0.75
	Gilroy	0.90	1.11	0.56
	Arleta	0.95	1.18	0.49
	Gerber South	1.18	1.49	0.59
	Woodland	1.02	1.27	0.65
	Diamond Springs	1.22	1.46	0.51
	Lompoc	1.14	1.31	0.16
	Santa Maria II	1.04	1.25	0.26
	Macdoel II	1.59	1.94	0.25
	Moreno Valley	1.02	1.24	0.53
	All Iran Stations	1.08	1.48	0.60

4.2 M5MT

To create a model tree, a data mining software known as WEKA (Waikato Environment for Knowledge Analysis) was used (Frank et al., 2016). This software contained a variety of machine learning algorithms, including M5MT. To assist with the modeling process, WEKA provided percentage split, train-test, and cross validation testing options.

The percentage split option partitioned input data into training and testing datasets based on a user specified percentage. In contrast, the train-test option involved manually inputting two sets of data (i.e., training and testing) separately.

The cross validation option involved using one set of data, which was broken up by a user specified amount of folds. For example, if 10 folds were specified, the software would split up the dataset into 10 sections. The software would then utilize 9 out of the 10 sections for model training and the remaining section for testing. The testing process would be repeated until all 10 sections have been utilized for testing.

Among the three abovementioned options, the train-test method was chosen due to its superior performance compared to the other options (see Table 8).

Table 8. Performance of M5MT for different testing options.

Methods	MAE (mm/d)	RMSE (mm/d)	R^2
Percentage Split	0.2550	0.3422	0.9904
Train-Test	0.2328	0.3189	0.9914
Cross Validation	0.2396	0.3337	0.9906

Presented in Tables 9-12 are the model trees from M5MT1-M5MT4 (left column) and corresponding linear models (right column). To interpret results in Tables 9-12, readers are referred to Appendix A. It is worth mentioning all M5MT models successfully found a relationship between all input variables and ET_0.

Figures 15-18 show a comparison of ET_0 estimates from the four M5MT models with those from the PM equation for all stations. M5MT1 and M5MT2 yielded ET_0 values closer to the 1:1 line compared to M5MT3 and M5MT4. Figures 19-22 illustrate the time series of estimated ET_0 from the four M5MT models. For comparison, observed ET_0 values from the PM equation are also graphed on the same figures. As shown, M5MT1-M5MT3 underestimated ET_0 in several California stations and M5MT4 overestimated ET_0 for most of the California stations. This is mainly due to measurement errors, and also the fact that no data from the California stations was used in the training process. Nonetheless, the predicted ET_0 values mirrored the daily fluctuations of the observed ET_0 values.

Figure 23 compare the performance of M5MT models using *MAE*, *RMSE*, and R^2. In all stations, M5MT1 performed the best. For California stations, average *MAE* of ET_0 estimates from M5MT2, M5MT3, and M5MT4 were respectively 16%, 47%, and 71% greater than that of M5MT1. Also, their average *RMSE* values were respectively 17%, 67%, and 93% larger than that of M5MT1. For Iran stations, M5MT1 had the lowest *MAE* and *RMSE* values (0.38 mm/d and 0.55 mm/d, respectively) and the highest R^2 (0.95). M5MT4 showed better results than M5MT2 and M5MT3 when tested with the Iran stations, with lower *MAE* (0.56 mm/d) and *RMSE* (0.82 mm/d), and higher R^2 (0.88).

Tables 13 and 14 show the statistical metrics of ET_0 estimates for the four M5MT models. The mean *MAE, RMSE,* and R^2 for all stations tested with M5MT1 were 0.76 mm/d, 0.94 mm/d, and 0.86, respectively. Average *MAE* values from all stations tested with M5MT2, M5MT3, and M5MT4 were respectively, 18%, 49%, and 70% greater than those of M5MT1. Moreover, *RMSE* values from all the stations tested with M5MT1 were respectively, 17%, 66%, and 89% less than those of M5MT2, M5MT3, and M5MT4.

Overall, M5MT1 outperformed M5MT2-M5MT4. The results from M5MT1 and M5MT2 tested with California stations showed M5MT to be an effective approach to estimate ET_0. The outcomes of M5MT2 also demonstrated M5MT's ability to estimate ET_0 with limited climatic data. M5MT4 model performance varied depending on the location of the testing data. M5MT4 performed better than M5MT2 and M5MT3 when tested with data from Iran. While, M5MT4 provided inaccurate ET_0 estimates, compared to M5MT1-M5MT3, when tested with California stations. This may suggest the inputs of M5MT4 to be region dependent.

Table 9. M5MT1 model tree and corresponding linear models.

<table>
<tr>
<td>
Rs ≤ 16.271 :
| Tmean ≤ 19.175 :
| | Tmean ≤ 11.65 :
| | | Ws ≤ 1.496 :
| | | | Rs ≤ 11.945 : LM1
| | | | Rs > 11.945 :
| | | | | RHmean ≤ 67.641 : LM2
| | | | | RHmean > 67.641 :
| | | | | | Tmean ≤ 9.1 : LM3
| | | | | | Tmean > 9.1 :
| | | | | | | Rs ≤ 13.929 :
| | | | | | | | RHmean ≤ 88.229 : LM4
| | | | | | | | RHmean > 88.229 : LM5
| | | | | | | Rs > 13.929 :
| | | | | | | | Tmean ≤ 9.75 : LM6
| | | | | | | | Tmean > 9.75 : LM7
| | | Ws > 1.496 :
| | | | RHmean ≤ 83.058 :
| | | | | Tmean ≤ 8.25 :
| | | | | | Ws ≤ 1.846 :
| | | | | | | Rs ≤ 10.086 : LM8
| | | | | | | Rs > 10.086 : LM9
| | | | | | Ws > 1.846 :
| | | | | | | RHmean ≤ 65.147 :
| | | | | | | | RHmean ≤ 56.173 :
| | | | | | | | | Rs ≤ 11.305 : LM10
| | | | | | | | | Rs > 11.305 : LM11
| | | | | | | | RHmean > 56.173 :
| | | | | | | | | Rs ≤ 11.237 : LM12
| | | | | | | | | Rs > 11.237 : LM13
| | | | | | | RHmean > 65.147 : LM14
| | | | | Tmean > 8.25 :
| | | | | | RHmean ≤ 67.436 :
| | | | | | | RHmean ≤ 59.008 : LM15
| | | | | | | RHmean > 59.008 : LM16
| | | | | | RHmean > 67.436 : LM17
| | | | RHmean > 83.058 :
| | | | | Rs ≤ 6.679 : LM18
| | | | | Rs > 6.679 :
| | | | | | Tmean ≤ 8.65 : LM19
| | | | | | Tmean > 8.65 : LM20
| | Tmean > 11.65 :
| | | Ws ≤ 2.127 :
| | | | Rs ≤ 12.346 : LM21
| | | | Rs > 12.346 :
| | | | | Rs ≤ 13.707 : LM22
| | | | | Rs > 13.707 :
| | | | | | Ws ≤ 0.625 : LM23
| | | | | | Ws > 0.625 :
| | | | | | | RHmean ≤ 80.479 :
| | | | | | | | Ws ≤ 1.496 : LM24
| | | | | | | | Ws > 1.496 :
| | | | | | | | | Tmean ≤ 14.25 : LM25
| | | | | | | | | Tmean > 14.25 :
</td>
<td>
LM num: 1
ET_0 = 0.2167 * Ws + 0.0057 * RHmean + 0.0523 * Tmean + 0.0668 * Rs - 0.7314

LM num: 2
ET_0 = 0.2462 * Ws + 0.0227 * RHmean + 0.049 * Tmean + 0.1254 * Rs - 2.4922

LM num: 3
ET_0 = 0.0174 * Ws - 0.0002 * RHmean + 0.0772 * Tmean + 0.0423 * Rs + 0.1846

LM num: 4
ET_0 = -0.025 * Ws + 0.0089 * RHmean + 0.047 * Tmean + 0.0736 * Rs - 0.5938

LM num: 5
ET_0 = -0.2025 * Ws + 0.0001 * RHmean + 0.047 * Tmean + 0.0736 * Rs + 0.1638

LM num: 6
ET_0 = 0.0174 * Ws + 0.0001 * RHmean + 0.1245 * Tmean + 0.157 * Rs - 1.7991

LM num: 7
ET_0 = 0.0174 * Ws + 0.0001 * RHmean + 0.1006 * Tmean + 0.1336 * Rs - 1.1001

LM num: 8
ET_0 = 0.0399 * Ws - 0.0127 * RHmean + 0.0441 * Tmean + 0.0292 * Rs + 1.3997

LM num: 9
ET_0 = 0.0399 * Ws + 0.0065 * RHmean + 0.0865 * Tmean + 0.0549 * Rs - 0.4197

LM num: 10
ET_0 = -0.0108 * Ws - 0.0463 * RHmean + 0.0107 * Tmean - 0.0454 * Rs + 5.1093

LM num: 11
ET_0 = -0.0108 * Ws - 0.0463 * RHmean + 0.0107 * Tmean - 0.0454 * Rs + 5.0685

LM num: 12
ET_0 = 0.0111 * Ws - 0.0407 * RHmean + 0.0107 * Tmean - 0.0213 * Rs + 4.3542

LM num: 13
ET_0 = 0.0111 * Ws - 0.0407 * RHmean + 0.0107 * Tmean - 0.0213 * Rs + 4.2997
</td>
</tr>
</table>

```
| | | | | | | | | | | Tmean ≤ 17.45 : LM26
| | | | | | | | | | | Tmean > 17.45 : LM27
| | | | | | | | RHmean > 80.479 : LM28
| | | Ws > 2.127 :
| | | | RHmean ≤ 76.486 : LM29
| | | | RHmean > 76.486 : LM30
| Tmean > 19.175 :
| | RHmean ≤ 15.45 : LM31
| | RHmean > 15.45 :
| | | Tmean ≤ 25.45 :
| | | | Ws ≤ 2.149 : LM32
| | | | Ws > 2.149 :
| | | | | RHmean ≤ 74.723 : LM33
| | | | | RHmean > 74.723 : LM34
| | | Tmean > 25.45 :
| | | | Ws ≤ 1.66 :
| | | | | Ws ≤ 0.898 :
| | | | | | Rs ≤ 12.007 : LM35
| | | | | | Rs > 12.007 :
| | | | | | | Ws ≤ 0.438 : LM36
| | | | | | | Ws > 0.438 :
| | | | | | | | Tmean ≤ 32.65 : LM37
| | | | | | | | Tmean > 32.65 : LM38
| | | | | Ws > 0.898 :
| | | | | | Tmean ≤ 34.75 : LM39
| | | | | | Tmean > 34.75 :
| | | | | | | RHmean ≤ 23.35 : LM40
| | | | | | | RHmean > 23.35 :
| | | | | | | | Tmean ≤ 1010.35 : LM41
| | | | | | | | Tmean > 1010.35 : LM42
| | | | Ws > 1.66 :
| | | | | RHmean ≤ 69.19 :
| | | | | | RHmean ≤ 23.9 : LM43
| | | | | | RHmean > 23.9 :
| | | | | | | Ws ≤ 2.999 :
| | | | | | | | Tmean ≤ 33.25 : LM44
| | | | | | | | Tmean > 33.25 :
| | | | | | | | | Tmean ≤ 1005.25 : LM45
| | | | | | | | | Tmean > 1005.25 : LM46
| | | | | | | Ws > 2.999 :
| | | | | | | | Tmean ≤ 32.15 : LM47
| | | | | | | | Tmean > 32.15 : LM48
| | | | | RHmean > 69.19 :
| | | | | | Tmean ≤ 29.45 : LM49
| | | | | | Tmean > 29.45 :
| | | | | | | RHmean ≤ 81.718 :
| | | | | | | | Ws ≤ 5.061 : LM50
| | | | | | | | Ws > 5.061 :
| | | | | | | | | Tmean ≤ 33.7 : LM51
| | | | | | | | | Tmean > 33.7 :
| | | | | | | | | | Tmean ≤ 35.3 : LM52
| | | | | | | | | | Tmean > 35.3 : LM53
| | | | | | | RHmean > 81.718 : LM54
Rs > 16.271 :
| Tmean ≤ 28.837 :
| | Tmean ≤ 21.15 :
```

LM num: 14
ET_0 = 0.1323 * Ws - 0.0062 * RHmean + 0.0107 * Tmean + 0.0467 * Rs + 0.9626

LM num: 15
ET_0 = 0.0545 * Ws - 0.0087 * RHmean + 0.2098 * Tmean - 0.0576 * Rs + 1.1154

LM num: 16
ET_0 = 0.2282 * Ws - 0.0069 * RHmean + 0.0368 * Tmean + 0.0041 * Rs + 1.4223

LM num: 17
ET_0 = 0.1653 * Ws - 0.0242 * RHmean + 0.0474 * Tmean + 0.0794 * Rs + 1.7702

LM num: 18
ET_0 = 0.0111 * Ws - 0.0029 * RHmean + 0.047 * Tmean + 0.104 * Rs + 0.2682

LM num: 19
ET_0 = 0.1052 * Ws - 0.0364 * RHmean + 0.0153 * Tmean + 0.0061 * Rs + 3.8513

LM num: 20
ET_0 = 0.0171 * Ws - 0.0049 * RHmean + 0.0151 * Tmean + 0.0061 * Rs + 1.6196

LM num: 21
ET_0 = 0.3303 * Ws - 0.0034 * RHmean + 0.0716 * Tmean + 0.0718 * Rs - 0.3503

LM num: 22
ET_0 = 0.3061 * Ws - 0.0004 * RHmean + 0.0572 * Tmean + 0.0112 * Rs + 0.556

LM num: 23
ET_0 = 0.2704 * Ws + 0.0136 * RHmean + 0.0413 * Tmean + 0.2062 * Rs - 2.7158

LM num: 24
ET_0 = 0.0434 * Ws - 0.0007 * RHmean + 0.0318 * Tmean + 0.0748 * Rs + 0.6326

LM num: 25
ET_0 = -0.0315 * Ws - 0.0007 * RHmean + 0.0202 * Tmean + 0.1944 * Rs - 0.7633

LM num: 26
ET_0 = 0.018 * Ws - 0.0007 * RHmean + 0.0094 * Tmean + 0.0451 * Rs + 1.7517

LM num: 27
ET_0 = 0.018 * Ws - 0.0007 * RHmean + 0.009 * Tmean + 0.1149 * Rs + 0.6364

<table>
<tr>
<td>| | | Tmean ≤ 15.05 :
| | | | Tmean ≤ 10.35 :
| | | | | RHmean ≤ 75.984 : LM55
| | | | | RHmean > 75.984 : LM56
| | | | Tmean > 10.35 :
| | | | | RHmean ≤ 74.45 :
| | | | | | Rs ≤ 19.273 : LM57
| | | | | | Rs > 19.273 :
| | | | | | | Tmean ≤ 13.15 : LM58
| | | | | | | Tmean > 13.15 :
| | | | | | | | RHmean ≤ 60.257 : LM59
| | | | | | | | RHmean > 60.257 : LM60
| | | | | RHmean > 74.45 : LM61
| | | Tmean > 15.05 :
| | | | Rs ≤ 19.827 :
| | | | | Ws ≤ 2.398 : LM62
| | | | | Ws > 2.398 :
| | | | | | Tmean ≤ 18.45 :
| | | | | | | RHmean ≤ 78.421 : LM63
| | | | | | | RHmean > 78.421 : LM64
| | | | | | Tmean > 18.45 :
| | | | | | | RHmean ≤ 67.942 :
| | | | | | | | RHmean ≤ 59.788 : LM65
| | | | | | | | RHmean > 59.788 : LM66
| | | | | | | RHmean > 67.942 :
| | | | | | | | RHmean ≤ 77.224 : LM67
| | | | | | | | RHmean > 77.224 : LM68
| | | | Rs > 19.827 :
| | | | | RHmean ≤ 75.239 :
| | | | | | Tmean ≤ 18.25 : LM69
| | | | | | Tmean > 18.25 :
| | | | | | | RHmean ≤ 66.375 :
| | | | | | | | Rs ≤ 23.347 : LM70
| | | | | | | | Rs > 23.347 :
| | | | | | | | | Rs ≤ 27.03 : LM71
| | | | | | | | | Rs > 27.03 : LM72
| | | | | | | RHmean > 66.375 : LM73
| | | | | RHmean > 75.239 : LM74
| | Tmean > 21.15 :
| | | Rs ≤ 21.308 :
| | | | Ws ≤ 2.668 :
| | | | | Tmean ≤ 26.15 :
| | | | | | Rs ≤ 19.202 :
| | | | | | | Ws ≤ 1.66 : LM75
| | | | | | | Ws > 1.66 : LM76
| | | | | | Rs > 19.202 :
| | | | | | | Ws ≤ 1.496 :
| | | | | | | | Rs ≤ 19.854 : LM77
| | | | | | | | Rs > 19.854 : LM78
| | | | | | | Ws > 1.496 :
| | | | | | | | RHmean ≤ 82.145 : LM79
| | | | | | | | RHmean > 82.145 : LM80
| | | | | Tmean > 26.15 : LM81
| | | | Ws > 2.668 : LM82
| | | Rs > 21.308 :
| | | | RHmean ≤ 70.957 :</td>
<td>LM num: 28
ET_0 = -0.1899 * Ws - 0.024 * RHmean + 0.067 * Tmean + 0.0233 * Rs + 3.0864

LM num: 29
ET_0 = 0.161 * Ws - 0.0384 * RHmean + 0.0944 * Tmean + 0.0399 * Rs + 2.7531

LM num: 30
ET_0 = 0.0503 * Ws - 0.0366 * RHmean + 0.0768 * Tmean + 0.0822 * Rs + 2.8575

LM num: 31
ET_0 = 0.2555 * Ws + 0.0571 * RHmean - 0.005 * Tmean + 0.1768 * Rs + 4.1943

LM num: 32
ET_0 = 0.267 * Ws - 0.0094 * RHmean + 0.0744 * Tmean + 0.1143 * Rs - 0.3165

LM num: 33
ET_0 = 0.2884 * Ws - 0.0618 * RHmean + 0.1621 * Tmean + 0.0723 * Rs + 2.1154

LM num: 34
ET_0 = 0.1218 * Ws - 0.0345 * RHmean + 0.0986 * Tmean + 0.0842 * Rs + 1.9539

LM num: 35
ET_0 = 0.8404 * Ws - 0.0054 * RHmean - 0.0008 * Tmean + 0.2178 * Rs + 0.056

LM num: 36
ET_0 = 1.1958 * Ws + 0.0035 * RHmean - 0.0001 * Tmean + 0.1544 * Rs + 0.0295

LM num: 37
ET_0 = 0.7798 * Ws - 0.0076 * RHmean + 0.0493 * Tmean + 0.1767 * Rs - 0.8888

LM num: 38
ET_0 = 0.5212 * Ws - 0.0095 * RHmean - 0.0009 * Tmean + 0.2822 * Rs - 0.3013

LM num: 39
ET_0 = 0.6783 * Ws - 0.0206 * RHmean + 0.0855 * Tmean + 0.1554 * Rs - 0.7063

LM num: 40
ET_0 = 0.7511 * Ws + 0.0322 * RHmean - 0.0002 * Tmean + 0.2172 * Rs - 1.3302

LM num: 41
ET_0 = 0.9783 * Ws + 0.0433 * RHmean + 0.0008 * Tmean + 0.3786 * Rs - 4.775</td>
</tr>
</table>

```
| | | | | Tmean ≤ 26.288 : LM83
| | | | | Tmean >  26.288 : LM84
| | | | RHmean >  70.957 :
| | | | | RHmean ≤ 78.741 :
| | | | | | Ws ≤ 2.55 :
| | | | | | | Tmean ≤ 24.55 : LM85
| | | | | | | Tmean >  24.55 : LM86
| | | | | | Ws >  2.55 : LM87
| | | | | RHmean >  78.741 : LM88
| Tmean >  28.837 :
| | RHmean ≤ 71.067 :
| | | Ws ≤ 4.171 :
| | | | Rs ≤ 22.693 :
| | | | | RHmean ≤ 22 : LM89
| | | | | RHmean >  22 :
| | | | | | Ws ≤ 2.263 :
| | | | | | | Ws ≤ 1.389 :
| | | | | | | | Tmean ≤ 31.05 :
| | | | | | | | | Ws ≤ 0.58 : LM90
| | | | | | | | | Ws >  0.58 :
| | | | | | | | | | RHmean ≤ 64.179 : LM91
| | | | | | | | | | RHmean >  64.179 : LM92
| | | | | | | | Tmean >  31.05 :
| | | | | | | | | RHmean ≤ 30.05 : LM93
| | | | | | | | | RHmean >  30.05 :
| | | | | | | | | | Tmean ≤ 32.1 : LM94
| | | | | | | | | | Tmean >  32.1 : LM95
| | | | | | | Ws >  1.389 :
| | | | | | | | Tmean ≤ 31.65 : LM96
| | | | | | | | Tmean >  31.65 :
| | | | | | | | | RHmean ≤ 29.45 : LM97
| | | | | | | | | RHmean >  29.45 : LM98
| | | | | | Ws >  2.263 :
| | | | | | | RHmean ≤ 27.95 : LM99
| | | | | | | RHmean >  27.95 :
| | | | | | | | Tmean ≤ 33.15 : LM100
| | | | | | | | Tmean >  33.15 :
| | | | | | | | | RHmean ≤ 55.235 : LM101
| | | | | | | | | RHmean >  55.235 : LM102
| | | | Rs >  22.693 :
| | | | | Tmean ≤ 32.05 :
| | | | | | RHmean ≤ 62.829 : LM103
| | | | | | RHmean >  62.829 : LM104
| | | | | Tmean >  32.05 :
| | | | | | RHmean ≤ 27.75 : LM105
| | | | | | RHmean >  27.75 :
| | | | | | | Tmean ≤ 35.05 :
| | | | | | | | RHmean ≤ 64.141 :
| | | | | | | | | Ws ≤ 1.988 : LM106
| | | | | | | | | Ws >  1.988 :
| | | | | | | | | | Rs ≤ 24.46 : LM107
| | | | | | | | | | Rs >  24.46 : LM108
| | | | | | | | RHmean >  64.141 : LM109
| | | | | | | Tmean >  35.05 :
| | | | | | | | Ws ≤ 2.398 :
| | | | | | | | | Rs ≤ 25.646 :
```

LM num: 42
ET_0 = 0.6124 * Ws + 0.0428 * RHmean + 0.0002 * Tmean + 0.2783 * Rs - 2.2927

LM num: 43
ET_0 = 0.4624 * Ws + 0.0703 * RHmean - 0 * Tmean + 0.237 * Rs - 2.0476

LM num: 44
ET_0 = 0.7394 * Ws - 0.0112 * RHmean + 0.0922 * Tmean + 0.2297 * Rs - 2.3566

LM num: 45
ET_0 = 0.7531 * Ws + 0.0445 * RHmean + 0.0007 * Tmean + 0.4069 * Rs - 4.7047

LM num: 46
ET_0 = 0.6017 * Ws + 0.0852 * RHmean + 0.0001 * Tmean + 0.2535 * Rs - 2.7791

LM num: 47
ET_0 = 0.3706 * Ws - 0.0487 * RHmean + 0.0918 * Tmean + 0.1721 * Rs + 1.9976

LM num: 48
ET_0 = 0.5979 * Ws + 0.0523 * RHmean + 0.0006 * Tmean + 0.3101 * Rs - 3.2486

LM num: 49
ET_0 = 0.219 * Ws - 0.0398 * RHmean + 0.0967 * Tmean + 0.0755 * Rs + 2.3572

LM num: 50
ET_0 = 0.2434 * Ws - 0.0577 * RHmean + 0.1472 * Tmean + 0.066 * Rs + 2.4549

LM num: 51
ET_0 = -0.0261 * Ws - 0.0626 * RHmean + 0.1712 * Tmean + 0.0201 * Rs + 4.5152

LM num: 52
ET_0 = 0.013 * Ws - 0.0635 * RHmean + 0.303 * Tmean + 0.0201 * Rs + 0.1058

LM num: 53
ET_0 = 0.013 * Ws - 0.0541 * RHmean + 0.3177 * Tmean + 0.0201 * Rs - 1.0526

LM num: 54
ET_0 = 0.1058 * Ws - 0.0645 * RHmean + 0.1375 * Tmean + 0.1678 * Rs + 2.2579

LM num: 55
ET_0 = 0.0133 * Ws + 0.0123 * RHmean + 0.1286 * Tmean + 0.067 * Rs - 1.0262

```
| | | | | | | | | | RHmean ≤ 51.456 : LM110
| | | | | | | | | | RHmean >  51.456 : LM111
| | | | | | | | | Rs >  25.646 :
| | | | | | | | | | Ws ≤ 1.496 : LM112
| | | | | | | | | | Ws >  1.496 :
| | | | | | | | | | | RHmean ≤ 52.578 :
| | | | | | | | | | | | RHmean ≤ 31.3 :
| | | | | | | | | | | | | Rs ≤ 28.414 : LM113
| | | | | | | | | | | | | Rs >  28.414 : LM114
| | | | | | | | | | | | RHmean >  31.3 : LM115
| | | | | | | | | | | RHmean >  52.578 :
| | | | | | | | | | | | Tmean ≤ 36.85 : LM116
| | | | | | | | | | | | Tmean >  36.85 : LM117
| | | | | | | | Ws >  2.398 :
| | | | | | | | | Rs ≤ 27.227 :
| | | | | | | | | | RHmean ≤ 30.6 : LM118
| | | | | | | | | | RHmean >  30.6 :
| | | | | | | | | | | Tmean ≤ 36.45 :
| | | | | | | | | | | | RHmean ≤ 64.476 : LM119
| | | | | | | | | | | | RHmean >  64.476 : LM120
| | | | | | | | | | | Tmean >  36.45 :
| | | | | | | | | | | | Rs ≤ 24.248 :
| | | | | | | | | | | | | Ws ≤ 2.974 :
| | | | | | | | | | | | | | Rs ≤ 23.444 : LM121
| | | | | | | | | | | | | | Rs >  23.444 : LM122
| | | | | | | | | | | | | Ws >  2.974 :
| | | | | | | | | | | | | | Tmean ≤ 38.05 :
| | | | | | | | | | | | | | | RHmean ≤ 61.445 : LM123
| | | | | | | | | | | | | | | RHmean >  61.445 : LM124
| | | | | | | | | | | | | | Tmean >  38.05 : LM125
| | | | | | | | | | | | Rs >  24.248 :
| | | | | | | | | | | | | Ws ≤ 3.266 :
| | | | | | | | | | | | | | Tmean ≤ 37.55 : LM126
| | | | | | | | | | | | | | Tmean >  37.55 : LM127
| | | | | | | | | | | | | Ws >  3.266 :
| | | | | | | | | | | | | | Tmean ≤ 37.75 : LM128
| | | | | | | | | | | | | | Tmean >  37.75 : LM129
| | | | | | | | | Rs >  27.227 :
| | | | | | | | | | RHmean ≤ 62.978 :
| | | | | | | | | | | RHmean ≤ 53.18 : LM130
| | | | | | | | | | | RHmean >  53.18 : LM131
| | | | | | | | | | RHmean >  62.978 : LM132
| | | Ws >  4.171 :
| | | | Tmean ≤ 33.35 : LM133
| | | | Tmean >  33.35 :
| | | | | Rs ≤ 23.814 :
| | | | | | RHmean ≤ 27.25 : LM134
| | | | | | RHmean >  27.25 :
| | | | | | | Ws ≤ 5.415 :
| | | | | | | | Rs ≤ 19.656 :
| | | | | | | | | RHmean ≤ 30.05 :
| | | | | | | | | | Tmean ≤ 1012.6 : LM135
| | | | | | | | | | Tmean >  1012.6 : LM136
| | | | | | | | | RHmean >  30.05 :
| | | | | | | | | | Tmean ≤ 37 :
| | | | | | | | | | | RHmean ≤ 66.567 : LM137
```

LM num: 56
ET_0 = -0.113 * Ws - 0.0373 * RHmean + 0.065 * Tmean + 0.0107 * Rs + 4.1903

LM num: 57
ET_0 = 0.1103 * Ws + 0.0218 * RHmean + 0.1029 * Tmean + 0.1756 * Rs - 3.3612

LM num: 58
ET_0 = 0.102 * Ws + 0.0146 * RHmean + 0.1241 * Tmean + 0.0225 * Rs - 0.0746

LM num: 59
ET_0 = -0.11 * Ws + 0.0081 * RHmean + 0.0363 * Tmean + 0.1703 * Rs - 1.4371

LM num: 60
ET_0 = 0.0825 * Ws - 0.019 * RHmean + 0.0363 * Tmean + 0.0276 * Rs + 3.4704

LM num: 61
ET_0 = 0.12 * Ws - 0.0665 * RHmean + 0.0881 * Tmean - 0.0322 * Rs + 6.9019

LM num: 62
ET_0 = 0.1382 * Ws - 0.0021 * RHmean + 0.057 * Tmean + 0.1808 * Rs - 1.3575

LM num: 63
ET_0 = 0.1033 * Ws - 0.008 * RHmean + 0.0196 * Tmean + 0.1781 * Rs + 0.0866

LM num: 64
ET_0 = 0.0491 * Ws - 0.0792 * RHmean + 0.1192 * Tmean + 0.062 * Rs + 6.1645

LM num: 65
ET_0 = 0.5904 * Ws - 0.0106 * RHmean + 0.121 * Tmean + 0.1415 * Rs - 2.0452

LM num: 66
ET_0 = 0.2287 * Ws - 0.0106 * RHmean + 0.1932 * Tmean + 0.2354 * Rs - 4.1627

LM num: 67
ET_0 = 0.1294 * Ws - 0.009 * RHmean + 0.0268 * Tmean + 0.172 * Rs + 0.1761

LM num: 68
ET_0 = 0.0467 * Ws - 0.0746 * RHmean + 0.0351 * Tmean + 0.0581 * Rs + 7.6209

LM num: 69
ET_0 = 0.1472 * Ws + 0.0099 * RHmean + 0.0882 * Tmean + 0.0532 * Rs + 0.2453

```
| | | | | | | | | | | RHmean > 66.567 : LM138
| | | | | | | | | | Tmean > 37 : LM139
| | | | | | | | Rs > 19.656 :
| | | | | | | | | RHmean ≤ 65.868 : LM140
| | | | | | | | | RHmean > 65.868 : LM141
| | | | | | | Ws > 5.415 :
| | | | | | | | RHmean ≤ 61.8 : LM142
| | | | | | | | RHmean > 61.8 : LM143
| | | | | Rs > 23.814 :
| | | | | | Tmean ≤ 35.95 :
| | | | | | | RHmean ≤ 65.888 : LM144
| | | | | | | RHmean > 65.888 : LM145
| | | | | | Tmean > 35.95 :
| | | | | | | RHmean ≤ 54.875 :
| | | | | | | | RHmean ≤ 27.35 : LM146
| | | | | | | | RHmean > 27.35 :
| | | | | | | | | Ws ≤ 5.9 :
| | | | | | | | | | Rs ≤ 29.123 : LM147
| | | | | | | | | | Rs > 29.123 : LM148
| | | | | | | | | Ws > 5.9 : LM149
| | | | | | | RHmean > 54.875 : LM150
| | RHmean > 71.067 :
| | | Rs ≤ 20.615 :
| | | | Rs ≤ 17.99 :
| | | | | Ws ≤ 4.029 :
| | | | | | Ws ≤ 2.263 : LM151
| | | | | | Ws > 2.263 : LM152
| | | | | Ws > 4.029 :
| | | | | | RHmean ≤ 81.631 :
| | | | | | | RHmean ≤ 76.841 : LM153
| | | | | | | RHmean > 76.841 : LM154
| | | | | | RHmean > 81.631 : LM155
| | | | Rs > 17.99 :
| | | | | Ws ≤ 3.343 : LM156
| | | | | Ws > 3.343 : LM157
| | | Rs > 20.615 :
| | | | Rs ≤ 23.848 :
| | | | | Ws ≤ 2.668 :
| | | | | | Tmean ≤ 31.95 :
| | | | | | | Tmean ≤ 29.85 :
| | | | | | | | Ws ≤ 2.075 : LM158
| | | | | | | | Ws > 2.075 : LM159
| | | | | | | Tmean > 29.85 :
| | | | | | | | Rs ≤ 22.335 :
| | | | | | | | | RHmean ≤ 79.56 :
| | | | | | | | | | Ws ≤ 1.653 : LM160
| | | | | | | | | | Ws > 1.653 : LM161
| | | | | | | | | RHmean > 79.56 : LM162
| | | | | | | | Rs > 22.335 : LM163
| | | | | | Tmean > 31.95 :
| | | | | | | RHmean ≤ 74.167 : LM164
| | | | | | | RHmean > 74.167 : LM165
| | | | | Ws > 2.668 :
| | | | | | RHmean ≤ 80.022 :
| | | | | | | Ws ≤ 3.926 : LM166
| | | | | | | Ws > 3.926 :
```

LM num: 70
ET_0 = 0.2585 * Ws - 0.0001 * RHmean + 0.023 * Tmean + 0.026 * Rs + 2.6351

LM num: 71
ET_0 = 0.1391 * Ws - 0.0027 * RHmean + 0.1912 * Tmean + 0.0048 * Rs + 0.895

LM num: 72
ET_0 = 0.04 * Ws - 0.0027 * RHmean + 0.0701 * Tmean - 0.0126 * Rs + 3.6056

LM num: 73
ET_0 = 0.0956 * Ws - 0.0372 * RHmean + 0.1832 * Tmean + 0.0126 * Rs + 2.659

LM num: 74
ET_0 = 0.0149 * Ws - 0.0911 * RHmean + 0.1625 * Tmean - 0.0586 * Rs + 8.7415

LM num: 75
ET_0 = 0.2068 * Ws + 0.0058 * RHmean + 0.0754 * Tmean + 0.1604 * Rs - 2.0667

LM num: 76
ET_0 = 0.2567 * Ws - 0.0337 * RHmean + 0.0652 * Tmean + 0.1668 * Rs + 1.0009

LM num: 77
ET_0 = 0.0831 * Ws - 0.0052 * RHmean + 0.0252 * Tmean + 0.074 * Rs + 1.7807

LM num: 78
ET_0 = 0.0733 * Ws - 0.0052 * RHmean + 0.0221 * Tmean + 0.0601 * Rs + 2.3334

LM num: 79
ET_0 = 0.0554 * Ws - 0.0098 * RHmean + 0.0257 * Tmean + 0.206 * Rs - 0.0295

LM num: 80
ET_0 = 0.0554 * Ws - 0.0986 * RHmean + 0.0406 * Tmean + 0.0758 * Rs + 9.3159

LM num: 81
ET_0 = 0.3841 * Ws - 0.0397 * RHmean + 0.117 * Tmean + 0.1742 * Rs - 0.129

LM num: 82
ET_0 = 0.2215 * Ws - 0.0565 * RHmean + 0.168 * Tmean + 0.1601 * Rs + 0.6541

LM num: 83
ET_0 = 0.1957 * Ws - 0.0409 * RHmean + 0.1334 * Tmean + 0.063 * Rs + 2.6507

```
|  |  |  |  |  |  |  |  RHmean ≤ 73.779 : LM167
|  |  |  |  |  |  |  |  RHmean >  73.779 :
|  |  |  |  |  |  |  |  |  Ws ≤ 5.693 : LM168
|  |  |  |  |  |  |  |  |  Ws >  5.693 :
|  |  |  |  |  |  |  |  |  |  RHmean ≤ 74.971 : LM169
|  |  |  |  |  |  |  |  |  |  RHmean >  74.971 : LM170
|  |  |  |  |  |  |  RHmean >  80.022 : LM171
|  |  |  |  Rs >  23.848 : LM172
```

LM num: 84
ET_0 = 0.3402 * Ws - 0.0628 * RHmean + 0.1113 * Tmean + 0.0844 * Rs + 3.9393

LM num: 85
ET_0 = 0.0204 * Ws - 0.0565 * RHmean + 0.1196 * Tmean + 0.0084 * Rs + 5.5466

LM num: 86
ET_0 = 0.1251 * Ws - 0.0382 * RHmean + 0.013 * Tmean + 0.0559 * Rs + 5.7266

LM num: 87
ET_0 = 0.0847 * Ws - 0.0805 * RHmean + 0.1937 * Tmean + 0.0054 * Rs + 5.6192

LM num: 88
ET_0 = 0.258 * Ws - 0.1163 * RHmean + 0.1343 * Tmean + 0.0047 * Rs + 9.4738

LM num: 89
ET_0 = 0.5067 * Ws + 0.0948 * RHmean - 0.017 * Tmean + 0.1804 * Rs + 15.5606

LM num: 90
ET_0 = 1.0375 * Ws - 0.0361 * RHmean + 0.0544 * Tmean + 0.0958 * Rs + 2.5772

LM num: 91
ET_0 = 0.8072 * Ws - 0.0216 * RHmean + 0.065 * Tmean + 0.1451 * Rs + 0.6261

LM num: 92
ET_0 = 0.5511 * Ws - 0.0271 * RHmean + 0.0684 * Tmean + 0.1631 * Rs + 0.5841

LM num: 93
ET_0 = 0.876 * Ws + 0.0328 * RHmean - 0.0004 * Tmean + 0.2212 * Rs - 1.0178

LM num: 94
ET_0 = 0.8708 * Ws - 0.0033 * RHmean - 0.0004 * Tmean + 0.2095 * Rs + 0.5698

LM num: 95
ET_0 = 0.9232 * Ws - 0.013 * RHmean - 0.0009 * Tmean + 0.1473 * Rs + 2.6076

LM num: 96
ET_0 = 0.6604 * Ws - 0.0644 * RHmean + 0.1063 * Tmean + 0.1957 * Rs + 1.1989

LM num: 97
ET_0 = 0.891 * Ws + 0.0586 * RHmean - 0.0005 * Tmean + 0.2091 * Rs - 1.2995

	LM num: 98 $ET_0 = 0.9545 * Ws - 0 * RHmean - 0.0005 * Tmean + 0.1792 * Rs + 1.1961$ LM num: 99 $ET_0 = 0.6077 * Ws + 0.1068 * RHmean - 0.0247 * Tmean + 0.2308 * Rs + 22.145$ LM num: 100 $ET_0 = 0.5196 * Ws - 0.0834 * RHmean + 0.1717 * Tmean + 0.2024 * Rs + 0.5766$ LM num: 101 $ET_0 = 0.6241 * Ws + 0.018 * RHmean - 0.0003 * Tmean + 0.1982 * Rs + 0.8349$ LM num: 102 $ET_0 = 0.6998 * Ws - 0.0353 * RHmean + 0.1858 * Tmean + 0.243 * Rs - 4.0606$ LM num: 103 $ET_0 = 0.5068 * Ws - 0.0292 * RHmean + 0.1963 * Tmean + 0.1273 * Rs - 1.5246$ LM num: 104 $ET_0 = 0.3723 * Ws - 0.1054 * RHmean + 0.2285 * Tmean + 0.059 * Rs + 4.1711$ LM num: 105 $ET_0 = 0.6272 * Ws + 0.1496 * RHmean - 0.0002 * Tmean + 0.245 * Rs - 4.2355$ LM num: 106 $ET_0 = 0.6572 * Ws - 0.0159 * RHmean + 0.1886 * Tmean + 0.2318 * Rs - 4.895$ LM num: 107 $ET_0 = 0.6852 * Ws - 0.0511 * RHmean + 0.2737 * Tmean + 0.0593 * Rs - 1.6114$ LM num: 108 $ET_0 = 0.2844 * Ws - 0.064 * RHmean + 0.2575 * Tmean + 0.0748 * Rs + 0.8417$ LM num: 109 $ET_0 = 0.4477 * Ws - 0.1174 * RHmean + 0.1342 * Tmean + 0.0892 * Rs + 7.2201$ LM num: 110 $ET_0 = 0.919 * Ws + 0.0151 * RHmean - 0.0007 * Tmean + 0.2179 * Rs + 0.2842$ LM num: 111 $ET_0 = 0.7666 * Ws - 0.0206 * RHmean + 0.1461 * Tmean + 0.1984 * Rs - 2.3622$

	LM num: 112 ET_0 = 0.7622 * Ws - 0.0124 * RHmean - 0.0014 * Tmean + 0.1568 * Rs + 3.9728 LM num: 113 ET_0 = 0.4364 * Ws + 0.0499 * RHmean - 0.0015 * Tmean + 0.1601 * Rs + 2.7322 LM num: 114 ET_0 = 0.4364 * Ws + 0.0233 * RHmean - 0.0015 * Tmean + 0.179 * Rs + 3.0575 LM num: 115 ET_0 = 0.5333 * Ws - 0.0266 * RHmean - 0.0016 * Tmean + 0.1976 * Rs + 4.0839 LM num: 116 ET_0 = 0.6503 * Ws - 0.0744 * RHmean + 0.1572 * Tmean + 0.1431 * Rs + 2.131 LM num: 117 ET_0 = 0.726 * Ws - 0.0355 * RHmean + 0.1324 * Tmean + 0.1809 * Rs - 0.2315 LM num: 118 ET_0 = 0.7315 * Ws - 0.0197 * RHmean - 0.0505 * Tmean + 0.2936 * Rs + 50.0822 LM num: 119 ET_0 = 0.6211 * Ws - 0.0602 * RHmean + 0.1865 * Tmean + 0.2142 * Rs - 1.4676 LM num: 120 ET_0 = 0.3365 * Ws - 0.1772 * RHmean + 0.0835 * Tmean + 0.1837 * Rs + 11.5674 LM num: 121 ET_0 = 0.4137 * Ws - 0.0213 * RHmean - 0.001 * Tmean + 0.304 * Rs + 1.6544 LM num: 122 ET_0 = 0.4137 * Ws - 0.0206 * RHmean - 0.001 * Tmean + 0.2882 * Rs + 2.0119 LM num: 123 ET_0 = 0.3953 * Ws - 0.0249 * RHmean + 0.041 * Tmean + 0.2461 * Rs + 1.8676 LM num: 124 ET_0 = 0.3953 * Ws - 0.0269 * RHmean + 0.041 * Tmean + 0.2356 * Rs + 2.2118 LM num: 125 ET_0 = 0.3953 * Ws - 0.0142 * RHmean + 0.055 * Tmean + 0.212 * Rs + 1.5143

	LM num: 126 ET_0 = 0.6104 * Ws - 0.0176 * RHmean - 0.0011 * Tmean + 0.2472 * Rs + 2.2277 LM num: 127 ET_0 = 0.5014 * Ws - 0.008 * RHmean - 0.001 * Tmean + 0.1791 * Rs + 3.8758 LM num: 128 ET_0 = 0.4234 * Ws - 0.0314 * RHmean - 0.0009 * Tmean + 0.141 * Rs + 6.4235 LM num: 129 ET_0 = 0.2873 * Ws - 0.0176 * RHmean - 0.0011 * Tmean + 0.1635 * Rs + 5.6695 LM num: 130 ET_0 = 0.791 * Ws - 0.007 * RHmean - 0.001 * Tmean + 0.252 * Rs + 0.8494 LM num: 131 ET_0 = 0.5896 * Ws - 0.0522 * RHmean + 0.1433 * Tmean + 0.1189 * Rs + 2.5565 LM num: 132 ET_0 = 0.2251 * Ws - 0.2206 * RHmean + 0.2509 * Tmean + 0.0675 * Rs + 11.9312 LM num: 133 ET_0 = 0.28 * Ws - 0.132 * RHmean + 0.2253 * Tmean + 0.1073 * Rs + 5.2344 LM num: 134 ET_0 = 0.4782 * Ws + 0.1608 * RHmean - 0.0004 * Tmean + 0.2276 * Rs - 3.2895 LM num: 135 ET_0 = 0.4884 * Ws + 0.0527 * RHmean - 0.0332 * Tmean + 0.0554 * Rs + 36.5765 LM num: 136 ET_0 = 0.3571 * Ws + 0.0835 * RHmean - 0.0504 * Tmean + 0.0554 * Rs + 53.5042 LM num: 137 ET_0 = 0.4414 * Ws - 0.045 * RHmean + 0.0489 * Tmean + 0.0554 * Rs + 6.418 LM num: 138 ET_0 = 0.4414 * Ws - 0.0485 * RHmean + 0.0723 * Tmean + 0.0554 * Rs + 5.7604 LM num: 139 ET_0 = 0.6744 * Ws - 0.0197 * RHmean - 0.001 * Tmean + 0.0554 * Rs + 5.6781

	LM num: 140 ET_0 = 0.5315 * Ws - 0.0195 * RHmean - 0.0012 * Tmean + 0.1101 * Rs + 5.5081 LM num: 141 ET_0 = 0.334 * Ws - 0.0348 * RHmean + 0.1638 * Tmean + 0.049 * Rs + 2.4769 LM num: 142 ET_0 = 0.4318 * Ws - 0.0433 * RHmean - 0.0017 * Tmean + 0.194 * Rs + 6.097 LM num: 143 ET_0 = 0.3588 * Ws - 0.0898 * RHmean + 0.1942 * Tmean + 0.1902 * Rs + 2.1411 LM num: 144 ET_0 = 0.5453 * Ws - 0.1068 * RHmean + 0.3348 * Tmean + 0.1179 * Rs - 1.072 LM num: 145 ET_0 = 0.5404 * Ws - 0.05 * RHmean + 0.0721 * Tmean + 0.0446 * Rs + 5.5416 LM num: 146 ET_0 = 0.7004 * Ws + 0.171 * RHmean - 0.0019 * Tmean + 0.218 * Rs - 2.6384 LM num: 147 ET_0 = 0.5496 * Ws - 0.021 * RHmean - 0.0021 * Tmean + 0.1798 * Rs + 4.8873 LM num: 148 ET_0 = 0.6089 * Ws - 0.0312 * RHmean - 0.0019 * Tmean + 0.1864 * Rs + 5.0853 LM num: 149 ET_0 = 0.6014 * Ws - 0.0207 * RHmean + 0.0932 * Tmean + 0.1956 * Rs + 0.6233 LM num: 150 ET_0 = 0.4919 * Ws - 0.0899 * RHmean + 0.2203 * Tmean + 0.1914 * Rs + 0.3419 LM num: 151 ET_0 = 0.276 * Ws + 0.0028 * RHmean + 0.0645 * Tmean + 0.2458 * Rs - 2.759 LM num: 152 ET_0 = 0.1123 * Ws - 0.0461 * RHmean + 0.1522 * Tmean + 0.1697 * Rs + 0.1263 LM num: 153 ET_0 = 0.1741 * Ws - 0.0963 * RHmean + 0.236 * Tmean + 0.1166 * Rs + 2.2372

	LM num: 154 ET_0 = 0.1133 * Ws - 0.0358 * RHmean + 0.0953 * Tmean + 0.224 * Rs + 0.4526 LM num: 155 ET_0 = 0.129 * Ws - 0.0867 * RHmean + 0.0977 * Tmean + 0.1845 * Rs + 5.0296 LM num: 156 ET_0 = 0.2182 * Ws - 0.0318 * RHmean + 0.1415 * Tmean + 0.1871 * Rs - 1.2798 LM num: 157 ET_0 = 0.1192 * Ws - 0.0797 * RHmean + 0.1243 * Tmean + 0.1793 * Rs + 3.6845 LM num: 158 ET_0 = 0.1582 * Ws - 0.0411 * RHmean + 0.0575 * Tmean + 0.0421 * Rs + 4.9139 LM num: 159 ET_0 = 0.1947 * Ws - 0.0229 * RHmean + 0.0575 * Tmean + 0.0906 * Rs + 2.6421 LM num: 160 ET_0 = 0.2179 * Ws - 0.0527 * RHmean + 0.1239 * Tmean + 0.0636 * Rs + 3.4265 LM num: 161 ET_0 = 0.1829 * Ws - 0.0319 * RHmean + 0.101 * Tmean + 0.0636 * Rs + 2.7614 LM num: 162 ET_0 = 0.1936 * Ws - 0.0911 * RHmean + 0.0998 * Tmean + 0.0636 * Rs + 7.3623 LM num: 163 ET_0 = 0.2265 * Ws - 0.0228 * RHmean + 0.1132 * Tmean + 0.198 * Rs - 1.3752 LM num: 164 ET_0 = 0.3288 * Ws - 0.0201 * RHmean + 0.2536 * Tmean + 0.06 * Rs - 3.2182 LM num: 165 ET_0 = 0.1895 * Ws - 0.0316 * RHmean + 0.1148 * Tmean + 0.1564 * Rs + 0.344 LM num: 166 ET_0 = 0.025 * Ws - 0.0633 * RHmean + 0.117 * Tmean + 0.1582 * Rs + 3.3244 LM num: 167 ET_0 = 0.0489 * Ws - 0.0293 * RHmean + 0.1146 * Tmean + 0.0348 * Rs + 4.0607

	LM num: 168 ET_0 = 0.0448 * Ws - 0.0458 * RHmean + 0.106 * Tmean + 0.1142 * Rs + 3.4281 LM num: 169 ET_0 = 0.1742 * Ws - 0.0252 * RHmean + 0.2516 * Tmean + 0.1498 * Rs - 4.2065 LM num: 170 ET_0 = 0.1226 * Ws - 0.0252 * RHmean + 0.0888 * Tmean + 0.1029 * Rs + 2.3879 LM num: 171 ET_0 = 0.0986 * Ws - 0.0648 * RHmean + 0.1221 * Tmean + 0.1429 * Rs + 3.3768 LM num: 172 ET_0 = 0.2209 * Ws - 0.0449 * RHmean + 0.1599 * Tmean + 0.1714 * Rs - 0.5402

Table 10. M5MT2 model tree and corresponding linear models.

```
Rs ≤ 16.271 :
|  Tmean ≤ 19.175 :
|  |  Tmean ≤ 11.05 :
|  |  |  Tmean ≤ 8.55 : LM1
|  |  |  Tmean >  8.55 :
|  |  |  |  Rs ≤ 12.768 : LM2
|  |  |  |  Rs >  12.768 : LM3
|  |  Tmean >  11.05 : LM4
|  Tmean >  19.175 :
|  |  Rs ≤ 12.217 :
|  |  |  Tmean ≤ 1014.05 :
|  |  |  |  Tmean ≤ 24.15 : LM5
|  |  |  |  Tmean >  24.15 :
|  |  |  |  |  Tmean ≤ 517.85 : LM6
|  |  |  |  |  Tmean >  517.85 : LM7
|  |  |  Tmean >  1014.05 :
|  |  |  |  Rs ≤ 9.683 : LM8
|  |  |  |  Rs >  9.683 : LM9
|  |  Rs >  12.217 :
|  |  |  Tmean ≤ 29.55 : LM10
|  |  |  Tmean >  29.55 :
|  |  |  |  Tmean ≤ 1011.65 :
|  |  |  |  |  Tmean ≤ 34.35 :
|  |  |  |  |  |  Rs ≤ 14.749 : LM11
|  |  |  |  |  |  Rs >  14.749 : LM12
|  |  |  |  |  Tmean >  34.35 :
|  |  |  |  |  |  Tmean ≤ 1000.2 :
|  |  |  |  |  |  |  Tmean ≤ 35.75 : LM13
|  |  |  |  |  |  |  Tmean >  35.75 : LM14
|  |  |  |  |  |  Tmean >  1000.2 : LM15
|  |  |  |  Tmean >  1011.65 : LM16
Rs > 16.271 :
|  Tmean ≤ 28.837 :
|  |  Tmean ≤ 21.15 :
|  |  |  Tmean ≤ 15.05 : LM17
|  |  |  Tmean >  15.05 :
|  |  |  |  Rs ≤ 19.827 : LM18
|  |  |  |  Rs >  19.827 :
|  |  |  |  |  Tmean ≤ 18.35 : LM19
|  |  |  |  |  Tmean >  18.35 :
|  |  |  |  |  |  Rs ≤ 23.345 : LM20
|  |  |  |  |  |  Rs >  23.345 : LM21
|  |  Tmean >  21.15 :
|  |  |  Rs ≤ 21.308 : LM22
|  |  |  Rs >  21.308 : LM23
|  Tmean >  28.837 :
|  |  Rs ≤ 23.821 :
|  |  |  Tmean ≤ 35.05 :
|  |  |  |  Rs ≤ 20.573 : LM24
|  |  |  |  Rs >  20.573 : LM25
|  |  |  Tmean >  35.05 :
|  |  |  |  Tmean ≤ 1010.65 :
|  |  |  |  |  Tmean ≤ 998.4 :
|  |  |  |  |  |  Tmean ≤ 36.25 : LM26
```

LM num: 1
ET_0 = 0.055 * Tmean + 0.0452 * Rs + 0.2615

LM num: 2
ET_0 = 0.1191 * Tmean + 0.0465 * Rs - 0.3114

LM num: 3
ET_0 = 0.1244 * Tmean + 0.1536 * Rs - 1.614

LM num: 4
ET_0 = 0.0898 * Tmean + 0.0948 * Rs - 0.3844

LM num: 5
ET_0 = 0.0907 * Tmean + 0.0095 * Rs + 0.5308

LM num: 6
ET_0 = 0.1668 * Tmean - 0.0826 * Rs - 0.102

LM num: 7
ET_0 = -0.0286 * Tmean + 0.25 * Rs + 28.9016

LM num: 8
ET_0 = -0.0245 * Tmean + 0.1368 * Rs + 25.4375

LM num: 9
ET_0 = -0.0545 * Tmean + 0.3132 * Rs + 54.5032

LM num: 10
ET_0 = 0.1099 * Tmean + 0.1697 * Rs - 1.8084

LM num: 11
ET_0 = 0.261 * Tmean + 0.0237 * Rs - 4.6786

LM num: 12
ET_0 = 0.0068 * Tmean + 0.2225 * Rs + 0.698

LM num: 13
ET_0 = 0.2878 * Tmean - 0.0926 * Rs - 3.5999

LM num: 14
ET_0 = -0.0001 * Tmean + 0.2954 * Rs + 1.6971

LM num: 15
ET_0 = -0.0001 * Tmean + 0.358 * Rs - 0.6375

LM num: 16
ET_0 = -0.0709 * Tmean + 0.3143 * Rs + 71.2821

LM num: 17
ET_0 = 0.1549 * Tmean + 0.0758 * Rs - 0.6739

LM num: 18
ET_0 = 0.1129 * Tmean + 0.1528 * Rs - 1.5611

<table>
<tr>
<td>| | | | | | Tmean > 36.25 : LM27
| | | | | Tmean > 998.4 : LM28
| | | | Tmean > 1010.65 : LM29
| | Rs > 23.821 :
| | | Tmean ≤ 33.362 : LM30
| | | Tmean > 33.362 :
| | | | Tmean ≤ 996.5 :
| | | | | Tmean ≤ 35.85 :
| | | | | | Rs ≤ 25.269 : LM31
| | | | | | Rs > 25.269 : LM32
| | | | | Tmean > 35.85 :
| | | | | | Rs ≤ 26.667 : LM33
| | | | | | Rs > 26.667 :
| | | | | | | Tmean ≤ 36.65 : LM34
| | | | | | | Tmean > 36.65 : LM35
| | | | Tmean > 996.5 :
| | | | | Tmean ≤ 1012.65 : LM36
| | | | | Tmean > 1012.65 : LM37</td>
<td>LM num: 19
ET_0 = 0.0816 * Tmean + 0.007 * Rs + 2.0797

LM num: 20
ET_0 = 0.017 * Tmean + 0.0118 * Rs + 3.3481

LM num: 21
ET_0 = 0.0177 * Tmean - 0.06 * Rs + 5.7062

LM num: 22
ET_0 = 0.1283 * Tmean + 0.1361 * Rs - 1.5961

LM num: 23
ET_0 = 0.1468 * Tmean + 0.0942 * Rs - 0.9974

LM num: 24
ET_0 = 0.0936 * Tmean + 0.2272 * Rs - 2.175

LM num: 25
ET_0 = 0.2001 * Tmean + 0.3041 * Rs - 7.1203

LM num: 26
ET_0 = -0 * Tmean + 0.4277 * Rs - 2.3509

LM num: 27
ET_0 = -0 * Tmean + 0.2871 * Rs + 2.3509

LM num: 28
ET_0 = -0 * Tmean + 0.2247 * Rs + 1.9173

LM num: 29
ET_0 = -0.1221 * Tmean + 0.1748 * Rs + 125.6793

LM num: 30
ET_0 = 0.3013 * Tmean + 0.1504 * Rs - 6.4609

LM num: 31
ET_0 = 0.2893 * Tmean + 0.033 * Rs - 3.3114

LM num: 32
ET_0 = 0.5463 * Tmean + 0.1636 * Rs - 14.7396

LM num: 33
ET_0 = 0.3421 * Tmean + 0.2588 * Rs - 9.9582

LM num: 34
ET_0 = -0 * Tmean + 0.0182 * Rs + 9.3817

LM num: 35
ET_0 = -0 * Tmean + 0.0182 * Rs + 10.0524

LM num: 36
ET_0 = -0.0407 * Tmean + 0.2123 * Rs + 43.2007

LM num: 37
ET_0 = -0.1143 * Tmean + 0.3193 * Rs + 114.1612</td>
</tr>
</table>

Table 11. M5MT3 model tree and corresponding linear models

<table>
<tr>
<td>
Tmean ≤ 23.85 :
| Tmean ≤ 13.85 :
| | Tmean ≤ 10.25 :
| | | RHmean ≤ 83.795 : LM1
| | | RHmean > 83.795 : LM2
| | Tmean > 10.25 : LM3
| Tmean > 13.85 :
| | Tmean ≤ 19.488 :
| | | RHmean ≤ 74.667 :
| | | | RHmean ≤ 70.126 :
| | | | | Tmean ≤ 18.35 : LM4
| | | | | Tmean > 18.35 : LM5
| | | | RHmean > 70.126 : LM6
| | | RHmean > 74.667 : LM7
| | Tmean > 19.488 :
| | | RHmean ≤ 72.906 : LM8
| | | RHmean > 72.906 : LM9
Tmean > 23.85 :
| RHmean ≤ 69.416 :
| | RHmean ≤ 18.65 :
| | | RHmean ≤ 11.55 :
| | | | RHmean ≤ 7.75 : LM10
| | | | RHmean > 7.75 : LM11
| | | RHmean > 11.55 : LM12
| | RHmean > 18.65 :
| | | Tmean ≤ 31.85 :
| | | | Tmean ≤ 28.15 : LM13
| | | | Tmean > 28.15 :
| | | | | RHmean ≤ 63.534 :
| | | | | | RHmean ≤ 58.207 : LM14
| | | | | | RHmean > 58.207 : LM15
| | | | | RHmean > 63.534 : LM16
| | | Tmean > 31.85 :
| | | | RHmean ≤ 30.15 : LM17
| | | | RHmean > 30.15 :
| | | | | Tmean ≤ 34.55 : LM18
| | | | | Tmean > 34.55 :
| | | | | | RHmean ≤ 60.804 :
| | | | | | | RHmean ≤ 58.384 :
| | | | | | | | RHmean ≤ 53.392 :
| | | | | | | | | RHmean ≤ 31.35 : LM19
| | | | | | | | | RHmean > 31.35 :
| | | | | | | | | | Tmean ≤ 36.85 : LM20
| | | | | | | | | | Tmean > 36.85 : LM21
| | | | | | | | RHmean > 53.392 : LM22
| | | | | | | RHmean > 58.384 : LM23
| | | | | | RHmean > 60.804 : LM24
| RHmean > 69.416 :
| | Tmean ≤ 29.45 : LM25
| | Tmean > 29.45 :
| | | RHmean ≤ 80.109 :
| | | | Tmean ≤ 31.65 : LM26
| | | | Tmean > 31.65 : LM27
| | | RHmean > 80.109 :
</td>
<td>
LM num: 1
ET_0 = 0.1051 * Tmean - 0.0117 * RHmean + 1.3417

LM num: 2
ET_0 = 0.0582 * Tmean - 0.0316 * RHmean + 3.3931

LM num: 3
ET_0 = 0.1463 * Tmean - 0.0296 * RHmean + 2.3315

LM num: 4
ET_0 = 0.1123 * Tmean - 0.0012 * RHmean + 1.1801

LM num: 5
ET_0 = 0.0105 * Tmean - 0.0469 * RHmean + 6.2457

LM num: 6
ET_0 = 0.1001 * Tmean - 0.0013 * RHmean + 1.0823

LM num: 7
ET_0 = 0.0876 * Tmean - 0.0484 * RHmean + 4.7497

LM num: 8
ET_0 = 0.184 * Tmean - 0.0374 * RHmean + 2.3942

LM num: 9
ET_0 = 0.1588 * Tmean - 0.0546 * RHmean + 3.9712

LM num: 10
ET_0 = -0.002 * Tmean + 0.1192 * RHmean + 2.8914

LM num: 11
ET_0 = -0.0393 * Tmean + 0.0066 * RHmean + 42.0496

LM num: 12
ET_0 = -0.0766 * Tmean + 0.1929 * RHmean + 78.2035

LM num: 13
ET_0 = 0.2032 * Tmean + 0.0003 * RHmean - 0.2284

LM num: 14
ET_0 = 0.0169 * Tmean + 0.006 * RHmean + 4.4895

LM num: 15
ET_0 = 0.0157 * Tmean + 0.0053 * RHmean + 5.187

LM num: 16
ET_0 = 0.3194 * Tmean - 0.1013 * RHmean + 3.2813

LM num: 17
ET_0 = -0.1276 * Tmean + 0.1691 * RHmean + 130.7925

LM num: 18
ET_0 = 0.5543 * Tmean + 0.0176 * RHmean - 12.079
</td>
</tr>
</table>

| | | | RHmean ≤ 84.591 : LM28 | | | | RHmean > 84.591 : | | | | | RHmean ≤ 87.71 : | | | | | | Tmean ≤ 34.35 : | | | | | | | Tmean ≤ 32.85 : | | | | | | | | Tmean ≤ 31.05 : LM29 | | | | | | | | Tmean > 31.05 : LM30 | | | | | | | Tmean > 32.85 : LM31 | | | | | | Tmean > 34.35 : LM32 | | | | | RHmean > 87.71 : | | | | | | Tmean ≤ 34.15 : LM33 | | | | | | Tmean > 34.15 : LM34	LM num: 19 ET_0 = -0.001 * Tmean - 0.0147 * RHmean + 8.9144 LM num: 20 ET_0 = -0.001 * Tmean - 0.0147 * RHmean + 10.0672 LM num: 21 ET_0 = -0.001 * Tmean + 0.0393 * RHmean + 8.3204 LM num: 22 ET_0 = 0.2932 * Tmean - 0.2086 * RHmean + 10.5241 LM num: 23 ET_0 = -0.0003 * Tmean + 0.2264 * RHmean - 4.4997 LM num: 24 ET_0 = -0.0002 * Tmean - 0.0386 * RHmean + 11.1191 LM num: 25 ET_0 = 0.1596 * Tmean - 0.0872 * RHmean + 6.7334 LM num: 26 ET_0 = 0.308 * Tmean - 0.1028 * RHmean + 3.7017 LM num: 27 ET_0 = 0.2144 * Tmean - 0.1418 * RHmean + 9.7024 LM num: 28 ET_0 = 0.1908 * Tmean - 0.1033 * RHmean + 7.3385 LM num: 29 ET_0 = -0.017 * Tmean - 0.0187 * RHmean + 6.8232 LM num: 30 ET_0 = -0.0013 * Tmean - 0.0187 * RHmean + 6.0764 LM num: 31 ET_0 = 0.012 * Tmean - 0.1332 * RHmean + 15.7479 LM num: 32 ET_0 = 0.1253 * Tmean - 0.1232 * RHmean + 11.1369 LM num: 33 ET_0 = -0.0308 * Tmean - 0.1851 * RHmean + 21.6473 LM num: 34 ET_0 = 0.2421 * Tmean - 0.0438 * RHmean - 0.0308

Table 12. M5MT4 model tree and corresponding linear models.

```
Tmean ≤ 23.85 :
|  Tmax ≤ 19.587 :
|  |  Ra ≤ 7.791 : LM1
|  |  Ra >  7.791 :
|  |  |  Ra ≤ 12.399 : LM2
|  |  |  Ra >  12.399 :
|  |  |  |  Tmax ≤ 14.15 : LM3
|  |  |  |  Tmax >  14.15 :
|  |  |  |  |  Tmin ≤ 9.3 : LM4
|  |  |  |  |  Tmin >  9.3 :
|  |  |  |  |  |  Tmin ≤ 12.5 : LM5
|  |  |  |  |  |  Tmin >  12.5 : LM6
|  Tmax >  19.587 :
|  |  Ra ≤ 11.661 :
|  |  |  Ra ≤ 8.569 :
|  |  |  |  Tmean ≤ 18.85 : LM7
|  |  |  |  Tmean >  18.85 : LM8
|  |  |  Ra >  8.569 : LM9
|  |  Ra >  11.661 :
|  |  |  Tmax ≤ 25.55 : LM10
|  |  |  Tmax >  25.55 :
|  |  |  |  Ra ≤ 13.463 : LM11
|  |  |  |  Ra >  13.463 :
|  |  |  |  |  Tmax ≤ 28.325 :
|  |  |  |  |  |  Tmin ≤ 15.1 :
|  |  |  |  |  |  |  Ra ≤ 16.043 : LM12
|  |  |  |  |  |  |  Ra >  16.043 : LM13
|  |  |  |  |  |  Tmin >  15.1 :
|  |  |  |  |  |  |  Ra ≤ 16.527 : LM14
|  |  |  |  |  |  |  Ra >  16.527 : LM15
|  |  |  |  |  Tmax >  28.325 :
|  |  |  |  |  |  Ra ≤ 15.671 : LM16
|  |  |  |  |  |  Ra >  15.671 : LM17
Tmean >  23.85 :
|  Tmax ≤ 39.938 :
|  |  Ra ≤ 11.752 :
|  |  |  Tmin ≤ 9.25 :
|  |  |  |  Ra ≤ 7.646 : LM18
|  |  |  |  Ra >  7.646 : LM19
|  |  |  Tmin >  9.25 :
|  |  |  |  Ra ≤ 9.879 :
|  |  |  |  |  Tmax ≤ 29.1 :
|  |  |  |  |  |  Tmean ≤ 26.4 :
|  |  |  |  |  |  |  Ra ≤ 9.265 : LM20
|  |  |  |  |  |  |  Ra >  9.265 : LM21
|  |  |  |  |  |  Tmean >  26.4 :
|  |  |  |  |  |  |  Tmean ≤ 1014.85 :
|  |  |  |  |  |  |  |  Ra ≤ 9.01 : LM22
|  |  |  |  |  |  |  |  Ra >  9.01 :
|  |  |  |  |  |  |  |  |  Tmean ≤ 1011.05 : LM23
|  |  |  |  |  |  |  |  |  Tmean >  1011.05 : LM24
|  |  |  |  |  |  |  Tmean >  1014.85 : LM25
|  |  |  |  |  Tmax >  29.1 :
|  |  |  |  |  |  Tmin ≤ 22.15 : LM26
```

LM num: 1
ET_0 = -0.0039 * Tmean + 0.0309 * Tmax + 0.0191 * Tmin + 0.2201 * Ra - 0.9325

LM num: 2
ET_0 = -0.0021 * Tmean + 0.0789 * Tmax + 0.0108 * Tmin + 0.1327 * Ra - 0.7615

LM num: 3
ET_0 = -0.0091 * Tmean + 0.1988 * Tmax - 0.0971 * Tmin + 0.1157 * Ra - 1.3236

LM num: 4
ET_0 = -0.0051 * Tmean + 0.112 * Tmax - 0.0014 * Tmin + 0.2485 * Ra - 2.4605

LM num: 5
ET_0 = -0.0051 * Tmean + 0.1719 * Tmax - 0.0007 * Tmin + 0.0169 * Ra - 0.4506

LM num: 6
ET_0 = -0.0051 * Tmean + 0.0289 * Tmax - 0.0007 * Tmin + 0.0169 * Ra + 2.1989

LM num: 7
ET_0 = -0.0005 * Tmean + 0.0862 * Tmax + 0.0005 * Tmin + 0.1569 * Ra - 1.244

LM num: 8
ET_0 = 0.1201 * Tmean + 0.0091 * Tmax + 0.0017 * Tmin - 0.3511 * Ra + 2.5801

LM num: 9
ET_0 = 0.2238 * Tmean + 0.0017 * Tmax - 0.1357 * Tmin + 0.2132 * Ra - 1.8172

LM num: 10
ET_0 = -0.0001 * Tmean + 0.1657 * Tmax - 0.056 * Tmin + 0.1742 * Ra - 2.0785

LM num: 11
ET_0 = -0.0001 * Tmean + 0.158 * Tmax - 0.0402 * Tmin + 0.261 * Ra - 3.1174

LM num: 12
ET_0 = -0.0151 * Tmean + 0.05 * Tmax + 0.0649 * Tmin + 0.0764 * Ra + 1.2339

LM num: 13
ET_0 = -0.0151 * Tmean + 0.1279 * Tmax - 0.0018 * Tmin + 0.5828 * Ra - 7.6865

<table>
<tr>
<td>| | | | | | Tmin > 22.15 : LM27
| | | | Ra > 9.879 :
| | | | | Tmax ≤ 34.55 :
| | | | | | Tmin ≤ 19.45 :
| | | | | | | Tmax ≤ 0.05 : LM28
| | | | | | | Tmax > 0.05 :
| | | | | | | | Tmax ≤ 33.35 :
| | | | | | | | | Ra ≤ 10.824 : LM29
| | | | | | | | | Ra > 10.824 : LM30
| | | | | | | | Tmax > 33.35 : LM31
| | | | | | Tmin > 19.45 : LM32
| | | | | Tmax > 34.55 : LM33
| | Ra > 11.752 :
| | | Tmax ≤ 28.35 :
| | | | Tmax ≤ 0.05 : LM34
| | | | Tmax > 0.05 :
| | | | | Ra ≤ 14.666 :
| | | | | | Tmin ≤ 2.25 : LM35
| | | | | | Tmin > 2.25 : LM36
| | | | | Ra > 14.666 :
| | | | | | Tmax ≤ 21.1 :
| | | | | | | Tmin ≤ 1.15 : LM37
| | | | | | | Tmin > 1.15 : LM38
| | | | | | Tmax > 21.1 : LM39
| | | Tmax > 28.35 :
| | | | Tmax ≤ 36.487 :
| | | | | Ra ≤ 14.109 :
| | | | | | Tmax ≤ 31.45 : LM40
| | | | | | Tmax > 31.45 :
| | | | | | | Tmax ≤ 35.45 : LM41
| | | | | | | Tmax > 35.45 :
| | | | | | | | Tmin ≤ 24.675 :
| | | | | | | | | Ra ≤ 12.339 : LM42
| | | | | | | | | Ra > 12.339 : LM43
| | | | | | | | Tmin > 24.675 : LM44
| | | | | Ra > 14.109 :
| | | | | | Tmax ≤ 32.15 : LM45
| | | | | | Tmax > 32.15 :
| | | | | | | Tmin ≤ 28.9 : LM46
| | | | | | | Tmin > 28.9 :
| | | | | | | | Ra ≤ 15.744 : LM47
| | | | | | | | Ra > 15.744 :
| | | | | | | | | Tmin ≤ 30.3 :
| | | | | | | | | | Tmax ≤ 35.3 :
| | | | | | | | | | | Tmin ≤ 29.9 : LM48
| | | | | | | | | | | Tmin > 29.9 :
| | | | | | | | | | | | Ra ≤ 16.289 : LM49
| | | | | | | | | | | | Ra > 16.289 : LM50
| | | | | | | | | | Tmax > 35.3 : LM51
| | | | | | | | | Tmin > 30.3 :
| | | | | | | | | | Ra ≤ 16.196 : LM52
| | | | | | | | | | Ra > 16.196 : LM53
| | | | Tmax > 36.487 :
| | | | | Tmin ≤ 29.35 :
| | | | | | Ra ≤ 14.918 :
| | | | | | | Tmin ≤ 24.45 : LM54</td>
<td>LM num: 14
ET_0 = -0.0154 * Tmean + 0.192 * Tmax - 0.0018 * Tmin + 0.1251 * Ra - 2.3983

LM num: 15
ET_0 = -0.1479 * Tmean + 0.2579 * Tmax - 0.0018 * Tmin + 0.0141 * Ra + 0.2144

LM num: 16
ET_0 = 0.0857 * Tmean + 0.1059 * Tmax - 0.0018 * Tmin + 0.0307 * Ra - 0.7997

LM num: 17
ET_0 = 0.3503 * Tmean + 0.0337 * Tmax - 0.3052 * Tmin + 0.0281 * Ra + 0.5695

LM num: 18
ET_0 = -0.0486 * Tmean + 0.001 * Tmax + 0.0812 * Tmin + 0.3109 * Ra + 48.6714

LM num: 19
ET_0 = -0.0758 * Tmean + 0.001 * Tmax + 0.1685 * Tmin + 0.3623 * Ra + 75.7266

LM num: 20
ET_0 = 0.0003 * Tmean + 0.0061 * Tmax + 0.0055 * Tmin + 0.1567 * Ra + 1.0434

LM num: 21
ET_0 = 0.0003 * Tmean + 0.0061 * Tmax + 0.0055 * Tmin + 0.1714 * Ra + 1.1579

LM num: 22
ET_0 = -0.0047 * Tmean + 0.0061 * Tmax - 0.2973 * Tmin + 0.3694 * Ra + 7.7452

LM num: 23
ET_0 = -0.0051 * Tmean + 0.0061 * Tmax - 0.3195 * Tmin + 0.3951 * Ra + 8.5883

LM num: 24
ET_0 = -0.0051 * Tmean + 0.0061 * Tmax - 0.3195 * Tmin + 0.3951 * Ra + 8.4212

LM num: 25
ET_0 = -0.0365 * Tmean + 0.0061 * Tmax - 0.0845 * Tmin + 0.301 * Ra + 38.3346

LM num: 26
ET_0 = 0.0424 * Tmean + 0.0034 * Tmax + 0.0061 * Tmin + 0.2371 * Ra - 0.0379

LM num: 27
ET_0 = 0.0001 * Tmean + 0.0034 * Tmax + 0.2073 * Tmin + 0.0336 * Ra - 1.5949</td>
</tr>
</table>

```
|  |  |  |  |  |  |  Tmin >  24.45 : LM55
|  |  |  |  |  |  Ra >  14.918 :
|  |  |  |  |  |  |  Tmin ≤ 27.45 : LM56
|  |  |  |  |  |  |  Tmin >  27.45 : LM57
|  |  |  |  |  Tmin >  29.35 : LM58
|  Tmax >  39.938 :
|  |  Ra ≤ 12.546 :
|  |  |  Ra ≤ 8.482 :
|  |  |  |  Ra ≤ 7.87 : LM59
|  |  |  |  Ra >  7.87 : LM60
|  |  |  Ra >  8.482 :
|  |  |  |  Tmin ≤ 28.1 : LM61
|  |  |  |  Tmin >  28.1 : LM62
|  |  Ra >  12.546 :
|  |  |  Tmax ≤ 42.55 :
|  |  |  |  Tmin ≤ 29.7 : LM63
|  |  |  |  Tmin >  29.7 : LM64
|  |  |  Tmax >  42.55 :
|  |  |  |  Ra ≤ 16.185 : LM65
|  |  |  |  Ra >  16.185 :
|  |  |  |  |  Tmax ≤ 45.3 :
|  |  |  |  |  |  Ra ≤ 16.39 : LM66
|  |  |  |  |  |  Ra >  16.39 : LM67
|  |  |  |  |  Tmax >  45.3 : LM68
```

LM num: 28
ET_0 = -0.0748 * Tmean + 0.0009 * Tmax - 0.0036 * Tmin + 0.5039 * Ra + 75.1206

LM num: 29
ET_0 = 0.0001 * Tmean + 0.0145 * Tmax - 0.054 * Tmin - 0.0947 * Ra + 5.4603

LM num: 30
ET_0 = -0.0002 * Tmean + 0.0105 * Tmax - 0.0701 * Tmin - 0.1344 * Ra + 6.0692

LM num: 31
ET_0 = 0.1156 * Tmean + 0.0507 * Tmax - 0.0309 * Tmin - 0.0478 * Ra + 0.7195

LM num: 32
ET_0 = 0.3243 * Tmean + 0.0009 * Tmax - 0.1955 * Tmin + 0.1093 * Ra - 1.8256

LM num: 33
ET_0 = 0 * Tmean + 0.1537 * Tmax - 0.0009 * Tmin + 0.029 * Ra - 1.1669

LM num: 34
ET_0 = -0.0979 * Tmean + 0.0007 * Tmax + 0.1843 * Tmin + 0.4766 * Ra + 96.568

LM num: 35
ET_0 = -0.0762 * Tmean - 0.0368 * Tmax + 0.0126 * Tmin + 0.0842 * Ra + 79.3077

LM num: 36
ET_0 = 0.0001 * Tmean - 0.0106 * Tmax - 0.0416 * Tmin + 0.4334 * Ra - 1.147

LM num: 37
ET_0 = -0.0163 * Tmean - 0.0502 * Tmax + 0.0549 * Tmin + 1.3162 * Ra + 0.1336

LM num: 38
ET_0 = -0.1242 * Tmean - 0.0136 * Tmax + 0.147 * Tmin + 0.1487 * Ra + 128.616

LM num: 39
ET_0 = 0.001 * Tmean + 0.1975 * Tmax + 0.0227 * Tmin + 0.1202 * Ra - 3.8525

LM num: 40
ET_0 = 0.2844 * Tmean + 0.0201 * Tmax - 0.282 * Tmin + 0.5213 * Ra - 4.8537

LM num: 41
ET_0 = -0.0055 * Tmean + 0.2207 * Tmax - 0.0975 * Tmin + 0.0859 * Ra - 1.5875

	LM num: 42 ET_0 = -0.0136 * Tmean + 0.7841 * Tmax - 0.0632 * Tmin + 0.0141 * Ra - 20.8525 LM num: 43 ET_0 = -0.0136 * Tmean + 0.2076 * Tmax - 0.0506 * Tmin + 0.2595 * Ra - 4.085 LM num: 44 ET_0 = -0.0136 * Tmean + 0.1934 * Tmax - 0.1409 * Tmin + 0.1364 * Ra + 0.3336 LM num: 45 ET_0 = 0 * Tmean + 0.2035 * Tmax - 0.0994 * Tmin + 0.0074 * Ra + 0.5209 LM num: 46 ET_0 = 0 * Tmean + 0.3293 * Tmax - 0.0974 * Tmin + 0.2211 * Ra - 6.8822 LM num: 47 ET_0 = 0 * Tmean + 0.4095 * Tmax - 0.253 * Tmin + 0.1368 * Ra - 4.4634 LM num: 48 ET_0 = 0 * Tmean + 0.1253 * Tmax - 0.0608 * Tmin + 0.0945 * Ra + 0.7244 LM num: 49 ET_0 =0 * Tmean + 0.1253 * Tmax - 0.0608 * Tmin + 0.0945 * Ra + 0.8233 LM num: 50 ET_0 = 0 * Tmean + 0.1253 * Tmax - 0.0608 * Tmin - 2.4711 * Ra + 42.8921 LM num: 51 ET_0 = 0 * Tmean + 0.1126 * Tmax - 0.3299 * Tmin + 0.0945 * Ra + 9.4554 LM num: 52 ET_0 = 0 * Tmean + 0.325 * Tmax - 0.252 * Tmin + 0.265 * Ra - 3.4958 LM num: 53 ET_0 = 0 * Tmean + 0.106 * Tmax - 0.0976 * Tmin + 1.5975 * Ra - 21.89 LM num: 54 ET_0 = -0.0041 * Tmean + 0.3827 * Tmax + 0.0555 * Tmin + 0.1906 * Ra - 11.7983 LM num: 55 ET_0 = -0.0041 * Tmean + 0.3576 * Tmax - 0.0977 * Tmin + 0.2136 * Ra - 8.0507

	LM num: 56 $ET_0 = -0.0041 * Tmean + 0.2525 * Tmax - 0.0312 * Tmin + 0.0171 * Ra - 2.0945$ LM num: 57 $ET_0 = 0.4947 * Tmean + 0.0262 * Tmax - 0.4129 * Tmin + 0.2685 * Ra - 3.6808$ LM num: 58 $ET_0 = -0.0055 * Tmean + 0.3603 * Tmax - 0.2839 * Tmin + 0.3018 * Ra - 3.9747$ LM num: 59 $ET_0 = 0.0048 * Tmean + 0.0159 * Tmax + 0.2523 * Tmin + 1.8181 * Ra - 16.5588$ LM num: 60 $ET_0 = 0.0003 * Tmean - 0.1337 * Tmax + 0.1811 * Tmin + 0.2034 * Ra + 5.0849$ LM num: 61 $ET_0 = -0.0001 * Tmean + 0.1697 * Tmax + 0.1427 * Tmin + 0.3542 * Ra - 8.5106$ LM num: 62 $ET_0 = -0.0001 * Tmean + 0.019 * Tmax + 0.0101 * Tmin + 0.6059 * Ra - 0.1704$ LM num: 63 $ET_0 = -0.0002 * Tmean + 0.4438 * Tmax - 0.0993 * Tmin + 0.4252 * Ra - 14.507$ LM num: 64 $ET_0 = -0.0002 * Tmean + 0.4729 * Tmax - 0.2464 * Tmin + 0.7366 * Ra - 16.6866$ LM num: 65 $ET_0 = -0.0002 * Tmean + 0.2871 * Tmax - 0.0013 * Tmin + 0.2557 * Ra - 7.6575$ LM num: 66 $ET_0 = -0.0014 * Tmean + 0.06 * Tmax - 0.075 * Tmin + 0.7214 * Ra - 2.9764$ LM num: 67 $ET_0 = -0.0037 * Tmean + 0.0547 * Tmax + 0.1074 * Tmin - 4.0612 * Ra + 71.9315$ LM num: 68 $ET_0 = -0.0005 * Tmean + 0.1471 * Tmax + 0.0922 * Tmin + 1.7732 * Ra - 27.6132$

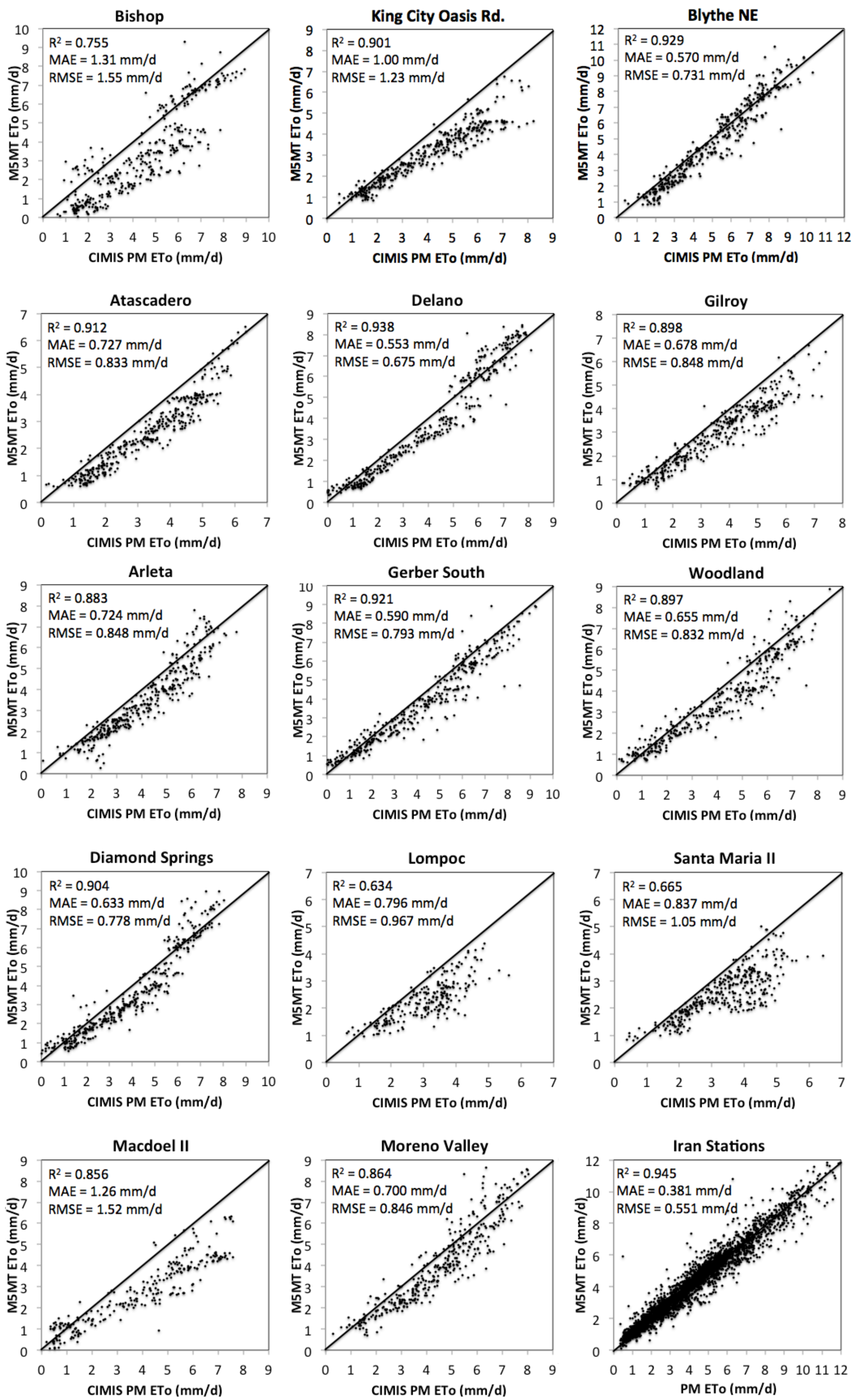


Figure 15. Estimated ET_0 values from M5MT1 versus observations.

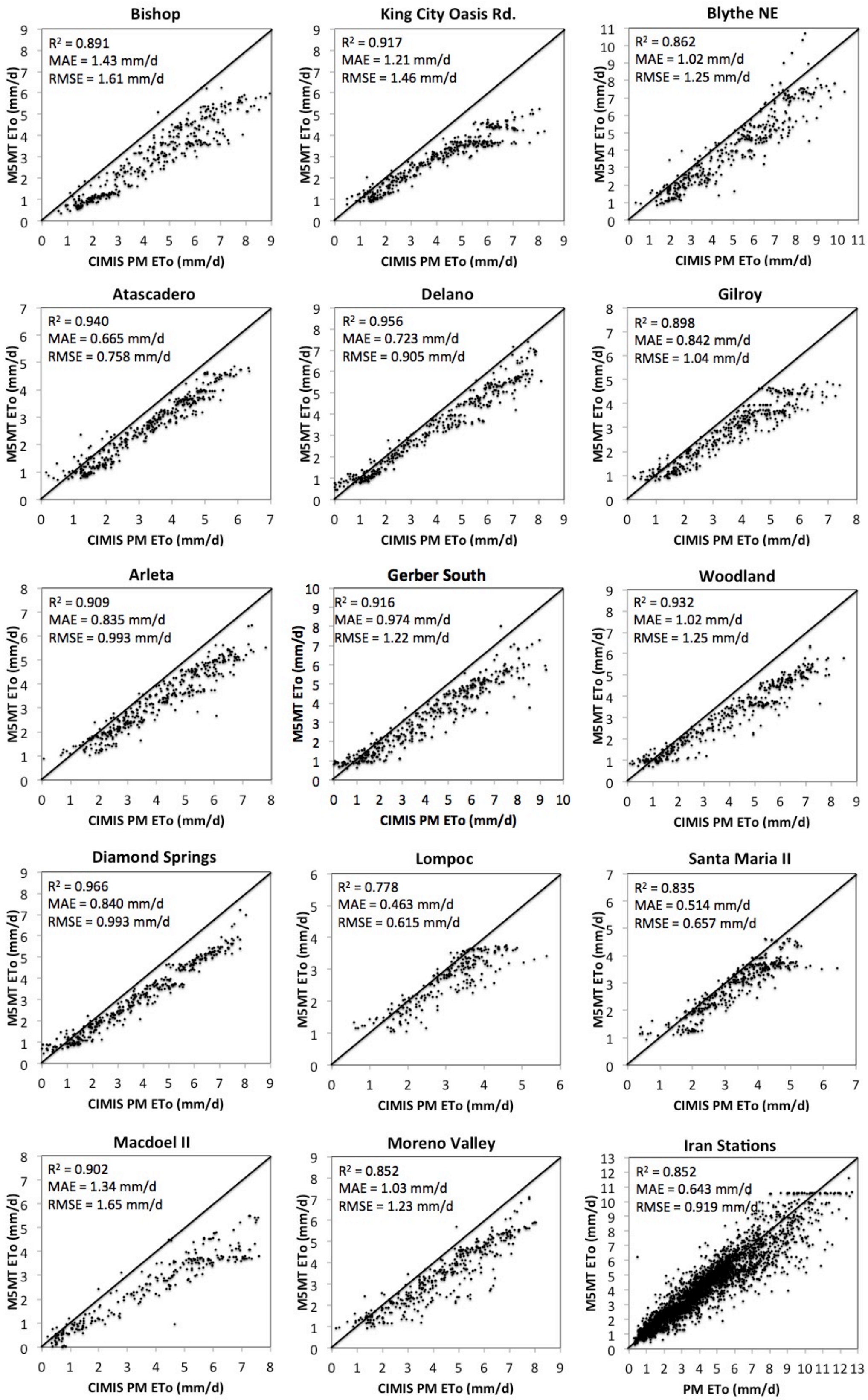


Figure 16. Estimated ET_0 values from M5MT2 versus observations.

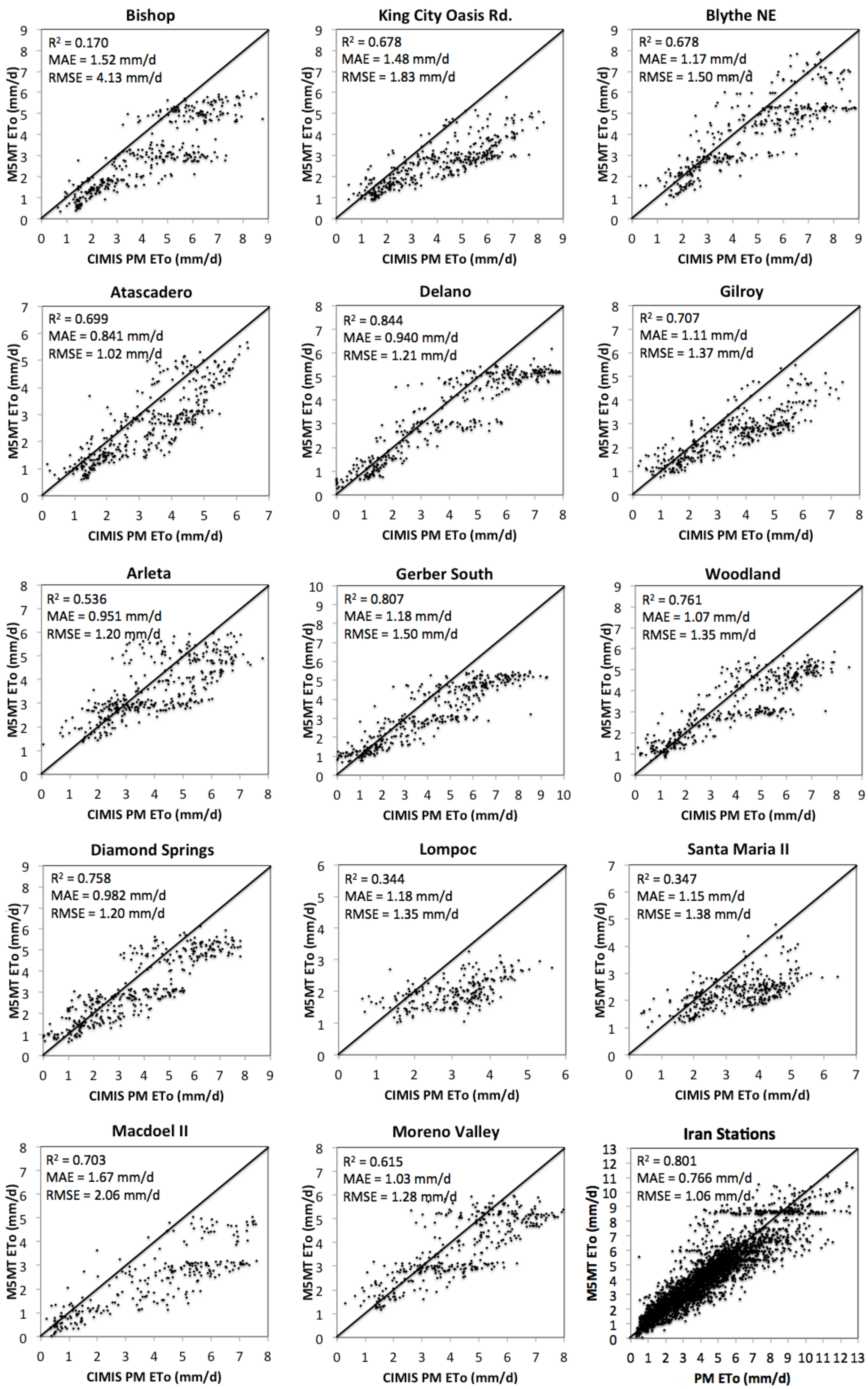


Figure 17. Estimated ET_0 values from M5MT3 versus observations.

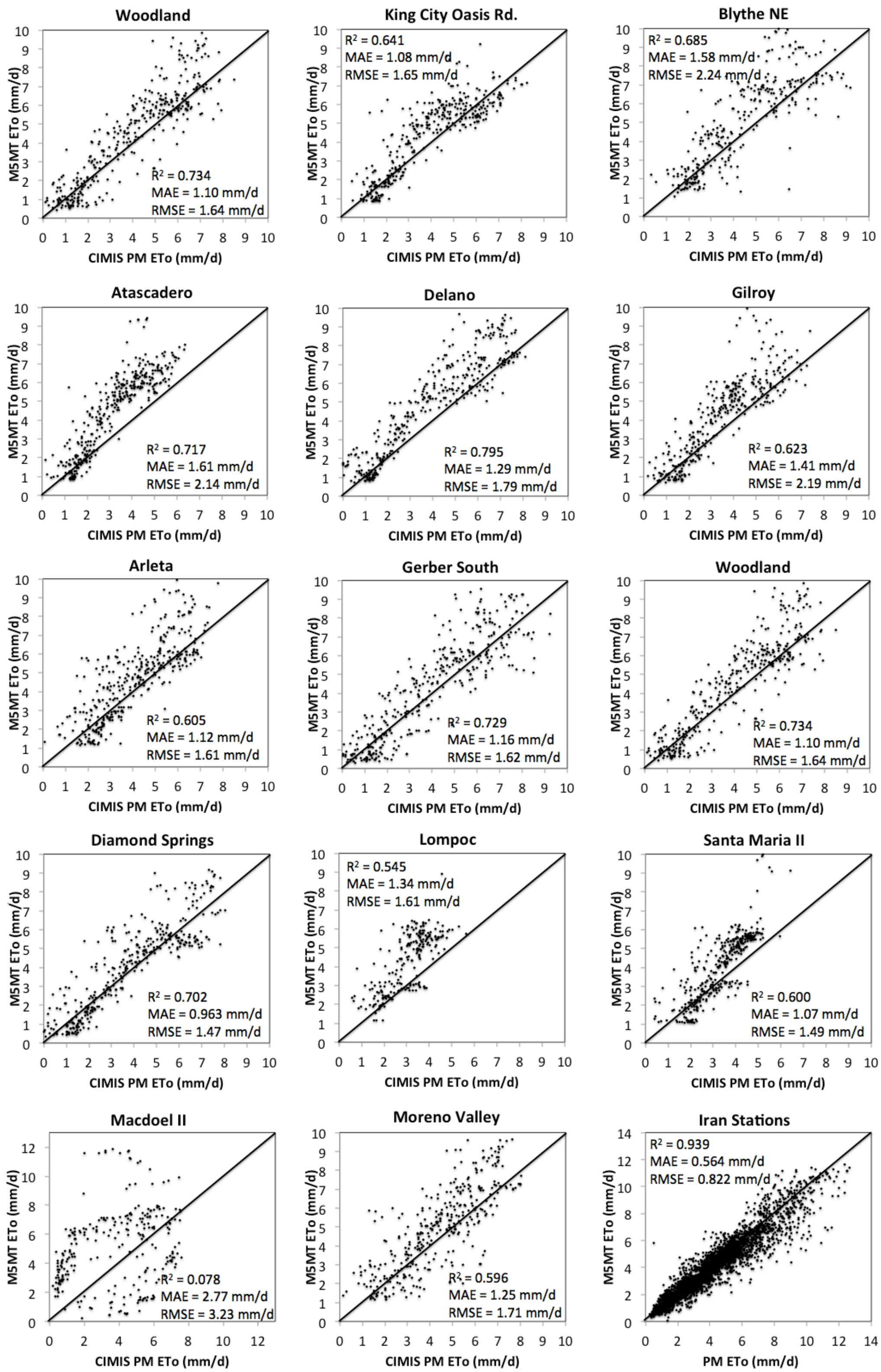
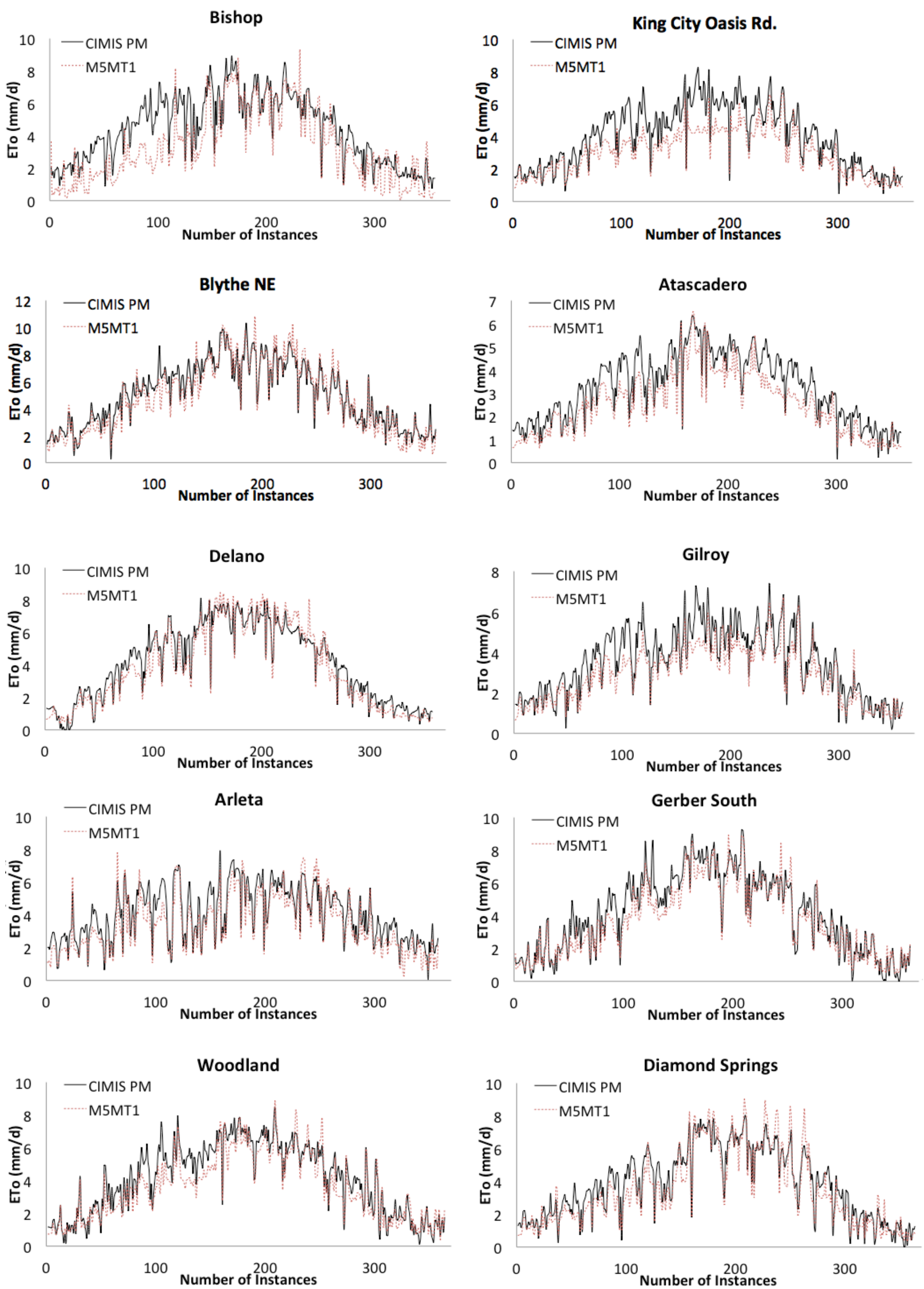


Figure 18. Estimated ET_0 values from M5MT4 versus observations.


Bishop
King City Oasis Rd.
Blythe NE
Atascadero
Delano
Gilroy
Arleta
Gerber South
Woodland
Diamond Springs
CIMIS PM
M5MT1
ETo (mm/d)
Number of Instances

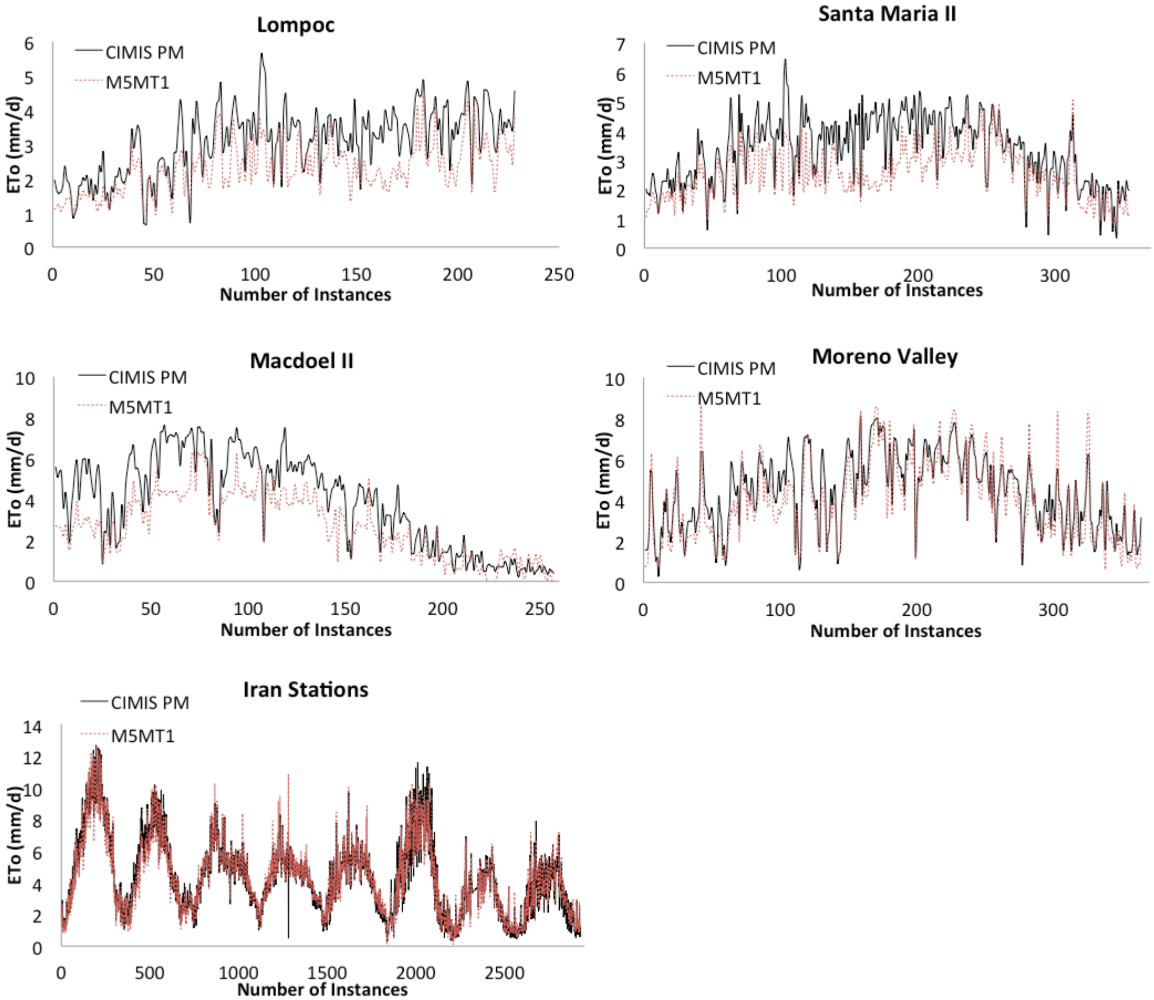
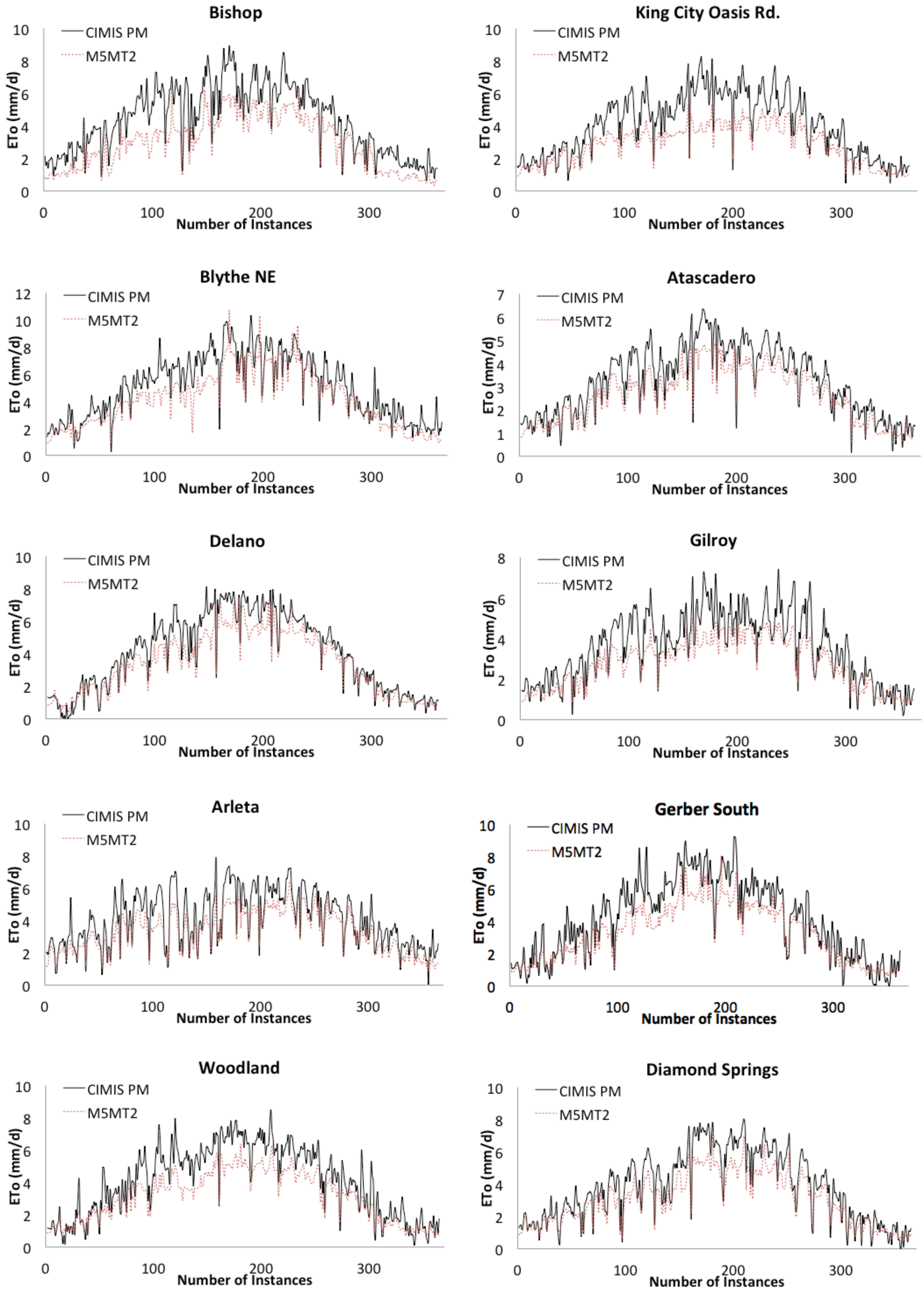


Figure 19. Time series of observed and estimated ET_0 values from M5MT1.


Bishop
King City Oasis Rd.
Blythe NE
Atascadero
Delano
Gilroy
Arleta
Gerber South
Woodland
Diamond Springs
CIMIS PM
M5MT2
ETo (mm/d)
Number of Instances
0
100
200
300

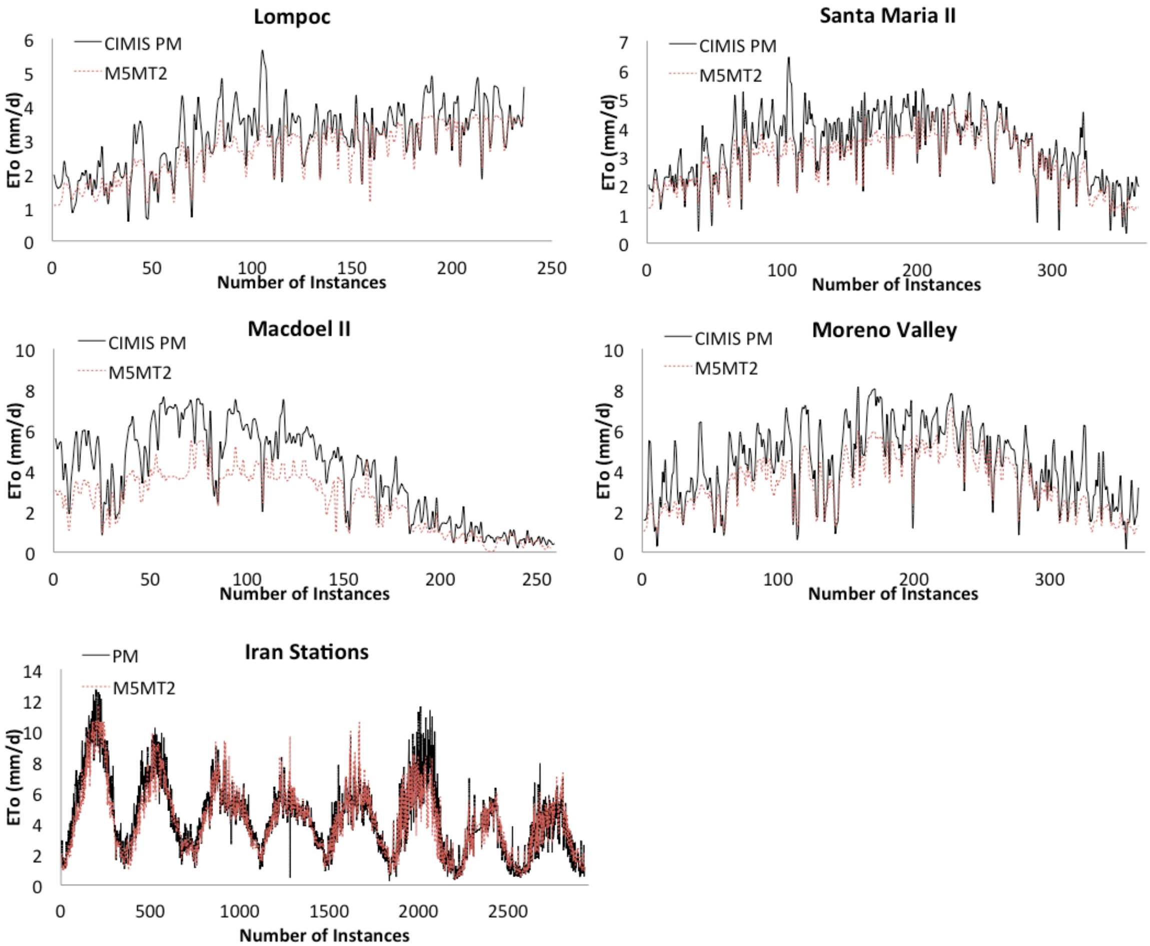
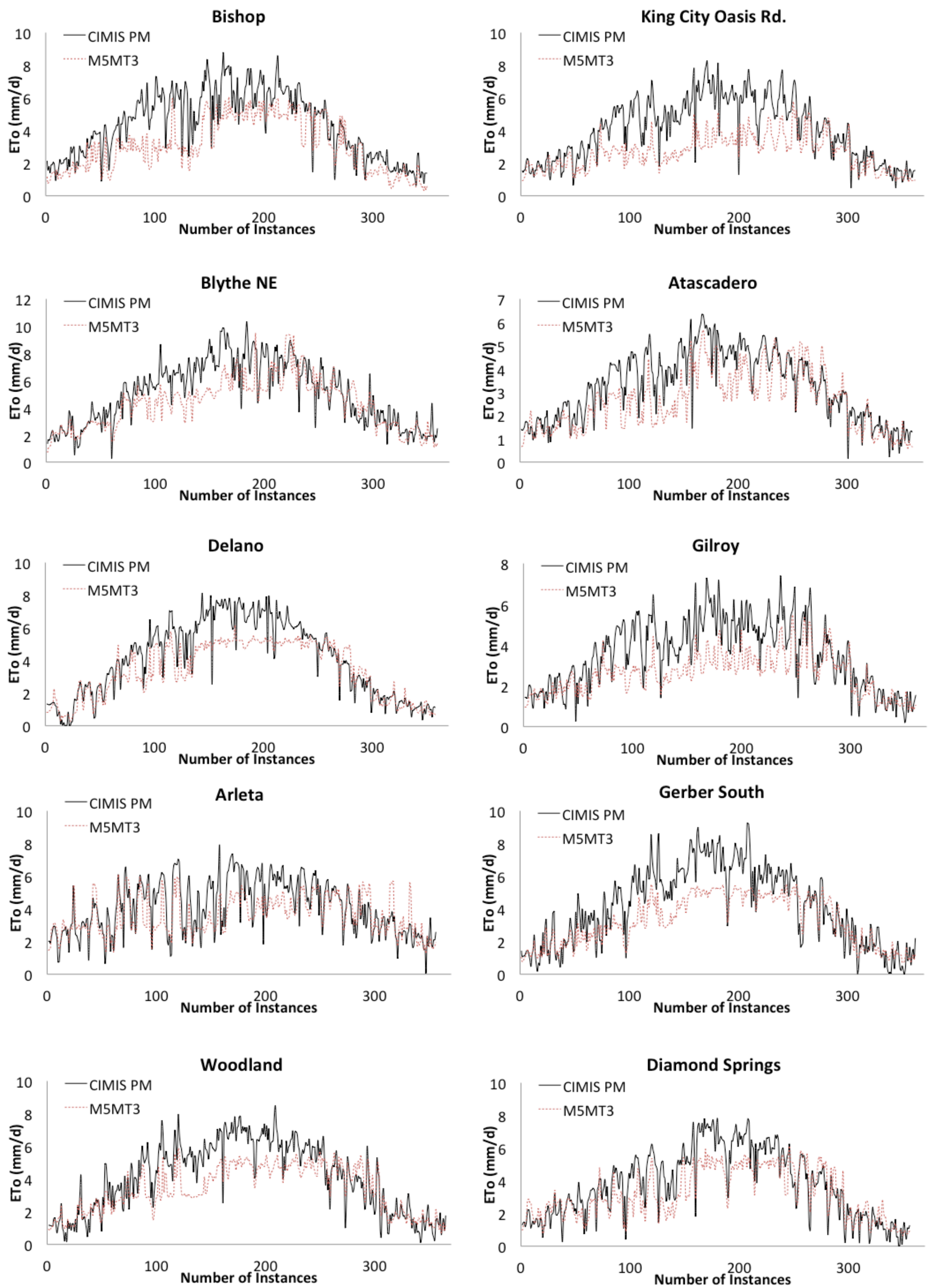


Figure 20. Time series of observed and estimated ET_0 values from M5MT2.


Bishop
King City Oasis Rd.
Blythe NE
Atascadero
Delano
Gilroy
Arleta
Gerber South
Woodland
Diamond Springs
CIMIS PM
M5MT3
ETo (mm/d)
Number of Instances

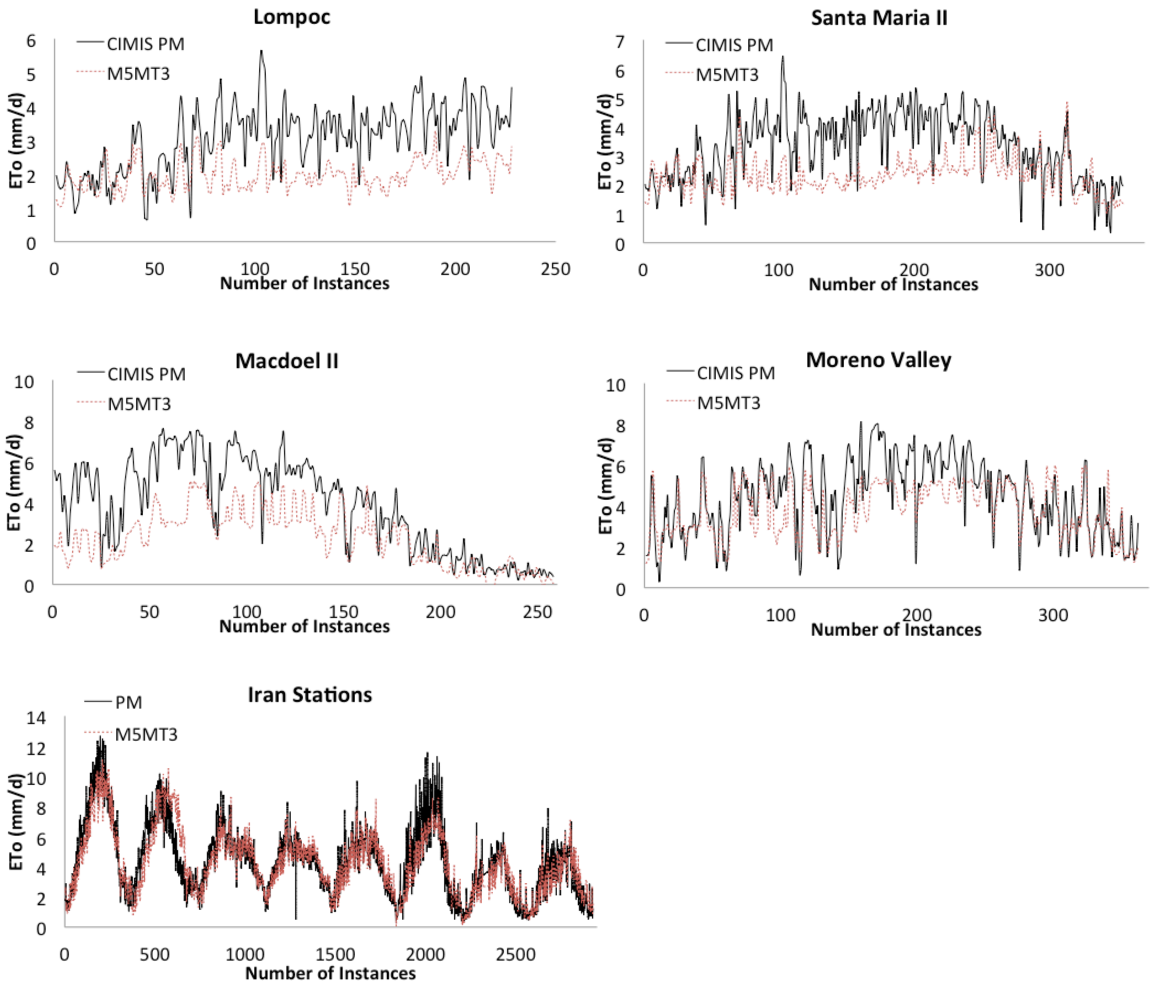
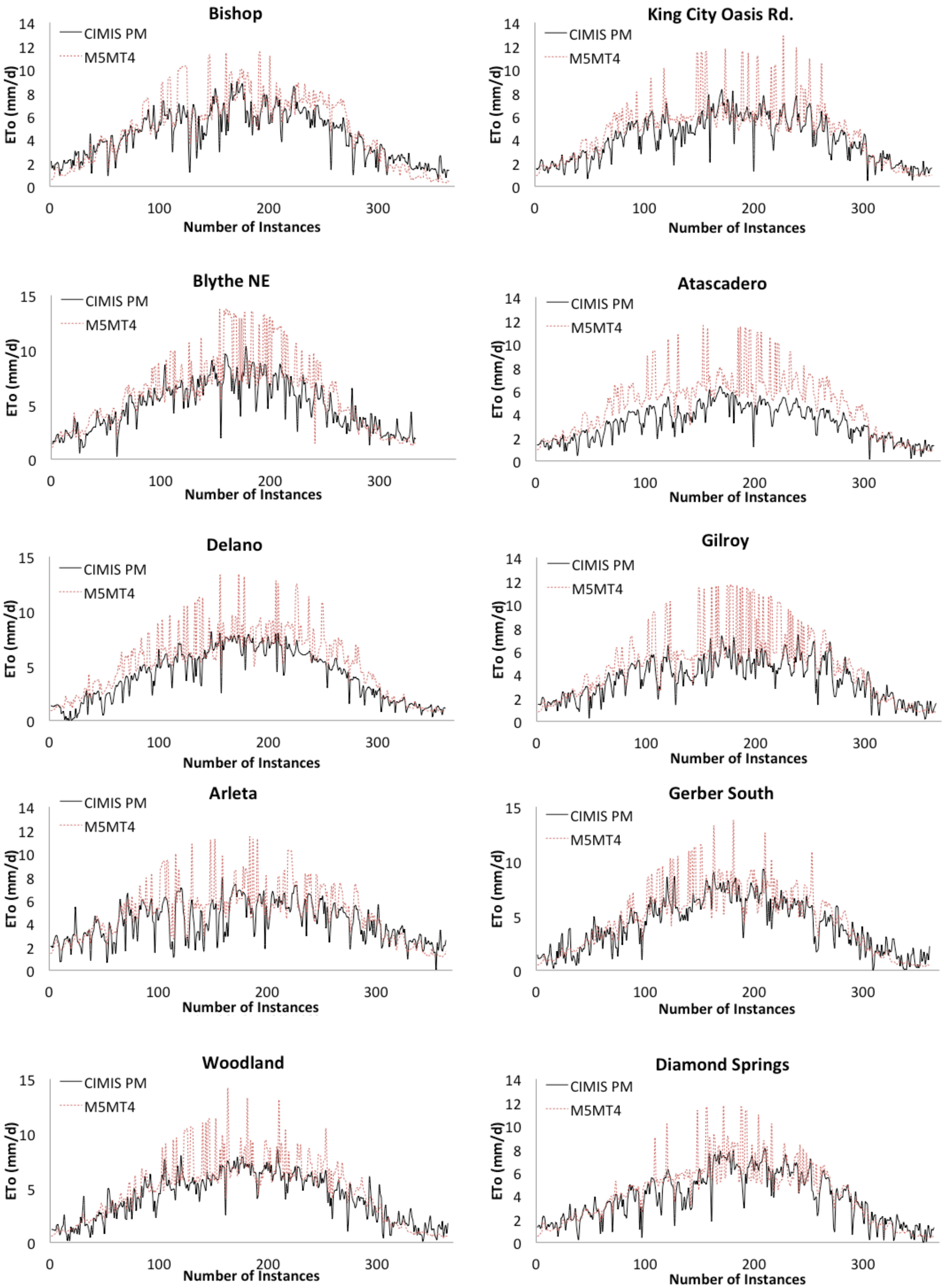


Figure 21. Time series of observed and estimated ET_0 values from M5MT3.


Bishop
King City Oasis Rd.
Blythe NE
Atascadero
Delano
Gilroy
Arleta
Gerber South
Woodland
Diamond Springs
CIMIS PM
M5MT4
ETo (mm/d)
Number of Instances

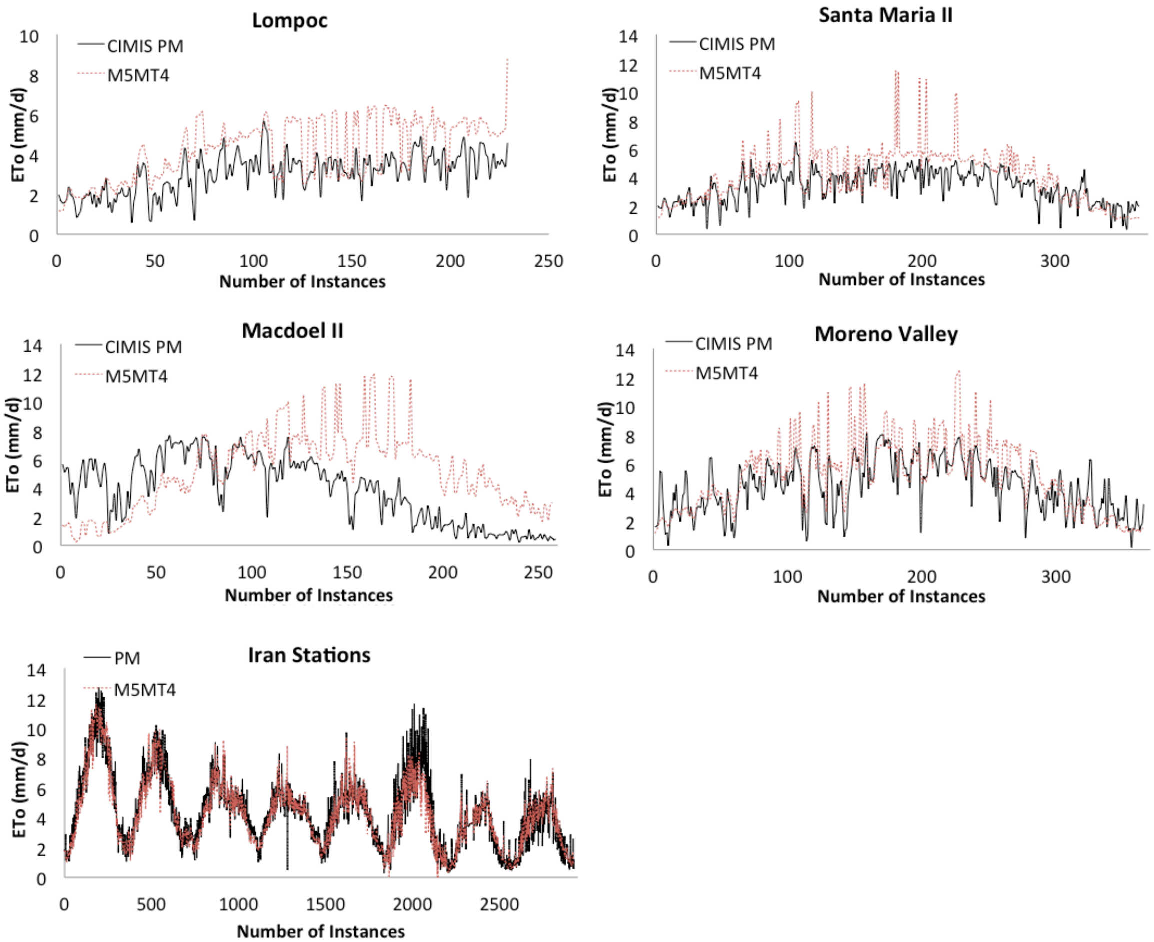


Figure 22. Time series of observed and estimated ET_0 values from M5MT4.

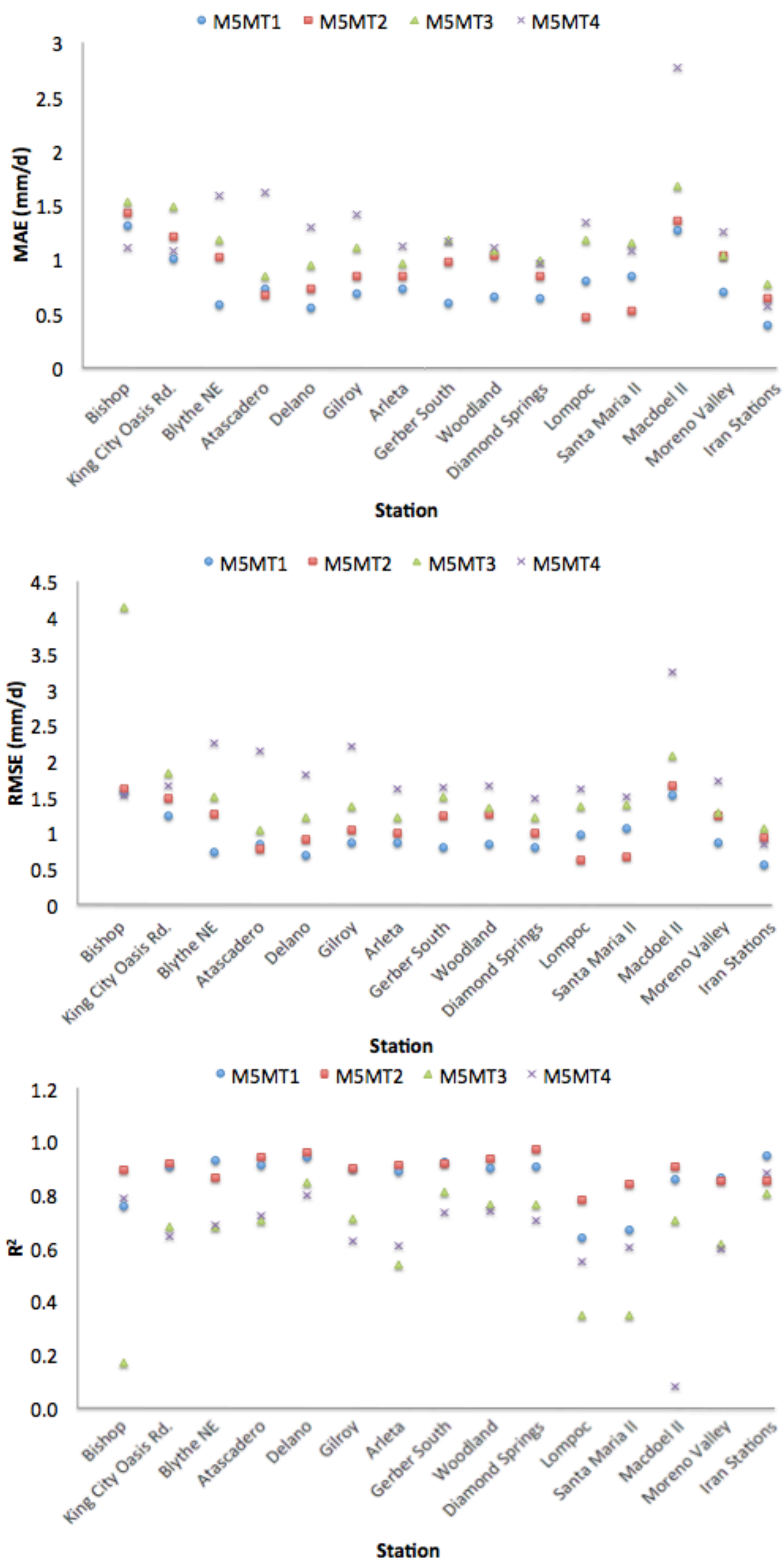


Figure 23. Comparing performance of M5MT models in different stations.

Table 13. Statistical metrics of ET_0 estimates from M5MT1 and M5MT2.

Model	Station	MAE (mm/d)	RMSE (mm/d)	R^2
M5MT1	Bishop	1.31	1.55	0.75
	King City Oasis Rd.	1.00	1.23	0.90
	Blythe NE	0.57	0.73	0.93
	Atascadero	0.73	0.83	0.91
	Delano	0.55	0.68	0.94
	Gilroy	0.68	0.85	0.90
	Arleta	0.72	0.85	0.88
	Gerber South	0.59	0.79	0.92
	Woodland	0.66	0.83	0.90
	Diamond Springs	0.63	0.78	0.90
	Lompoc	0.80	0.97	0.63
	Santa Maria II	0.84	1.05	0.66
	Macdoel II	1.26	1.52	0.86
	Moreno Valley	0.70	0.85	0.86
	All Iran Stations	0.38	0.55	0.94
M5MT2	Bishop	1.43	1.61	0.89
	King City Oasis Rd.	1.21	1.46	0.92
	Blythe NE	1.02	1.25	0.86
	Atascadero	0.67	0.76	0.94
	Delano	0.72	0.90	0.96
	Gilroy	0.84	1.04	0.90
	Arleta	0.84	0.99	0.91
	Gerber South	0.97	1.22	0.92
	Woodland	1.02	1.25	0.93
	Diamond Springs	0.84	0.99	0.97
	Lompoc	0.46	0.61	0.78
	Santa Maria II	0.51	0.66	0.84
	Macdoel II	1.34	1.65	0.90
	Moreno Valley	1.03	1.23	0.85
	All Iran Stations	0.64	0.92	0.85

Table 14. Statistical metrics of ET_0 estimates from M5MT3 and M5MT4.

Model	Station	MAE (mm/d)	RMSE (mm/d)	R^2
M5MT3	Bishop	1.52	4.13	0.17
	King City Oasis Rd.	1.48	1.83	0.68
	Blythe NE	1.17	1.50	0.68
	Atascadero	0.84	1.02	0.70
	Delano	0.94	1.21	0.84
	Gilroy	1.11	1.37	0.71
	Arleta	0.95	1.20	0.54
	Gerber South	1.18	1.50	0.81
	Woodland	1.07	1.35	0.76
	Diamond Springs	0.98	1.20	0.76
	Lompoc	1.18	1.35	0.34
	Santa Maria II	1.15	1.38	0.35
	Macdoel II	1.67	2.06	0.70
	Moreno Valley	1.03	1.28	0.61
	All Iran Stations	0.77	1.06	0.80
M5MT4	Bishop	1.10	1.53	0.78
	King City Oasis Rd.	1.08	1.65	0.64
	Blythe NE	1.58	2.24	0.68
	Atascadero	1.61	2.14	0.72
	Delano	1.29	1.79	0.79
	Gilroy	1.41	2.19	0.62
	Arleta	1.12	1.61	0.60
	Gerber South	1.16	1.62	0.73
	Woodland	1.10	1.64	0.73
	Diamond Springs	0.96	1.47	0.70
	Lompoc	1.34	1.61	0.55
	Santa Maria II	1.07	1.49	0.60
	Macdoel II	2.77	3.23	0.08
	Moreno Valley	1.25	1.71	0.60
	All Iran Stations	0.56	0.82	0.88

4.3 GEP

The GeneXpro program was used to create a GEP model for each input combination. The parameters that were specified by the user were terminal and function sets, the head size, the number of chromosomes, the number of genes, and the linking function.

The terminal and function sets were specified to create the genes of the chromosomes. The terminal function set consisted of W_s, RH_{mean}, T_{mean}, T_{max}, T_{min}, R_s, and R_a. The function set consisted of the following arithmetic functions: +, -, x, /, square root, exponential, natural log, x^2, x^3, $x^{1/3}$, sine, cosine, and arctangent. Next, the length of the head and the number of genes per chromosome were specified. The head size dictated the amount of terminal and functional symbols used in each chromosome and the number of genes determined how many sub-expression trees would be generated in the model.

In this study, the head size was set to a default value of seven. 30 chromosomes and three genes were used (Ferreira, 2001; Emamgolizadeh et al., 2015). The addition mathematical operator was utilized as the linking function (Zakaria et al., 2010; Hashmi et al., 2011; Azamathulla and Ahmad, 2012; Azamathulla and Jarrett, 2013). Lastly, RMSE was used as the fitness function. The applied parameters are summarized in Table 15.

Table 15. Summary of GEP model parameters.

Parameter	Parameter Setting
Terminal set	W_s, RH_{mean}, T_{mean}, T_{max}, T_{min}, R_s, and R_a
Function set	+, -, x, /, sqrt, exp, ln, x^2, x^3, $x^{1/3}$, sin, cos, and arctan
Head size	7
Number of chromosomes	30
Number of genes	3
Linking function	Addition

Table 16 presents the ET_0 mathematical expressions generated by the four GEP models. GEP1-GEP3 models successfully determined a relationship between all input variables and ET_0. The GEP4 model, however, was not able to establish a relationship between all input variables (T_{mean}, T_{min}, T_{max}, and R_a) and ET_0, resulting in the omission T_{min}.

Figures 24-27 show a comparison between ET_0 estimates from the four GEP models and those from the PM equation for all stations. As indicated, it was observed that the majority of the ET_0 estimates from GEP1 and GEP2 conformed to the 45-degree line better than those from GEP3 and GEP4. Figures 28-31 indicate the time series of the calculated ET_0 values from the four GEP models. For comparison, the estimated ET_0 values from the PM equation were also plotted on the same figures. Underestimation of ET_0 was observed from GEP2 and GEP3, while GEP1 and GEP4 mostly overestimated ET_0. The under- and overestimation is most likely attributed to measurement errors and the exclusion of California stations from the training process. However, most of the ET_0 estimates from the GEP models coincided with the daily variations in observed ET_0 values.

The performance of GEP models in different stations was compared in Figure 32 using *MAE*, *RMSE,* and R^2. GEP1 provided the best results in all stations. Pertaining to the California stations, average *MAE* of ET_0 estimates from GEP2, GEP3, and GEP4 were respectively, 31%, 91%, and 139% higher than those from GEP1. Also, their average *RMSE* were respectively 22%, 79%, and 125% greater than the average *RMSE* value of GEP1. When tested with data from Iran, GEP1 resulted in the lowest *MAE* and *RMSE* values (0.81 mm/d and 1.06 mm/d, respectively) and the highest R^2 (0.81). Performances of GEP2 and GEP4 tested with Iran stations were comparable, with similar *MAE*, *RMSE*, and R^2 values. This indicates GEP may be a plausible approach to estimate ET_0 for areas where only air temperature is available.

Tables 17 and 18 present statistical metrics of ET_0 estimates from the four GEP models. The average *MAE* for all stations tested with GEP2, GEP3, and GEP4 were respectively, 28%, 91%, and 126% greater than that of GEP1. Also, average *RMSE* values for all stations tested with GEP2, GEP3, and GEP4 were respectively, 20%, 75%, and 111% greater than that of GEP1.

Generally, results showed GEP1 to be the most accurate model, followed by GEP2. The results from GEP1 and GEP2 tested with California stations indicated GEP to be a viable approach to estimate ET_0. Metrics from GEP2 further showed GEP's capability to successfully estimate ET_0 with limited climatic data. GEP4 model performance varied when tested with data from different locations. GEP4 performed comparably to GEP2 when tested with Iran stations, while GEP4 performed not as good as GEP1-GEP3 when tested with California stations. This may indicate the inputs of GEP4 to be region dependent.

Table 16. GEP equations corresponding to each model.

Model	Equation
GEP1	$ET_0 = \frac{\left(W_s + (W_s + 3.66)\right) + \left(\left(\frac{RH_{mean}}{9.42} - R_s\right)\right.}{-3.66} + \arctan\left(\left(\sqrt{\exp\left(\left(\frac{RH_{mean}}{4.15}\right)^2 - \left((W_s + 4.15) + 4.15\right)\right)}\right)\right) + W_s - cos\left(arctan(R_s - 9.23) - \frac{T_{mean}}{9.82}\right)$
GEP2	$ET_0 = \left[sin\left(\frac{(7.63 - T_{mean}) - 8.14}{8.14}\right)\right]^3 + arctan\left[cos\left(3.23 + \left(\frac{T_{mean} - 9.07}{12.3}\right)\right)\right] + R_s\left[cos\left(arctan\left((R_s - 3.4) - R_s\right)\right)\right]$
GEP3	$ET_0 = [arctan(T_{mean})]^3 + 2cos\left[0.03\left(T_{mean}arctan(T_{mean}) - (RH_{mean} + 3.14)\right)\right]$
GEP4	$ET_0 = \frac{R_a}{5.81 + cos\left(arctan(T_{max}) + \frac{T_{mean}}{-7.76}\right)} + \frac{R_a}{exp\left[\frac{R_a + 6.02}{T_{max} - 5.47}\right] + 5.47} + \left[arctan\left(cos\left(19.06(e^{T_{max}} - 16.84)\right)\right)\right]^3$

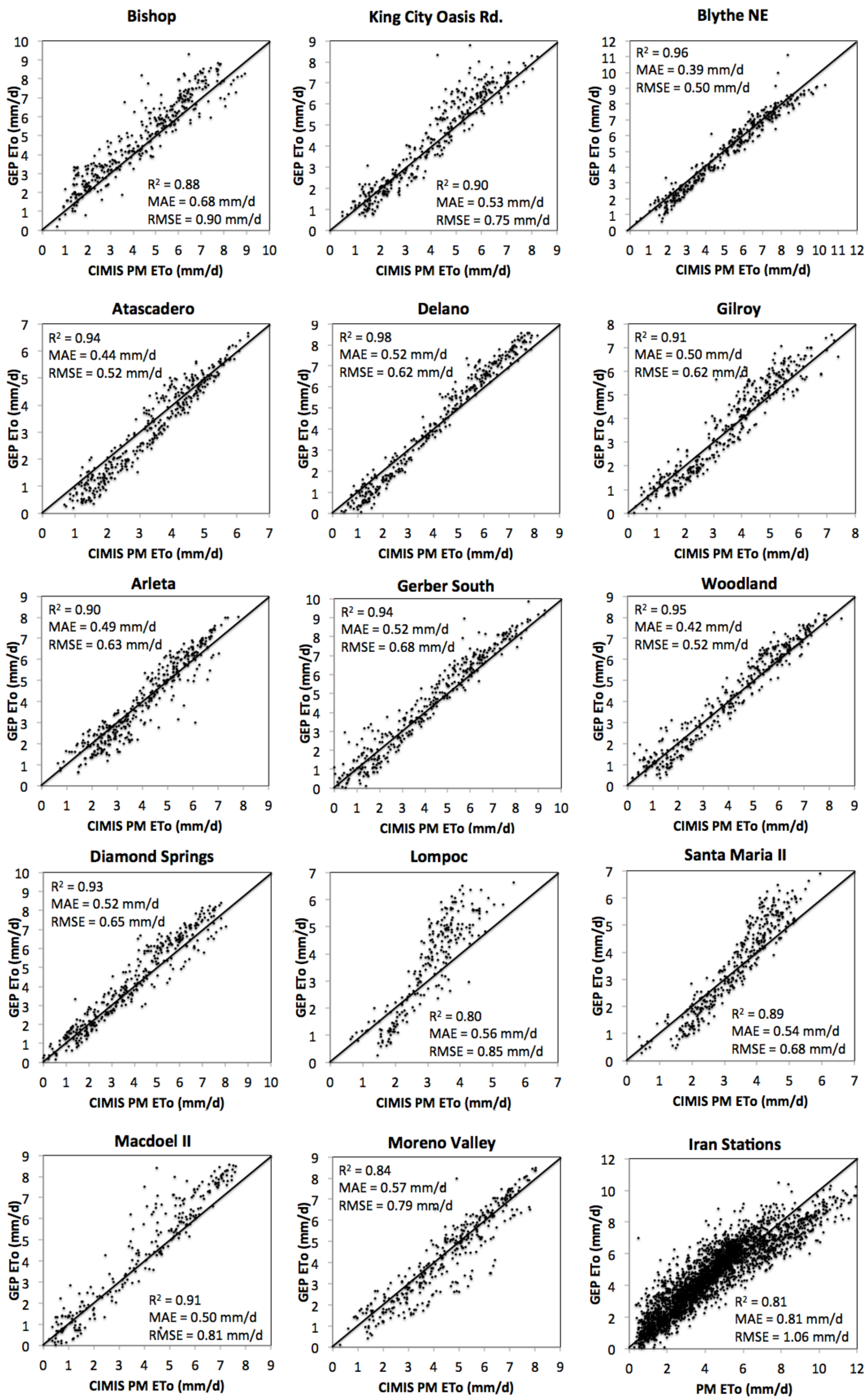


Figure 24. Estimated ET_0 values from GEP1 versus observations.

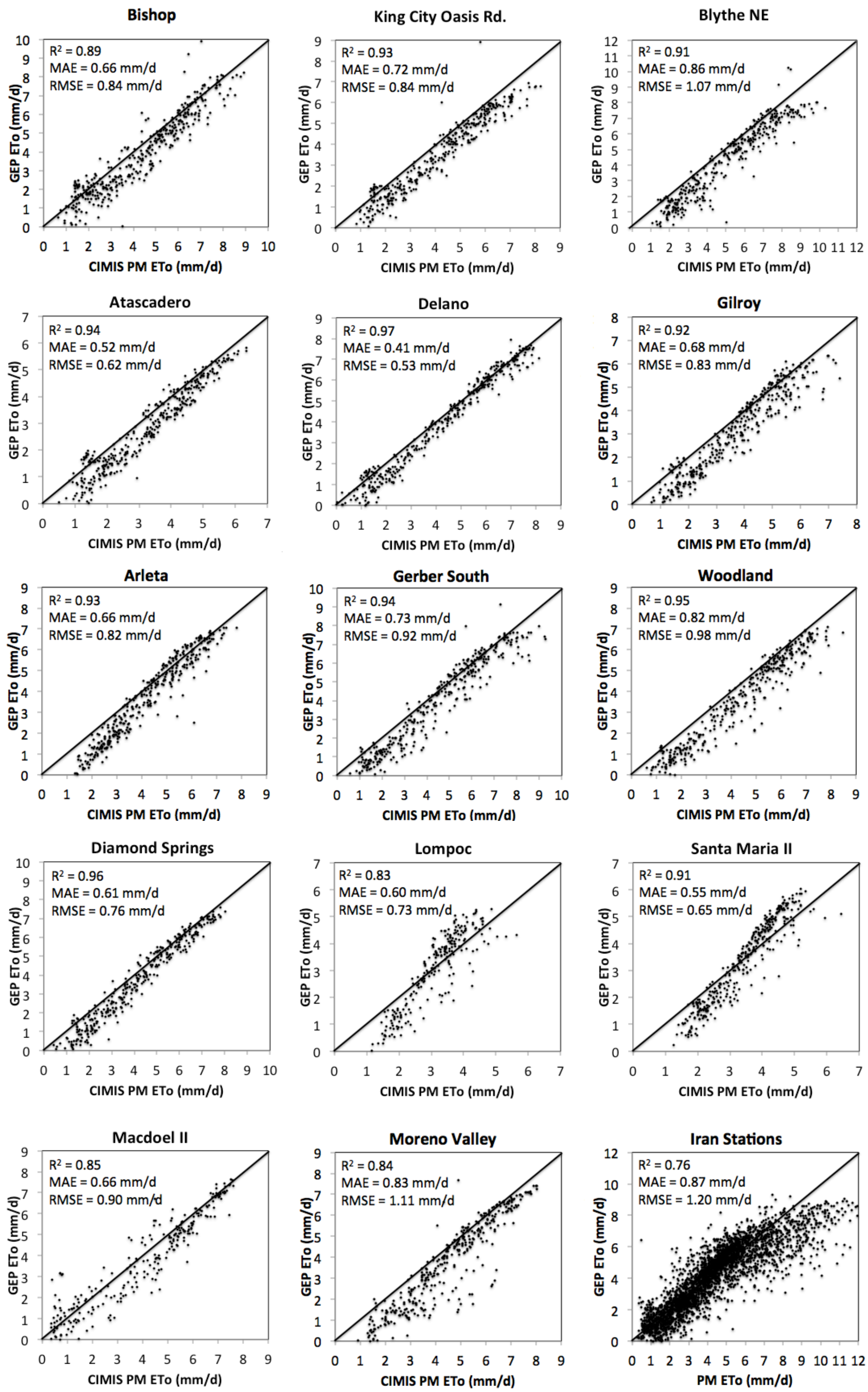


Figure 25. Estimated ET_0 values from GEP2 versus observations.

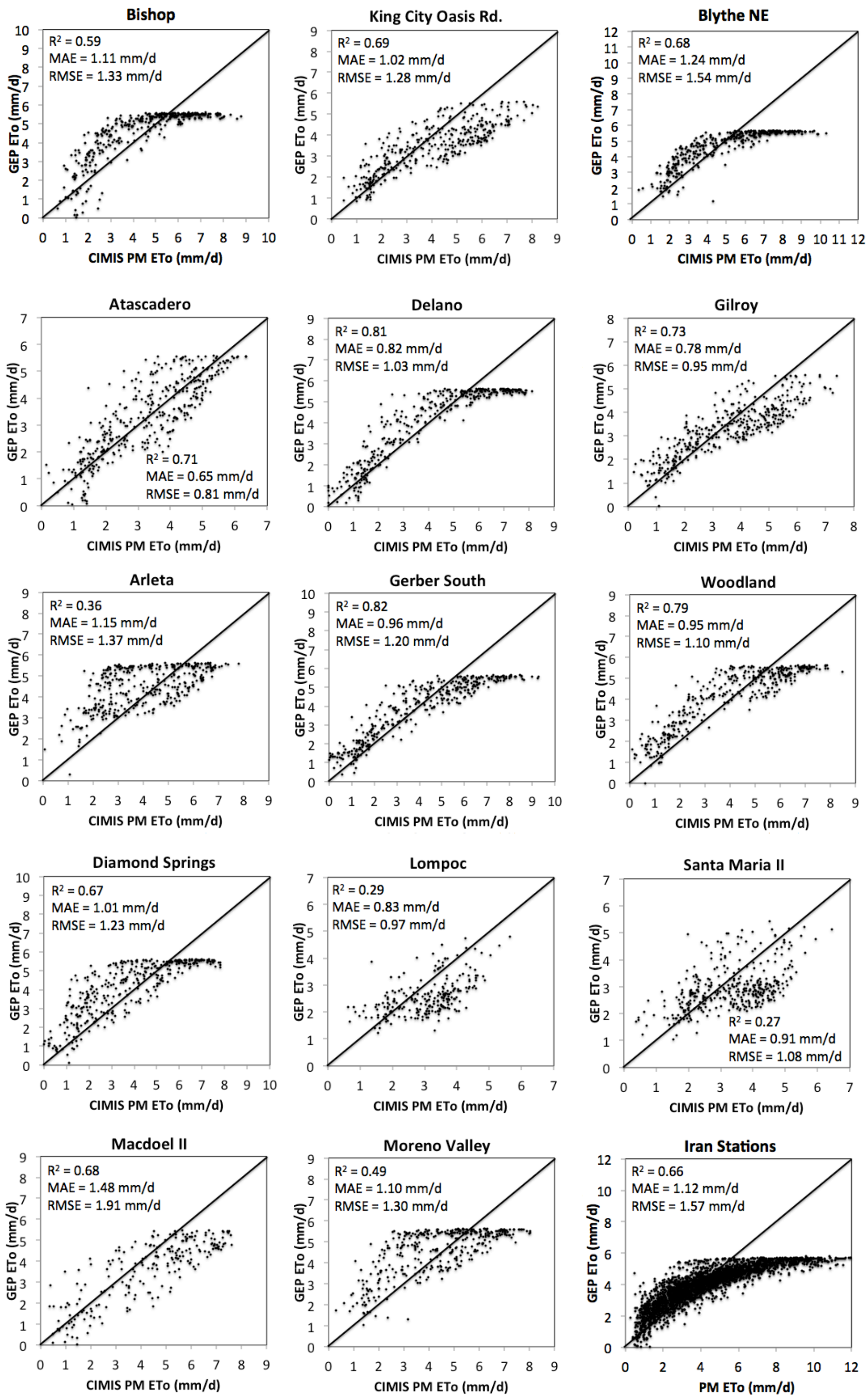


Figure 26. Estimated ET_0 values from GEP3 versus observations.

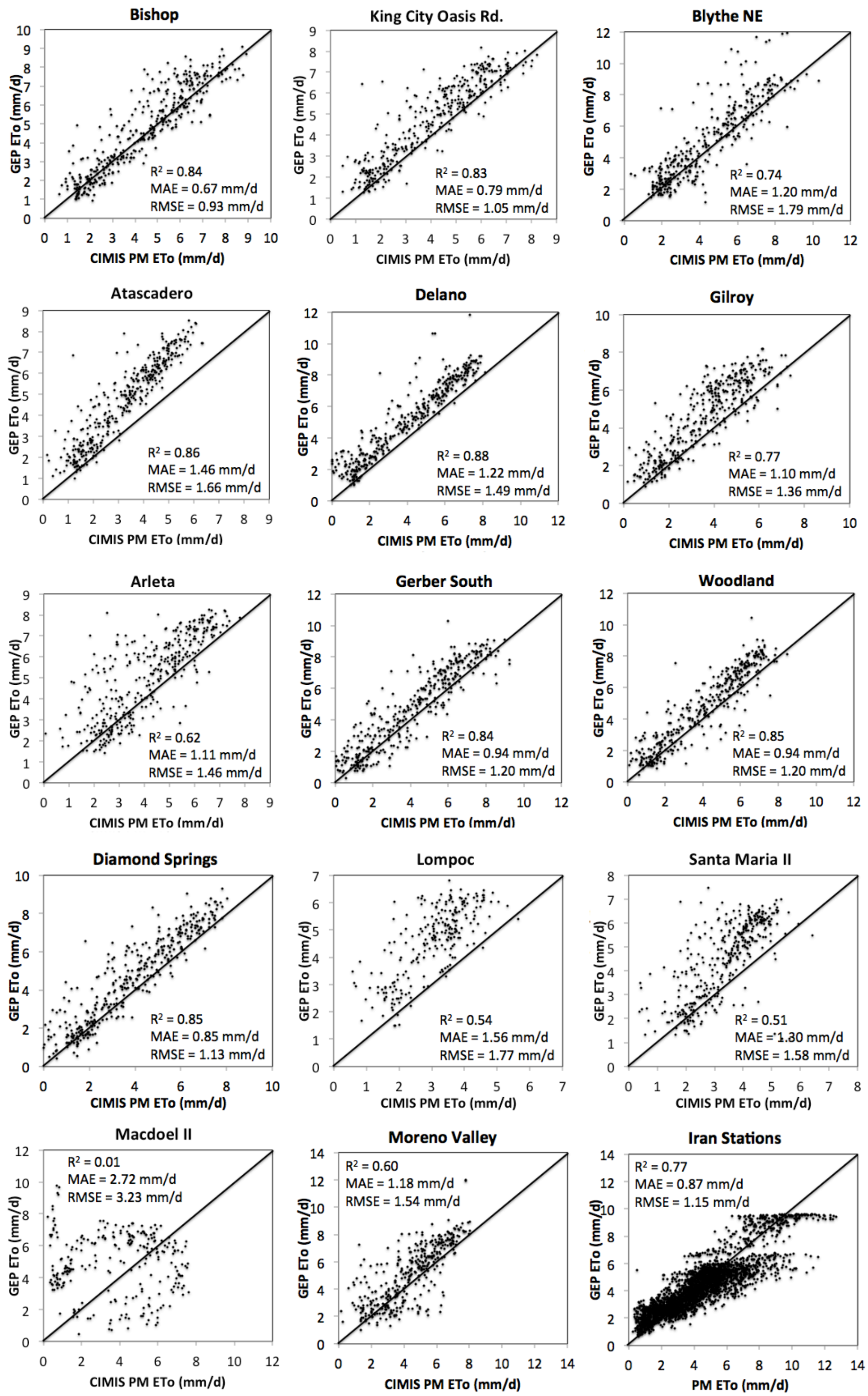
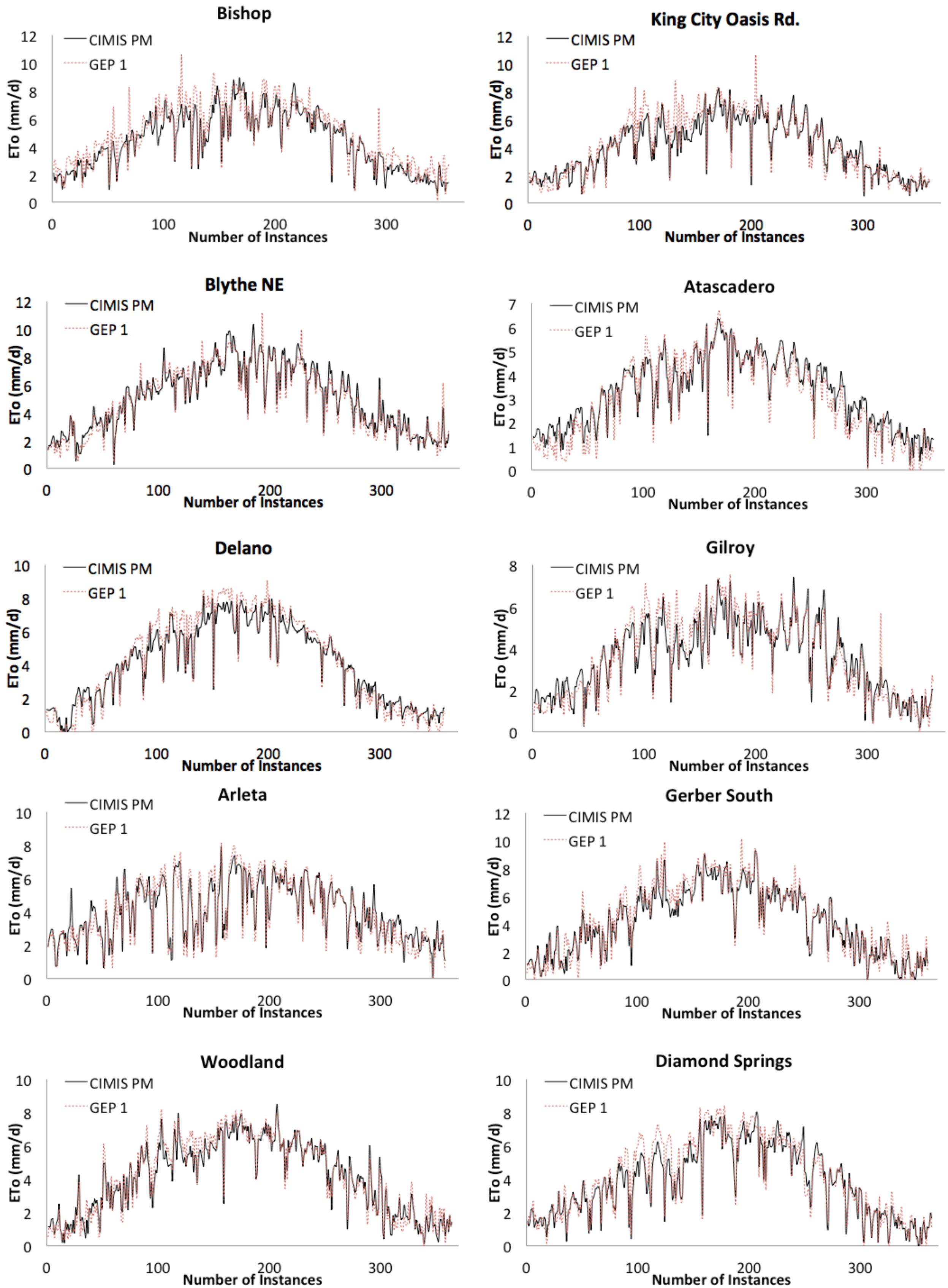


Figure 27. Estimated ET_0 values from GEP4 versus observations.



Bishop
CIMIS PM
GEP 1
ETo (mm/d)
Number of Instances
King City Oasis Rd.
CIMIS PM
GEP 1
ETo (mm/d)
Number of Instances
Blythe NE
CIMIS PM
GEP 1
ETo (mm/d)
Number of Instances
Atascadero
CIMIS PM
GEP 1
ETo (mm/d)
Number of Instances
Delano
CIMIS PM
GEP 1
ETo (mm/d)
Number of Instances
Gilroy
CIMIS PM
GEP 1
ETo (mm/d)
Number of Instances
Arleta
CIMIS PM
GEP 1
ETo (mm/d)
Number of Instances
Gerber South
CIMIS PM
GEP 1
ETo (mm/d)
Number of Instances
Woodland
CIMIS PM
GEP 1
ETo (mm/d)
Number of Instances
Diamond Springs
CIMIS PM
GEP 1
ETo (mm/d)
Number of Instances

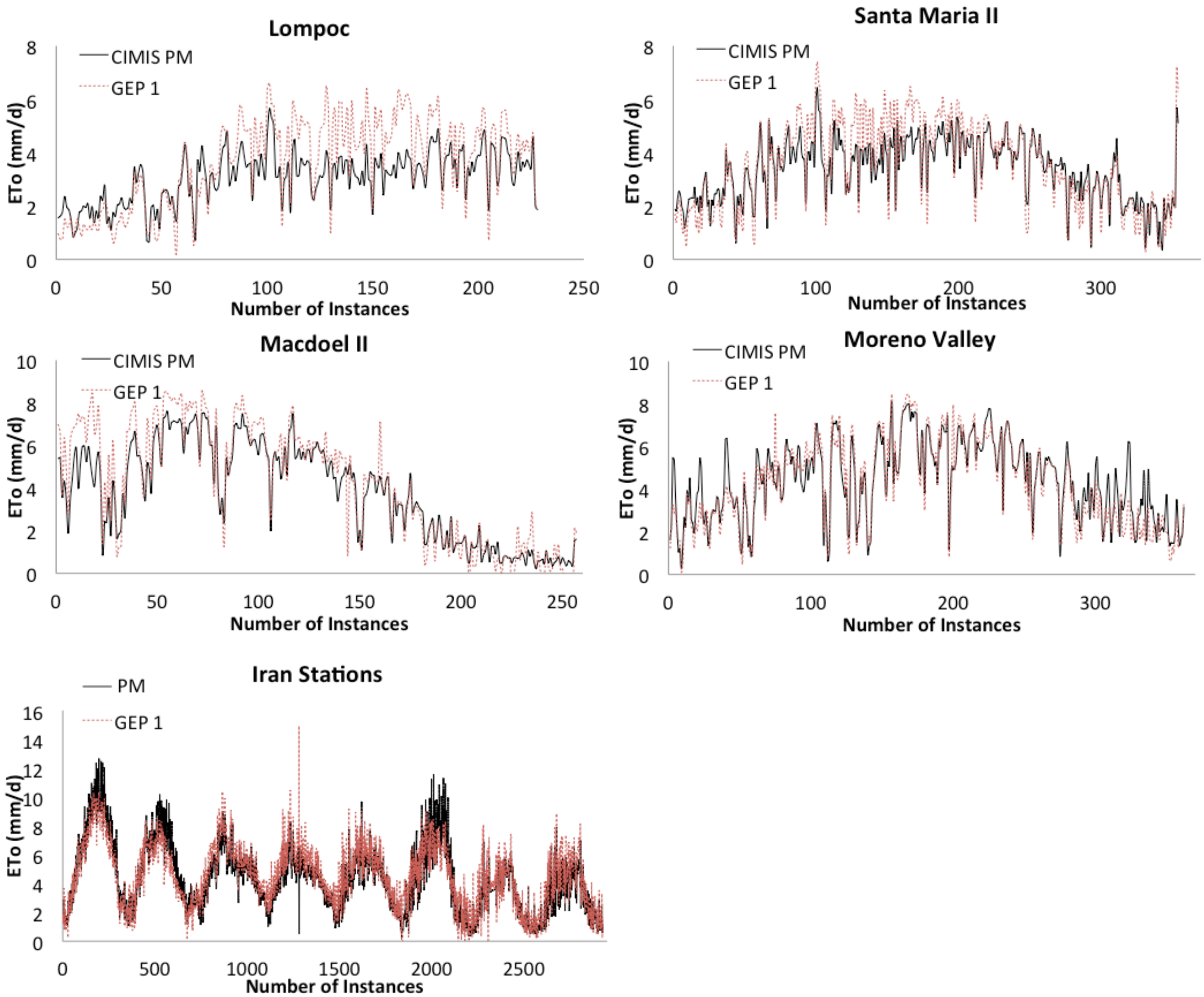
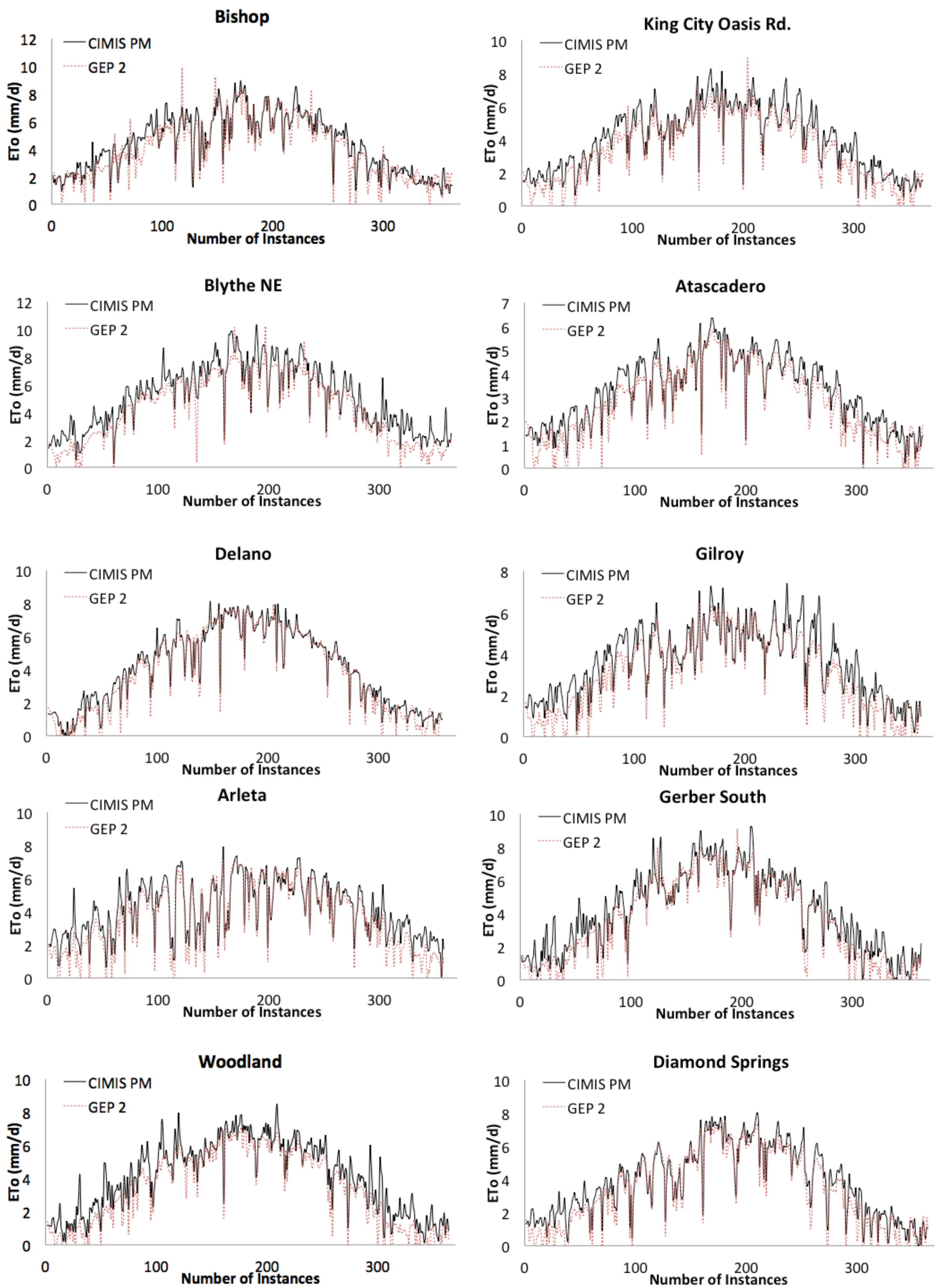


Figure 28. Time series of observed and estimated ET_0 values from GEP1.


Bishop
King City Oasis Rd.
Blythe NE
Atascadero
Delano
Gilroy
Arleta
Gerber South
Woodland
Diamond Springs
CIMIS PM
GEP 2
ETo (mm/d)
Number of Instances

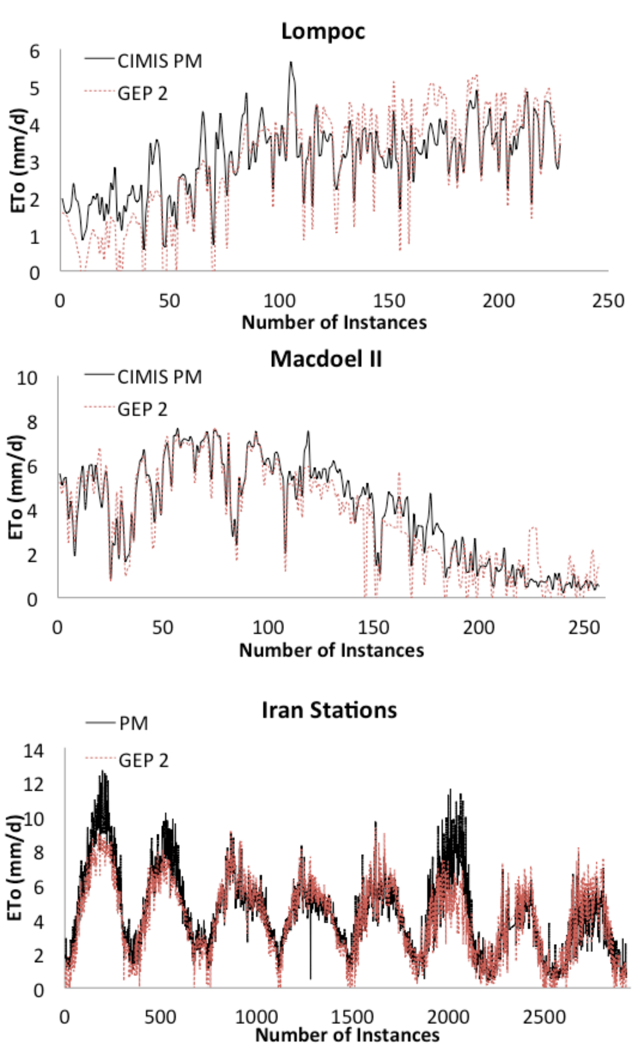
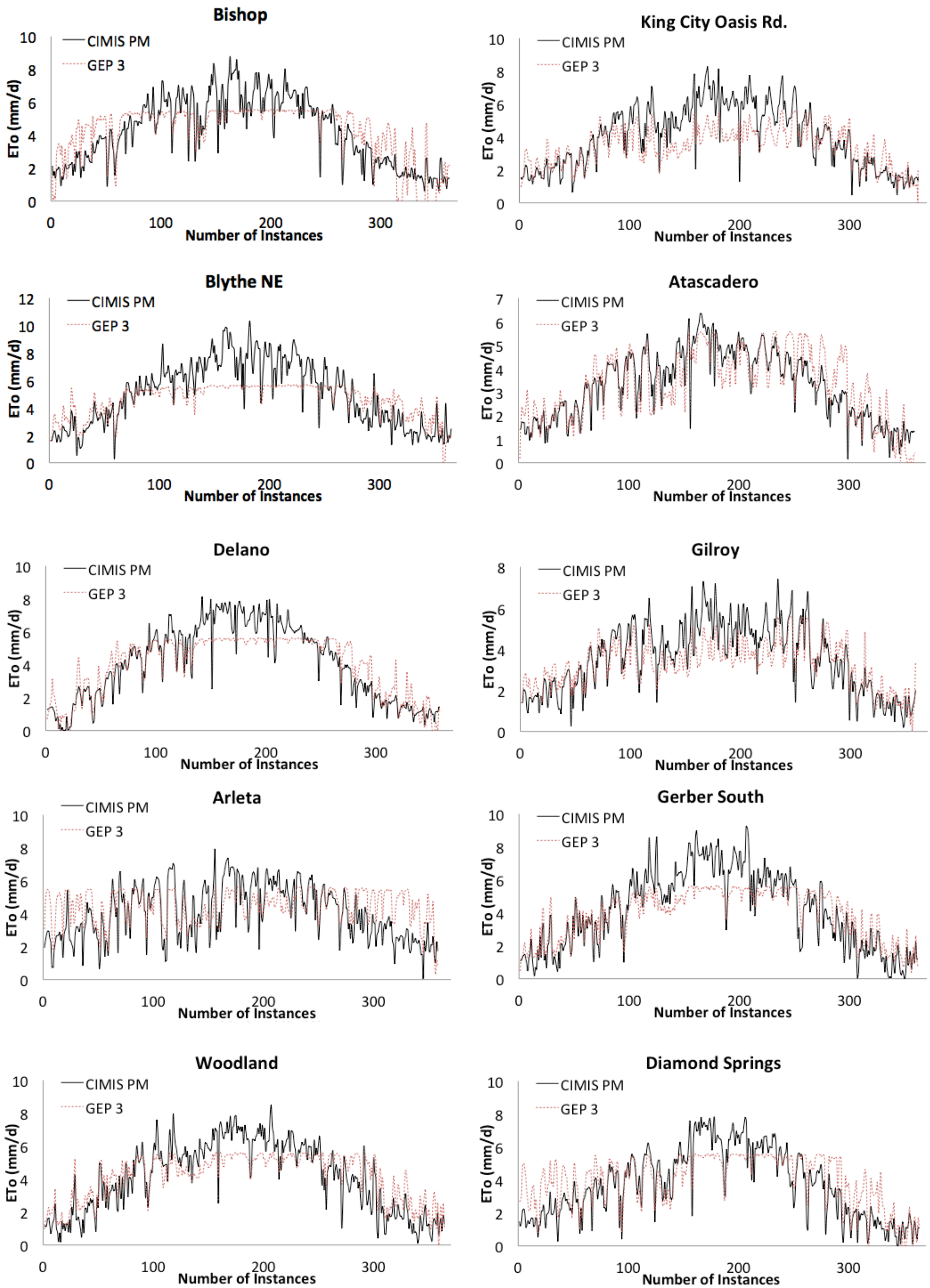


Figure 29. Time series of observed and estimated ET_0 values from GEP2.


Bishop
King City Oasis Rd.
Blythe NE
Atascadero
Delano
Gilroy
Arleta
Gerber South
Woodland
Diamond Springs
CIMIS PM
GEP 3
ETo (mm/d)
Number of Instances

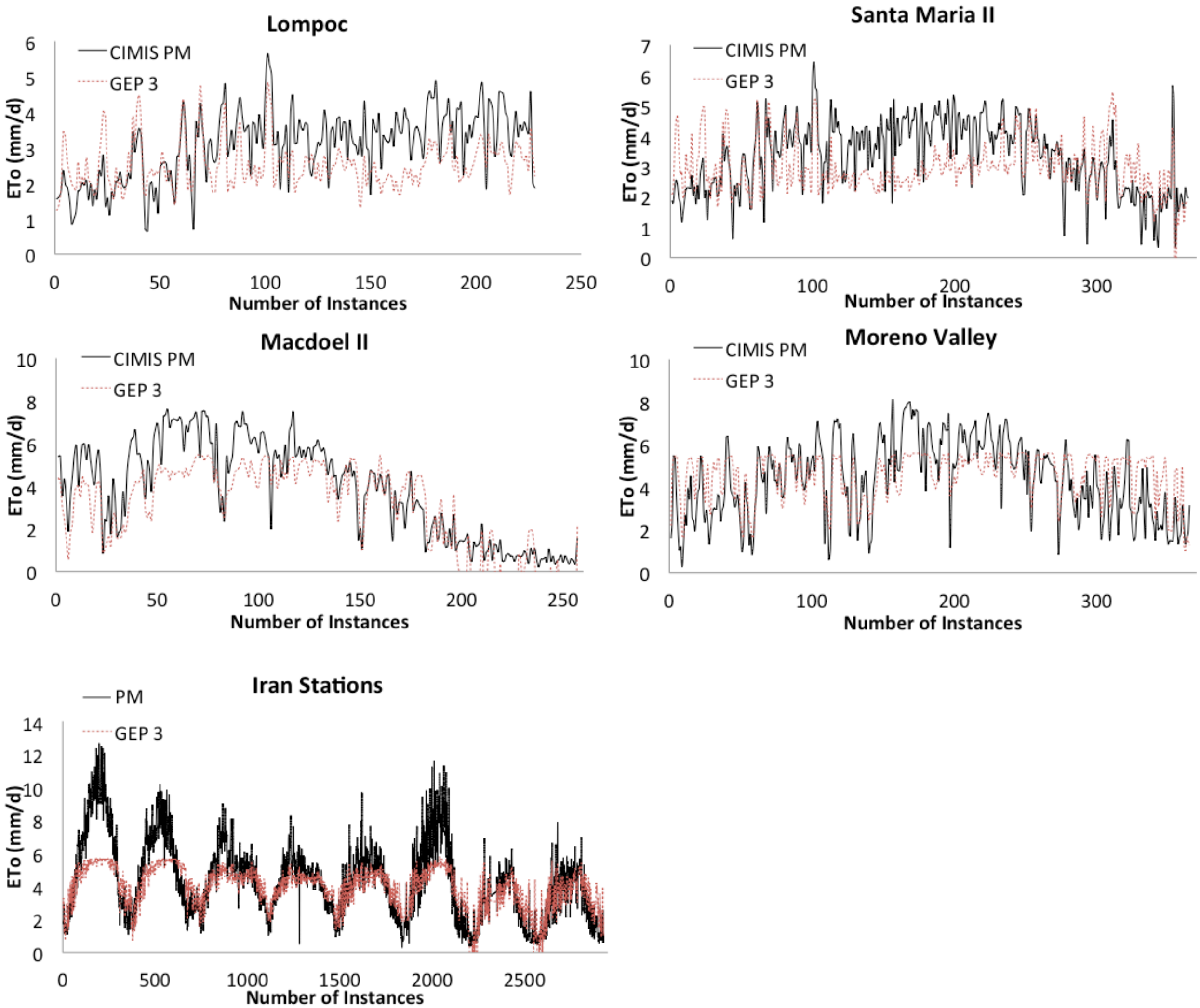
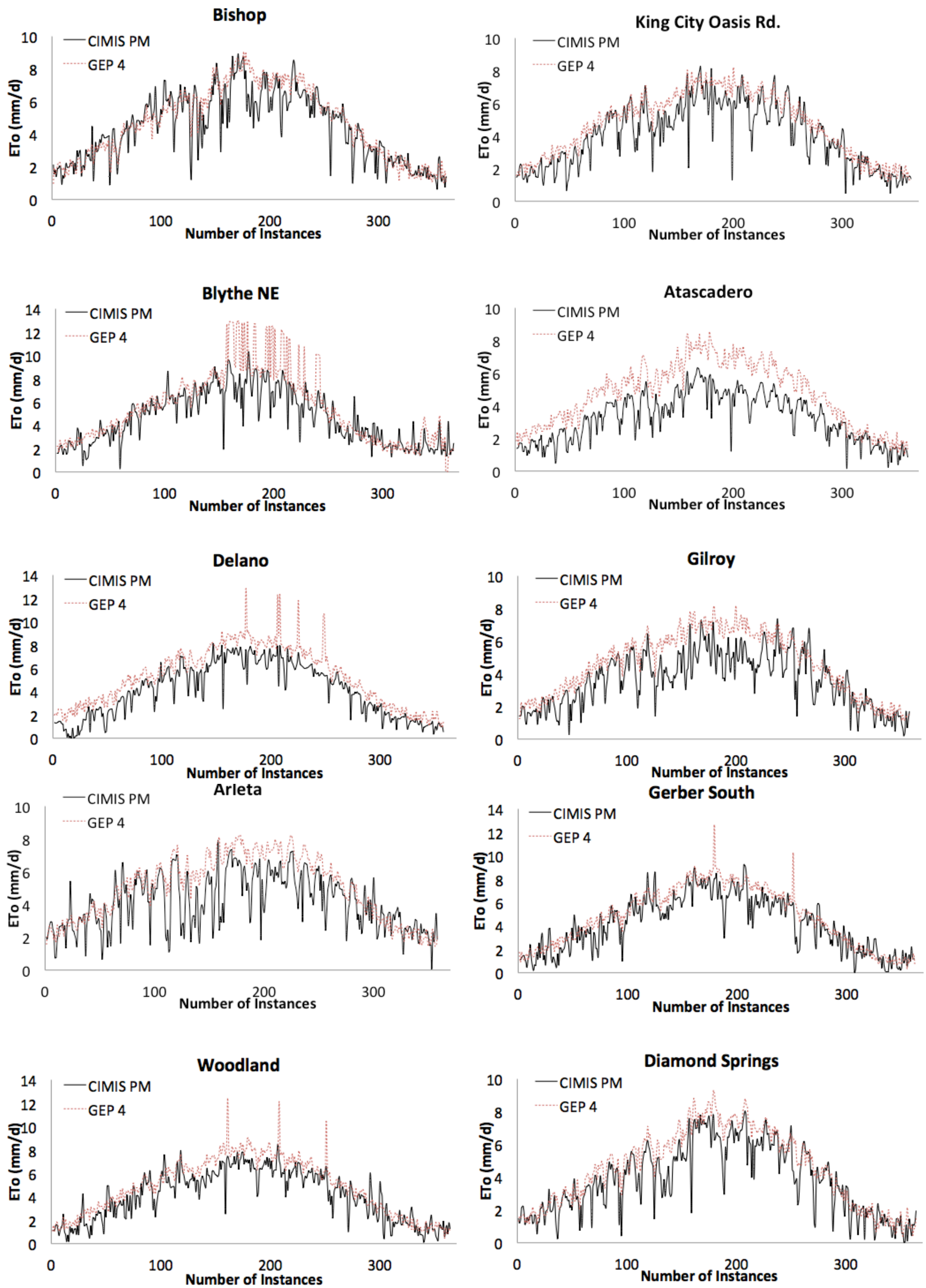


Figure 30. Time series of observed and estimated ET_0 values from GEP3.


Bishop
King City Oasis Rd.
Blythe NE
Atascadero
Delano
Gilroy
Arleta
Gerber South
Woodland
Diamond Springs
CIMIS PM
GEP 4
ETo (mm/d)
Number of Instances

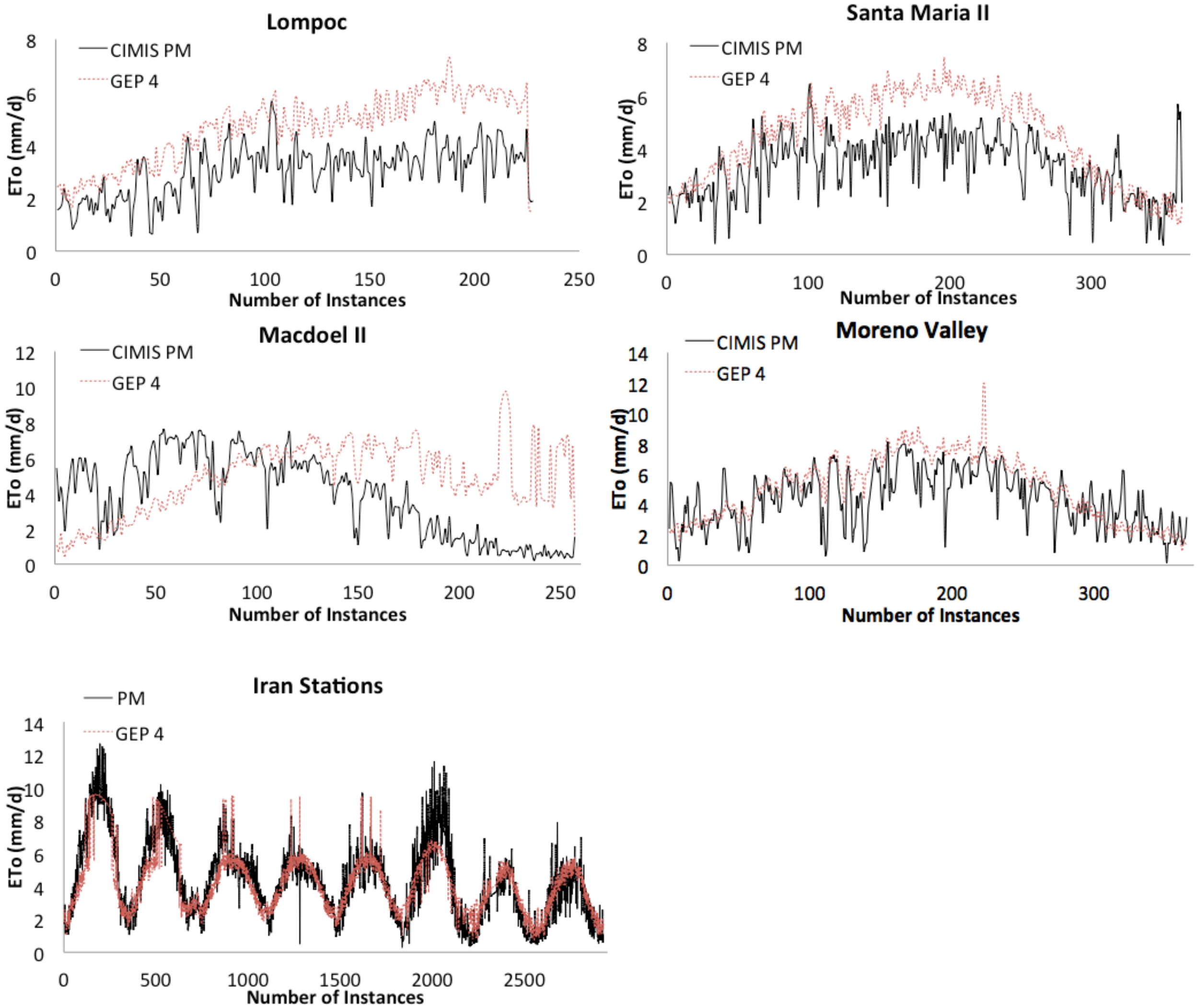


Figure 31. Time series of observed and estimated ET_0 values from GEP4.

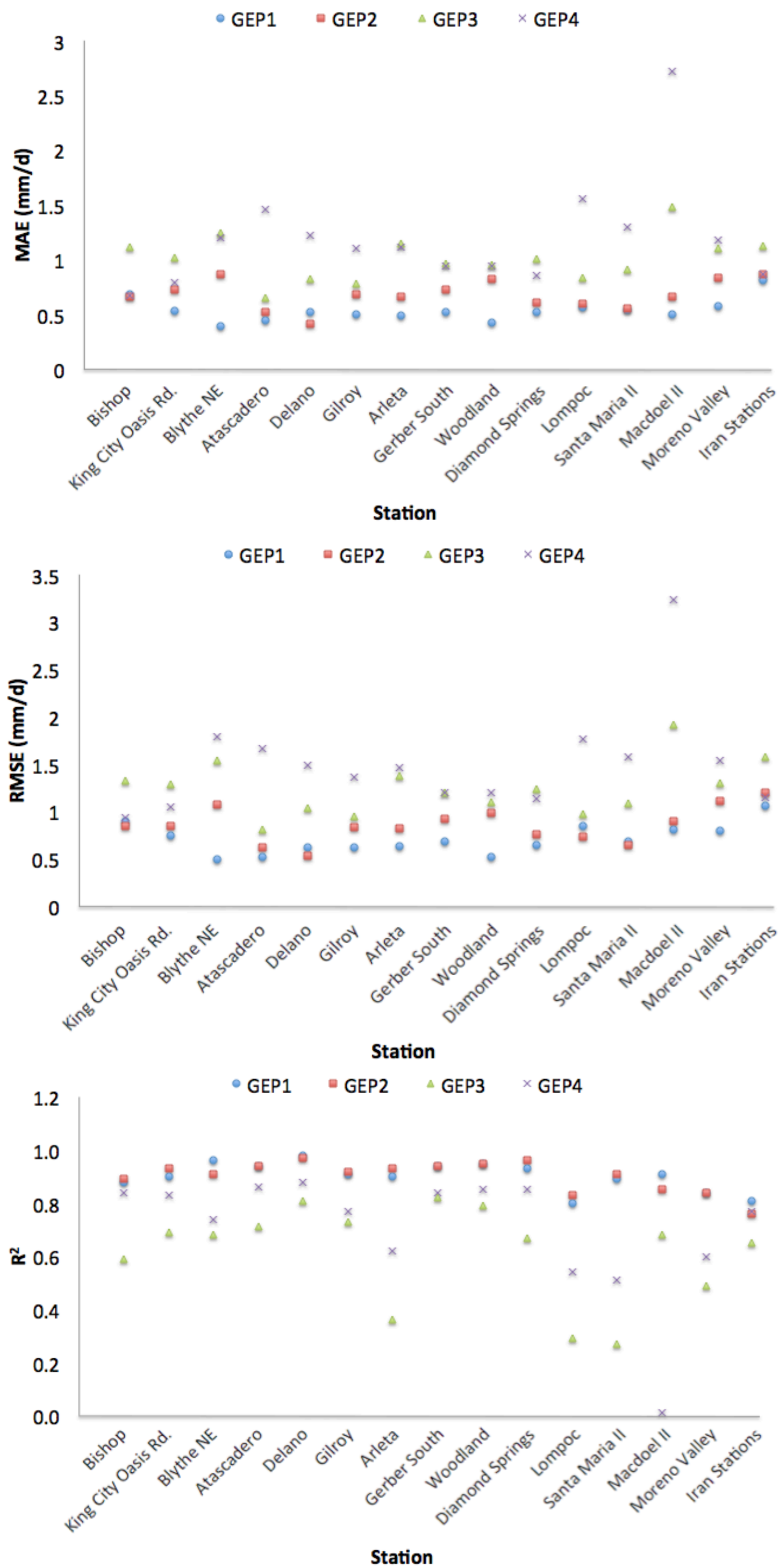


Figure 32. Comparing performance of GEP models in different stations.

Table 17. Statistical metrics of ET_0 estimates from GEP1 and GEP2.

Model	Station	MAE (mm/d)	RMSE (mm/d)	R^2
GEP1	Bishop	0.68	0.90	0.88
	King City Oasis Rd.	0.53	0.75	0.90
	Blythe NE	0.39	0.50	0.96
	Atascadero	0.44	0.52	0.94
	Delano	0.52	0.62	0.98
	Gilroy	0.50	0.62	0.91
	Arleta	0.49	0.63	0.90
	Gerber South	0.52	0.68	0.94
	Woodland	0.42	0.52	0.95
	Diamond Springs	0.52	0.65	0.93
	Lompoc	0.56	0.85	0.80
	Santa Maria II	0.54	0.68	0.89
	Macdoel II	0.50	0.81	0.91
	Moreno Valley	0.57	0.79	0.84
	All Iran Stations	0.81	1.06	0.81
GEP2	Bishop	0.66	0.84	0.89
	King City Oasis Rd.	0.72	0.84	0.93
	Blythe NE	0.86	1.07	0.91
	Atascadero	0.52	0.62	0.94
	Delano	0.41	0.53	0.97
	Gilroy	0.68	0.83	0.92
	Arleta	0.66	0.82	0.93
	Gerber South	0.73	0.92	0.94
	Woodland	0.82	0.98	0.95
	Diamond Springs	0.61	0.76	0.96
	Lompoc	0.60	0.73	0.83
	Santa Maria II	0.55	0.65	0.91
	Macdoel II	0.66	0.90	0.85
	Moreno Valley	0.83	1.11	0.84
	All Iran Stations	0.87	1.20	0.76

Table 18. Statistical metrics of ET_0 estimates from GEP3 and GEP4.

Model	Station	MAE (mm/d)	RMSE (mm/d)	R^2
GEP3	Bishop	1.11	1.33	0.59
	King City Oasis Rd.	1.02	1.28	0.69
	Blythe NE	1.24	1.54	0.68
	Atascadero	0.65	0.81	0.71
	Delano	0.82	1.03	0.81
	Gilroy	0.78	0.95	0.73
	Arleta	1.15	1.37	0.36
	Gerber South	0.96	1.20	0.82
	Woodland	0.95	1.10	0.79
	Diamond Springs	1.01	1.23	0.67
	Lompoc	0.83	0.97	0.29
	Santa Maria II	0.91	1.08	0.27
	Macdoel II	1.48	1.91	0.68
	Moreno Valley	1.10	1.30	0.49
	All Iran Stations	1.12	1.57	0.65
GEP4	Bishop	0.67	0.93	0.84
	King City Oasis Rd.	0.79	1.05	0.83
	Blythe NE	1.20	1.79	0.74
	Atascadero	1.46	1.66	0.86
	Delano	1.22	1.49	0.88
	Gilroy	1.10	1.36	0.77
	Arleta	1.11	1.46	0.62
	Gerber South	0.94	1.20	0.84
	Woodland	0.94	1.20	0.85
	Diamond Springs	0.85	1.13	0.85
	Lompoc	1.56	1.77	0.54
	Santa Maria II	1.30	1.58	0.51
	Macdoel II	2.72	3.23	0.01
	Moreno Valley	1.18	1.54	0.60
	All Iran Stations	0.87	1.15	0.77

4.4 Comparing Performance of MARS, M5MT, and GEP

The average performance of the MARS, M5MT, and GEP approaches for all stations are compared in Table 19 and Figure 33. According to Table 19, the GEP approach performed better than MARS and M5MT. Also, the MARS approach provided better results than M5MT. Average *MAE* and *RMSE* of ET_0 estimates from the MARS approach decreased respectively, by 9% and 18% compared to M5MT. Compared to GEP, MARS average *MAE* (*RMSE*) was 9% (5%) higher and R^2 was 29% lower. Also, M5MT average *MAE* (*RMSE*) was 19% (25%) higher and R^2 was 3% lower than GEP.

Comparing the average accuracies of the four different configurations, Table 19 shows configuration 1 to provide the best ET_0 estimates, followed by configuration 2. This was true for all approaches. These results demonstrated that the use of only two input parameters, T_{mean} and R_s (configuration 2), can provide results comparable to the use of four input parameters, T_{mean}, RH_{mean}, W_s, and R_s (configuration 1). The results also suggested that RH_{mean}, T_{min}, T_{max}, or R_a used with T_{mean} (configurations 3 and 4) did not have a significant role in predicting ET_0.

Of the four GEP models, GEP1 achieved the best accuracy, followed by GEP2, GEP3, and GEP4. GEP1 had the lowest *MAE* (0.53 mm/d) and *RMSE* (0.71 mm/d). *MAE* and *RMSE* of GEP2 increased by 28% and 20%, respectively compared to those of GEP1. *MAE* and *RMSE* of GEP3 increased by 91% and 75%, respectively, and R^2 decreased by 45% in comparison with GEP1. Furthermore, compared to GEP1, *MAE* and *RMSE* of GEP4 increased by 126% and 111%, respectively, and R^2 decreased by 29%.

Based on the previously mentioned results, GEP demonstrated superior accuracy in estimating ET_0 compared to MARS and M5MT. It was observed that for all approaches, the performance of configuration 1 surpassed the performances of configurations 2, 3, and 4. The results also indicated that configuration 2 can provide similar results to configuration 1, and yielded better results than configurations 3 and 4.

Table 19. Mean R^2, *MAE*, and *RMSE* of ET_0 estimates from MARS, M5MT and GEP models (tested with Iran and California stations).

Approach	Model	Tested with Iran and California		
		MAE (mm/d)	**RMSE (mm/d)**	**R^2**
MARS	MARS1	0.71	0.84	0.75
	MARS2	0.85	1.01	0.70
	MARS3	1.08	1.33	0.49
	MARS4	1.10	1.35	0.49
M5MT	M5MT1	0.76	0.94	0.86
	M5MT2	0.90	1.10	0.89
	M5MT3	1.13	1.56	0.63
	M5MT4	1.29	1.78	0.65
GEP	GEP1	0.53	0.71	0.90
	GEP2	0.68	0.85	0.90
	GEP3	1.01	1.24	0.62
	GEP4	1.20	1.50	0.70

Best Approach
Best Model Performance
Second Best Model Performance

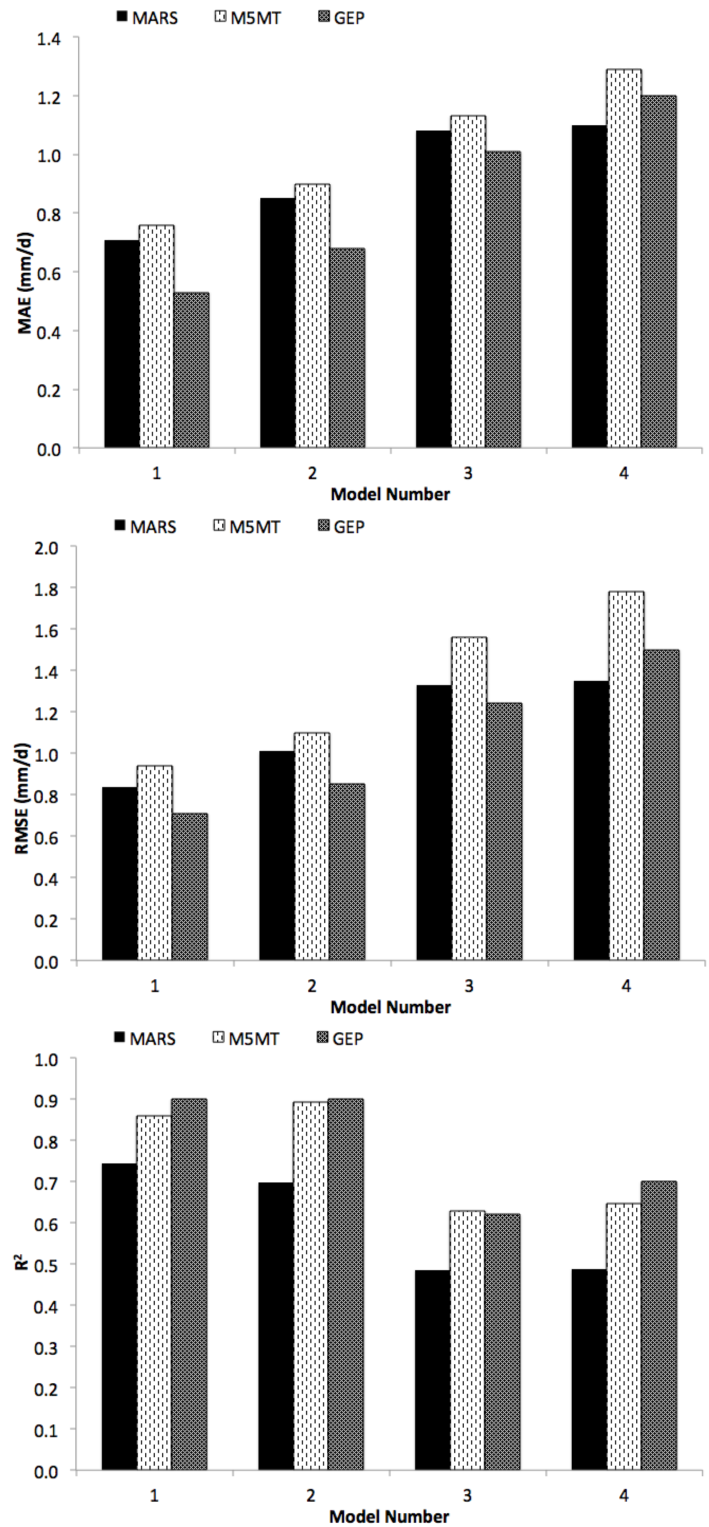


Figure 33. Performance of MARS, M5MT, and GEP models tested with Iran and California stations.

5. CONCLUSION

In this study, Multivariate Adaptive Regression Splines (MARS), M5 Model Tree (M5MT), and Gene Expression Programming (GEP) were used to estimate ET_0 from climatic data. Four different input configurations were used: daily mean air temperature (T_{mean}), daily mean wind speed (W_s), daily mean relative humidity (RH_{mean}), and solar radiation (R_s) [configuration 1]; daily mean air temperature (T_{mean}) and solar radiation (R_s) [configuration 2]; daily mean air temperature (T_{mean}) and daily mean relative humidity (RH_{mean}) [configuration 3]; daily maximum, minimum, and mean air temperature (T_{max}, T_{min}, and T_{mean}) and extraterrestrial radiation (R_a) [configuration 4]. MARS, M5MT and GEP were trained with climatic data from 8 weather stations in Iran for the years 2000-2007. These models were tested with the data from the same weather stations in Iran for the year 2008 and fourteen stations in California for the year 2015. Their performance was evaluated based on mean absolute error (*MAE*), root mean square error (*RMSE*), and coefficient of determination (R^2).

It was shown that configuration 1 yielded the most accurate results with the lowest *MAE* and *RMSE* in all approaches. In GEP, configuration 1 reduced *RMSE* by 20%, 75%, and 111% compared to configurations 2, 3, and 4, respectively. Similarly, in MARS, the reduction in *RMSE* was 20% 59%, and 61%, respectively. Finally, the decrease in *RMSE* was respectively, 17%, 66%, and 89% in M5MT. Therefore, it was concluded that configuration 1 consisted of variables that have the most significant effect on ET_0. Moreover, the accuracy of configuration 1 was not region dependent and it generated the best results in both Iran and California.

M5MT and GEP showed configuration 4 to be region dependent, performing better when tested locally. M5MT4 tested with Iran stations decreased *RMSE* by 125% compared to testing with California stations. Likewise, GEP4 tested with Iran stations reduced *RMSE* by 33%. Thus, configuration 4 is suggested when only air temperature is obtainable for a specific region.

This study demonstrated that modeling ET_0 is possible through the use of MARS, M5MT, and GEP. However, results showed GEP to be the best approach, followed by MARS. Compared to MARS, the average *MAE* and *RMSE* of ET_0 estimates from GEP were respectively 9% and 5% lower. Also, average *MAE* and *RMSE* of ET_0 estimates from GEP were respectively, 19% and 25% lower than those of M5MT. As a result, GEP is the recommended approach to most accurately model ET_0.

Summarizing, configuration 1 was not region dependent and provided the most accurate ET_0 estimates compared to configurations 2, 3, and 4. Configuration 4 was shown to be region dependent and would be useful in areas where only temperature data is available. GEP provided better results over MARS and M5MT, and is the suggested approach to model ET_0.

Future work should be directed to validate GEP, MARS and M5MT approaches in other regions as well. Also, their performance should be compared with commonly used equations such as Makkink, Romanenko and Hargreaves-Semani equations.

APPENDIX: M5MT Output Reference Guide

1. Starting from the left column of Table 10, "$R_s \leq 16.271$" is stated. This means R_s is the first node with a branching decision statement ≤ 16.271 (refer to Figure 1d for general configuration of model tree).
2. Following the symbol "|" down the same indentation, "$R_s > 16.271$" is stated. This represents the second branch of the node previously established (R_s) and its corresponding decision statement.
3. Following the symbol "|" down the same indentation, there are no other decision statements. This concludes the first tier of the model tree.
4. The second indentation from the top left of Table 10 starts the second tier. "$T_{mean} \leq 19.175$" and "$T_{mean} > 19.175$" are stated and T_{mean} is a node in the second tier with decision statements ≤ 19.175 and > 19.175. A data value from the testing dataset would reach this node if the statement "$R_s \leq 16.271$" is true.
5. Similarly, following the same indentation down the table, "$T_{mean} \leq 28.837$" and "$T_{mean} > 28.837$" are also in the second tier, which is another T_{mean} node with decision statements ≤ 28.837 and > 28.837. A data value from the testing dataset would reach this node if the statement "$R_s > 16.271$" is true.
6. Following the symbol "|" down the same indentation, there are no other decision statements. This concludes the second tier of the model tree.
7. The third indentation from the top left of Table 10 represents the nodes in the third tier and so on.

Note: Statements, such as "$T_{mean} \leq 8.55$: LM1", signify the node does not have any more branches below, and linear model 1 (LM1) can be used.

REFERENCES

Adamowski, J., H.F. Chan, S.O. Prasher, and V.N. Sharda (2012). Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for runoff forecasting in Himalayan micro-watersheds with limited data, *Journal of Hydroinformatics*, 14 (3), 731–744.

Allen, R., L.S. Pereira, D. Raes, and M. Smith (1998). Crop evapotranspiration. Guidelines for computing crop evapotranspiration. FAO Irrigation and Drainage Paper 56, Rome, Italy.

Atiaa, A. M., and H. B. Ghalib (2008). Rainfall-Runoff modeling by using M5 model trees technique: an example of Tigris catchment area in Baghdad, Middle of Iraq, *MARSH BULLETIN*, 3 (2), 125–135.

Azamathulla, H., and Z. Ahmad (2012). Gene-expression programming for transverse mixing coefficient, *Journal of Hydrology*, 434-435, 142-148.

Azamathulla, H., and R.D. Jarrett (2013). Use of Gene expression programming to estimate Manning's roughness coefficient for high gradient streams. *Water Resource Management,* 27, 715-729.

Bansal, P. and J. Salling. "Multivariate Adaptive Regression Splines (MARS)." Data Mining, UT ECE. 16 Feb. 2013.

Bhattacharya, B., and D.P. Solomatine (2005). Neural networks and M5 model trees in modelling water level–discharge relationship, *Neurocomputing,* 63, 381-396.

Emamgolizadeh, S., S.M. Bateni, D. Shahsavani, T. Ashrafi, and H. Ghorbani. Estimation of soil cation exchange capacity using Genetic Expression Programming (GEP) and Multivariate Adaptive Regression Splines (MARS). J. Hydrol. (2015), http://dx.doi.org/10.1016/j.jhydrol.2015.08.025

Etemad-Shahidi, A., and J. Mahjoobi (2009). Comparison between M5 model tree and neural networks for prediction of significant wave height in Lake Superior, *Ocean Engineering,* 36.15, 1175-1181.

"Evapotranspiration – The Water Cycle." Evapotranspiration – The Water Cycle, from USGS Water-Science School. 2 Dec. 2016. USGS. <http://water.usgs.gov/eduwatercycleevapotranspiration.html>.

Ferreira, C. (2001). Gene Expression Programming: A New Adaptive Algorithm for Solving Problems, *Complex Systems*, 13 (2), 87–129.

Frank, E., M.A. Hall, and I.H. Witten, 2016. The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

"Fresh Water Crisis." National Geographic: Environment. 2 May 2016. National Geographic. <https://environment.nationalgeographic.com>.

Friedman, J.H. "Multivariate Adaptive Regression Splines (with discussion)". Annals of Statistics, 1991. <http://www.salfordsystems.com/doc/MARS.pdf>

Garcia-Navarro, M.C., R.Y. Evans, and R.S. Montserrat (2004). Estimation of relative water use among ornamental landscape species, *Scientia Horticulturae*, 99 (2), 163–174.

Guven, A. and M. Gunal (2008). Genetic programming for prediction of local scour downstream of grade-control structures, *Journal of Irriation and Drainage Engineering*, 134 (2), 241–249.

Guven, A., A. Aytek, M.I. Yuce, and H. Aksoy (2008). Genetic programming-based empirical model for daily reference evapotranspiration estimation, *Clean Soil, Air, Water,* 36, 905–912.

Hargreaves, G.H. and Z.A. Samani (1985). Reference crop evapotranspiration from temperature, *American Society of Agricultural Engineers*, 1 (2), 96-99.

Hashmi, M.Z., A.Y. Shamseldin, B.W. Melville (2011). Statistical downscaling of watershed precipitation using gene expression programming (GEP), *Environmental Modeling Software*, 26, 1639-1646.

Jekabsons, G., ARESLab: Adaptive Regression Spines toolbox for Matlab/Octave, 2016, available at http://www.cs.rtu.lv/jekabsons/.

Kisi, O. (2015). Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree, *Journal of Hydrology*, 528, 312-320.

Kisi, O. and K.S. Parmar (2016). Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution, *Journal of Hydrology*, 534, 104–112.

Lui, J., A.J. B Zehnder, and H. Yang (2009). Global consumptive water use for crop production: the importance of green water and virtual water, *Water Resources Research*, 45, W05428, doi: 10.1029/2007WR006051.

Makkink, G.F. (1957). Testing the Penman formula by means of lysimeters, *Journal of the Institution of Water Engineering,* 11 (3), 277-288.

Oltean, M., and C. Grosan (2003). A Comparison of Several Linear Genetic Programming Techniques, *Complex Systems*, 14, 285–313.

Pal, M. and S. Deswal (2009). M5 model tree based modeling of reference evapotranspiration, *Hydrological Processes*, 23, 1437–1443.

Postel, S. "Climate Change Poses Existential Water Risks." National Geographic: Voices. 2 May 2016. National Geographic. <voices.nationalgeographic.com>.

Pour Ali Baba, A., J. Shiri, O. Kisi, A. Fakheri Fard, S. Kim, and R. Amini (2013). Estimating daily reference evapotranspiration using available and estimated climatic data by adaptive neuro-fuzzy inference system and artificial neural networks, *Hydrology Research*, 44 (1), 131–146.

Rahimikhoob, A. (2008). Comparative study of Hargreaves's and artificial neural network's methodologies in estimating reference evapotranspiration in a semiarid environment, *Irrigation Science*, 26 (3), 253–259.

Rahimikhoob, A., M. Asadi, and M. Mashal (2013). A Comparison Between Conventional and M5 Model Tree Methods for Converting Pan Evaporation to Reference Evapotranspiration for Semi-Arid Region, *Water Resour Manage*, 27, 4815–4826.

Romanenko, V.A. (1961). Computation of the autumn soil moisture using a universal relationship for a large area, In: Proceedings of Ukranian Hydrometeorological Research Institute, no. 3 Kiev.

Sattari, M.T., M. Pal, K. Yurkekl, and A. Unlukara (2013). M5 model trees and neural network based modeling of ET_0 in Ankara, Turkey, *Turkish Journal of Engineering and Environmental Sciences*, 37.2, 211–219.

Shahidian, S., R.P. Serralheiro, J. Serrano, J. Teixeira, N. Haie, and F. Santos. *Hargreaves and other reduced-set methods for calculating evapotranspiration.* InTech, 2012.

Sharda, V.N., R.M. Patel, S.O. Prasher, P.R. Ojasvi, and C. Prakash (2006). Modeling runoff from middle Himalayan watersheds employing artificial intelligence techniques, *Agricultural Water Management*, 83, 233 – 242.

Shiri, J. and O. Kisi (2011). Application of Artificial Intelligence to Estimate Daily Pan Evaporation Using Available and Estimated Climatic Data in the Khozestan Province (South Western Iran), *Journal of Irrigation and Drainage Engineering*, 137, 412–425.

Shiri, J., O. Kisi, G. Landeras, J.J. Lopez, A.H. Nazemi, and L. Stuyt (2011). Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain), *Journal of Hydrology*, doi: 10.1016/j.jhydrol.2011.11.004, in press.

Shiri, J., A. A. Sadraddini, A.H. Nazemi, O. Kisi, P. Marti, A.F. Fard, and G. Landeras (2013). Evaluation of different data management scenarios for estimating daily reference evapotranspiration, *Hydrology Research*, 44 (6), 1058–1070.

Shiri, J., A.A. Sadraddini, A.H. Nazemi, O. Kisi, G. Landeras, A.F. Fard, and P. Marti (2014). Generalizability of Gene Expression Programming-based approaches for estimating daily reference evapotranspiration in coastal stations of Iran, *Journal of Hydrology*, 508, 1–11.

Traore, S., Y. Wang, and T. Kerh (2010). Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone, *Agricultural Water Management*, 97, 707–714.

Verstraeten, W., F. Veroustraete, and J. Feyen (2008). Assessment of Evapotranspiration and Soil Moisture Content Across Different Scales of Observation, *Sensors*, 8, 70–117.

Wang, Y., and I.H. Witten (1997). Induction of model trees for predicting continuous lasses. In: Proceedings of the Poster Papers of the European Conference on Machine Learning. University of Economics, Faculty of Informatics and Statistics, Prague.

Witten, I.H., and E. Frank (2005). Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann: San Francisco.

Xu, T., S. Bateni, S. Margulis, L. Song, S. Liu, and Z. Xu (2016). Partitioning Evapotranspiration into Soil Evaporation and Canopy Transpiration via a Two-Source Variational Data Assimilation System, *Journal of Hydrometeor*, doi: 10.1175/JHM-D-15-01781, in press.

Yang, C.C., S.O. Prasher, R. Lacroix, and S.H. Kim (2004). Application of multivariate adaptive regression splines (MARS) to simulate soil temperature, *Transactions of the ASAE*, 47 (3), 881-887.

Zakaria, N.A., H. Azamathulla, C.K. Chang, A.A. Ghani (2010). Gene expression programming for total bed material load estimation – a case study, *Sci. Total. Environ.,* 408, 5078-5085.