Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Daily reference evapotranspiration prediction using empirical and data-driven approaches: A case study of Adana plain

Deniz Levent Koç[a*], Semin Topaloğlu Paksoy[b]

[a]*Department of Agricultural Structures and Irrigation, Agriculture Faculty, Çukurova University, 01250, Sarıçam, Adana, TÜRKIYE*
[b]*Department of Econometrics, Economics and Administrative Sciences Faculty, Çukurova University, 01250, Sarıçam, Adana, TÜRKIYE*

ARTICLE INFO
Research Article
Corresponding Author: Deniz Levent Koç, E-mail: leventk@cu.edu.tr
Received: 09 May 2024 / Revised: 12 September 2024 / Accepted: 30 September 2024 / Online: 14 January 2025

Cite this article
Koç D L, Topaloğlu Paksoy S (2025). Daily reference evapotranspiration prediction using empirical and data-driven approaches: A case study of Adana plain. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi),* 31(1):207-229. DOI: 10.15832/ankutbd.1481207

ABSTRACT
Precise determination of the reference evapotranspiration (ET_0) is vital to studying the hydrological cycle. In addition, it plays a significant role in properly managing and allocating water resources in agriculture. The objective of this research was to examine the effectiveness of five different data-driven techniques, including artificial neural networks "multilayer perceptron" (ANN), gene expression programming (GEP), random forest (RF), support vector machine "radial basis function" (SVM), and multiple linear regression (MLR) to model the daily ET_0. These methods were also compared with Hargreaves-Samani (HS), Oudin, Ritchie, Makkink (MAK), and Jensen Haise (JH) empirical models and their calibrated versions. The empirical models JH and MAK performed better than the models HS and Oudin after being calibrated by linear regression. All data-driven methods with four inputs were superior to the original and calibrated empirical models. Generally, data-driven models provided increased accuracy and enhanced generalization in predicting daily reference evapotranspiration compared to empirical models. The RF and ANN methods generally demonstrated better estimation accuracy than other data-driven methods. The performance of the RF and ANN models that utilized T_{max}, T_{min}, and R_s inputs, as well as those that incorporated T_{max}, T_{min}, R_s, and U_2 inputs, proved to be superior to their corresponding MLR-based and GEP-based models for predicting ET_0 in the Adana plain, which is characterized by a Mediterranean climate. Nevertheless, the GEP and MLR methods have the advantage of utilizing explicit algebraic equations, making them more convenient to apply, especially in the context of agricultural irrigation practices.

Keywords: Reference evapotranspiration, Data-driven approaches, Empirical models, Calibration, Adana plain

1. Introduction

Evapotranspiration is vital in maintaining the land's water and energy balance, significantly influencing water resource management, irrigation, environmental studies, and hydrological systems (Sabziparvar & Tabari 2010; Izadifar 2010). Various methods, including high-cost micrometeorological techniques, remote sensing, and water budget measurements, can be utilized to measure evapotranspiration. However, these approaches have inherent limitations, such as high expenses, maintenance requirements, and complexity (Liu & Zhu 2018; Niaghi et al. 2021).

Using mathematical models based on measured meteorological parameters is a cost-effective way to estimate reference evapotranspiration (ET_0) (Negm et al. 2018). Crop evapotranspiration (ET_c) can be calculated by multiplying the crop-specific coefficient known as K_c by ET_0 The FAO56 Penman-Monteith (PM) equation is a recognized method for estimating ET_0 (Allen et al. 1998) and can be used in diverse environments and climate conditions (Landeras et al. 2008; Shiri et al. 2012). However, the most crucial drawback of the FAO56-PM equation is that it requires air temperatures (T), solar radiation (R_s), relative humidity (RH), and wind speed (U) data. Many models have been developed to estimate ET_0 using reduced climate data due to incomplete meteorological data in some areas (Tabari et al. 2013a). Empirical methods often necessitate less data and fewer meteorological parameters, making them particularly applicable in agricultural settings for farmers and water managers. However, it is essential to note that the accuracy of empirical ET_0 methods can vary significantly due to differences in data requirements and theoretical assumptions (Dong et al. 2024). These methods are often tailored to specific geographical areas and local weather patterns, necessitating local calibration for optimal performance, unlike the standardized FAO56-PM equation (Gao et al. 2015; Pereira et al. 2015; Ferreira et al. 2019). HS, Oudin, JH, Ritchie, and MAK models are among the commonly used models (Tabari et al. 2012; Ferreira et al. 2019). FAO recommends that the HS equation be applied to calculate ET_0 in case of the absence of R_s, RH, and U data, which are crucial inputs of the FAO56-PM model (Allen et al. 1998). To date, several studies have been conducted aiming to calibrate the HS, Ritchie, MAK, Oudin, and JH models (Citakoglu et al. 2014; Almorox

& Grieser 2016; Feng et al. 2016; Feng et al. 2017; Shiri 2017; Cobaner et al. 2017; Ferreira et al. 2018; Banda et al. 2018; Gomariz Castillo et al. 2018; Srivastava et al., 2018; Ferreira et al. 2019; Khodayvandie et al. 2022). These papers highly recommended a regional or local calibration of the empirical equations used in the studies.

Another option for predicting ET_0 involves utilizing data-driven or soft computing methods such as ANN, GEP, MLR, SVM, and RF. There has been a significant emphasis on using these methods to estimate ET_0 in the past few years (Shiri 2018; Dou & Yang 2018; Gavili et al. 2018; Abrishami et al. 2019; Mattar & Alazba 2019; Reis et al. 2019; Yirga 2019; Ferreira et al. 2019; Wang et al. 2019; Mohsin & Lone, 2021; Jang et al. 2021; Niaghi et al. 2021; Achite et al. 2022; Dimitriadou & Nikolakopoulos 2022; Wang et al. 2022; Bayram & Çıtakoğlu 2023).

Data-driven approaches provide several advantages, including simplified development processes compared to physically based models. These models are independent of underlying boundary conditions, other assumptions, or initial forcing and can effectively operate at localized positions (Prasad et al. 2017). However, data-driven models require large amounts of high-quality data and can become overfitted to the training data. Developing and implementing data-driven models, particularly deep learning models, may necessitate substantial computational resources and time.

However, despite all the disadvantages, data-driven approaches have become increasingly popular for predicting ET_0, leading water resource and irrigation engineering experts to employ these methods in different applications. Huo et al. (2012) compared the performances of ANN models with those of MLR, Penman, PT and HS models in an arid area of northwest China. The results showed that ANN models estimated ET_0 more accurately than other models. Yassin et al. (2016) conducted a comparative analysis between ANN and GEP models to estimate potential evapotranspiration (ET_0) in arid environments. Based on the findings, it was observed that ANN-based models exhibited a marginally higher level of precision than GEP-based models. Antonopoulos & Antonopoulos (2017) used ANN, PT, MAK, HS and mass-transfer models to predict ET_0 in West Macedonia of northern Greece. In the study, the ANN model with four input variables estimated daily ET_0 most accurately. In Guangxi, which is situated in the southwest of China's Pearl River basin, Wang et al. (2019) utilized RF-based and GEP-based models to estimate ET_0. The findings concluded that the RF-based ET_0 models outperformed the GEP-based ones, albeit by a small margin. Mattar & Alazba (2019) modelled ET_0 from various combinations of climatic variables using GEP and MLR techniques in Egypt, where climatic conditions changed from warm to temperate. The study's findings indicated that GEP models outperformed MLR, HS, and MAK models. In the semiarid region of Brazil, Reis et al. (2019) conducted a study comparing the performance of empirical equations with those of ANN, ELM (extreme learning machine), and MLR models. The results showed that ANN, ELM, and MLR models had similar accuracy and were also more accurate than HS and calibrated HS models. Üneş et al. (2020) compared various models for modeling daily ET_0 and found that data-driven techniques, such as ANFIS and radial basis function SVM, outperformed empirical equations. ANFIS showed the highest correlation coefficient while radial basis function SVM had the minimum errors. Turc empirical formula was found to be better than other empirical equations. Chen et al. (2020) evaluated the performance of deep learning methods for ET_0 estimation in the Northeast Plain, China, using incomplete meteorological data. Three deep learning models, deep neural network (DNN), temporal convolution neural network (TCN), and long short-term memory neural network (LSTM), were developed and compared with classical machine learning models (SVM, RF) and empirical equations. The results showed that the deep learning models outperformed the empirical models, especially when temperature-based features were available. Kaya et al. (2021) found that support vector regression (SVR), multi-layer perceptron (MLP), and multiple linear regression (MLR) models outperformed commonly used empirical formulas in estimating daily ET_0 in the Košice City area of eastern Slovakia. The results from the Hargreaves-Samani equation closely aligned with the FAO 56-PM equation, demonstrating superior performance compared to other empirical equations. Niaghi et al. (2021) conducted a study to evaluate several machine learning techniques, comprising GEP, SVM, MLR, and RF, using three different input combinations to estimate ET_0 in the Red River Valley in the USA. The findings indicated that the RF model was the most effective approach for all input combinations among the four models considered. In Algeria's semiarid region, Achite et al. (2022) conducted a study to explore the potential of ANN and GEP models to estimate ET_0 by utilizing various combinations of climatic variables. The results showed that modelling ET_0 utilizing the ANN technique gave better estimates than the GEP models.

Sarıgöl & Katipoğlu (2024) evaluated the effectiveness of hybrid machine learning models for predicting monthly evaporation in the Southeast Anatolia Project Area. The research found that combining the gradient boosting machines (GBM) technique with signal decomposition methods generally provided more accurate evaporation estimations than using the GBM model alone.

As the reviewed literature shows, the studies of combined application of data-driven techniques to estimate ET_0 and applications of GEP, RF, MLR, and SVM approaches for modelling ET_0 are minimal, especially the RF method was rarely applied in hydrological and irrigation research. Also, in Türkiye, very few studies applied data-driven methods to estimate ET_0 and used the calibration of empirical models to estimate ET_0. As far as we know, only three studies have been carried out to adjust empirical models (Citakoglu et al. 2014; Çobaner et al. 2017; Gharehbaghi & Kaya 2022) and no studies have also addressed evaluating the GEP models to estimate daily ET_0 and developing the models' mathematical expressions in Türkiye. Other researchers preferred to compare the outcomes of the original empirical models solely. One of the unique features of this paper is the calibration of five widely used empirical models for estimating ET_0 and comparing them with data-driven models.

Forecasting reference evapotranspiration accurately is crucial for optimizing agricultural production and effective water resource management, which highlights the importance of addressing gaps mentioned above in the literature through this study.

2. Material and Methods

2.1. Study area and data used

The Adana Plain, situated in the Eastern Mediterranean region of Türkiye, was the focus of the study. Adana's climatic conditions are categorized under the Csa per the Köppen-Geiger classification system. This type of climate is known for its mild winters and extremely hot, dry summers. It is considered a typical Mediterranean climate (TSMS 2024). Adana Plain covers 27% of the province's territory. The basin's southern part is called Çukurova, and the northern part is called Anavarza (CSB 2023). Turkey's leading producer of citrus fruits, watermelons, soybeans, and peanuts is the Adana Plain (Kades 2019). Irrigation is necessary for crop production in Adana Plain due to inadequate precipitation levels and distribution during the growing season. The study's timeframe is from April through October, encompassing the growth season of the main crops in Adana Plain. This study used long-term daily climate data (2000-2021). The data was gathered from the Adana weather station in the Adana Plain (TSMS 2022) (Figure 1). According to the long-term climate data (1929-2022), the region experiences an annual precipitation of 668.8 mm, and approximately 50% of this precipitation falls during the winter months of December, January, and February. Temperatures are highest in July and August, with an average daily temperature of 28.2 and 28.7 °C. (TSMS 2023).

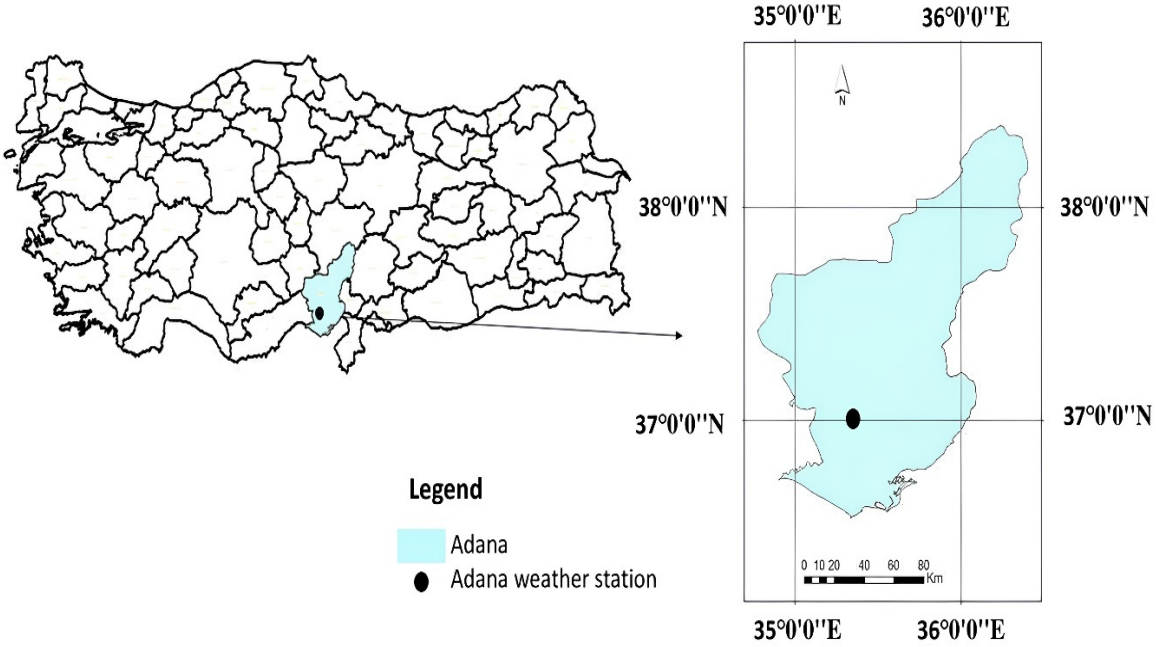


Figure 1- Position of the Adana climate station in Adana Plain in Türkiye (37 00' 14" N; 35 20' 39" E; altitude: 24 m)

The statistics regarding the data used can be found in Table 1. The average highest temperature (T_{max}) and lowest temperature (T_{min}) during the training period are 30.8 and 19.5 ºC, respectively. The mean relative humidity (RH) is 64.7%, solar radiation (R_s) is 19.6 MJ m^{-2} d^{-1} and wind speed at a two-meter height (U_2) is 0.92 m s^{-1} during the training period. The RH has a skewness coefficient of -0.94, while the U_2 has a skewness coefficient 0.41. The U_2 has a kurtosis coefficient 1.34, and T_{min} kurtosis coefficient of -0.56. In the validation period, the average values for T_{max}, T_{min}, RH, and R_s are 31.5 ºC, 20.6 ºC, 63.8 %, and 20.8 MJ m^{-2} d^{-1}, respectively. However, the average U_2 during the validation period is notably higher at 1.33 m s^{-1}. The coefficient of variation for U_2 is higher in the validation period, indicating more variability in wind speed during this period. The coefficients of variation of other climate parameters are fairly similar. The skewness coefficients of U_2 differ, while the other parameters are similar in both the training and validation periods. The kurtosis coefficients of T_{max} and U_2 differ, while the other parameters are similar in both the training and validation periods. The mean ET_0 is slightly higher in the validation period. However, the coefficient of variation is the same in both periods, indicating similar variability in ET_0 in both periods.

Table 1- Key statistical information of the dataset utilized in the study during the training and validation periods

Period	*Statistic*	T_{max} (°C)	T_{min} (°C)	*RH* (%)	R_s (MJ m^{-2} d^{-1})	U_2 (m s^{-1})	ET_0 (mm d^{-1})
Training	Maximum	42.1	29.8	94.5	30.4	3.07	8.9
	Minimum	15.0	3.5	16.0	0.8	0.10	0.9
	Mean	30.8	19.5	64.7	19.6	0.92	4.2
	Standard Deviation	4.64	4.80	11.23	5.11	0.39	1.17
	Skewness	-0.67	-0.42	-0.94	-0.60	0.41	-0.19
	Kurtosis	-0.09	-0.56	1.02	0.23	1.34	-0.38
	Coefficient of Variation	0.15	0.25	0.17	0.26	0.43	0.28
Validation	Maximum	45.1	29.8	87.0	30.5	3.96	10.5
	Minimum	16.5	6.6	24.0	0.3	0.45	1.1
	Mean	31.5	20.6	63.8	20.8	1.33	4.7
	Standard Deviation	4.57	4.65	10.21	5.58	0.44	1.31
	Skewness	-0.72	-0.52	-1.08	-0.84	1.58	-0.21
	Kurtosis	0.43	-0.53	0.94	0.44	4.99	-0.31
	Coefficient of Variation	0.15	0.23	0.16	0.27	0.33	0.28

Notes: ET_0 represents the ET_0 values estimated by the FAO56-PM equation

2.2. Empirical models for ET_0 prediction

This study employed the FAO56-PM ET_0 model as the benchmark for assessing the empirical and data-driven models, a common process in literature (Equation 1).

$$ET_0 = \frac{0.408\text{SVPC}(R_n - G) + \gamma \frac{900}{T + 273} U_2(e_{sat} - e_{act})}{\text{SVPC} + {}_{\varphi}(1 + 0.34U_2)} \quad (1)$$

ET_0 refers to the reference evapotranspiration in mm d^{-1}, R_n represents the net radiation in MJ m^{-2} d^{-1}, G denotes the soil heat flux density in MJ m^{-2} d^{-1}, T signifies the mean daily air temperature at the height of 2 meters in °C, U_2 represents the wind speed at a height of 2 meters in m s^{-1}, e_{sat} refers to the saturation vapour pressure in kPa, e_{act} denotes the actual vapour pressure in kPa, SVPC represents the slope vapour pressure curve in kPa $°C^{-1}$, and φ signifies the psychrometric constant in kPa $°C^{-1}$.

All parameters' daily values were calculated using the equations outlined in the FAO Irrigation and Drainage Paper 56 written by Allen et al. (1998).

Five empirical models were employed in this study, and concise descriptions of the methodologies utilized are presented below. The equations corresponding to each model can be found in Table 2.

2.2.1. Hargreaves-Samani model

Hargreaves and Samani (1985) introduced an equation for estimating ET_0 based on daily or mean values of maximum and minimum temperature. The equation was formulated using eight years of daily lysimeter data collected in Davis, California, and subsequently applied to compute ET_0 values for various locations.

2.2.2. Oudin model

In 2005, Oudin et al. suggested that the Penman-Monteith equation may not be the best choice for use in rainfall-runoff models. Instead, they put forward the Oudin method as an alternative.

2.2.3. Ritchie model

Due to a lack of meteorological data, Jones & Ritchie (1990) devised a simpler formula that correlates ET_0 with only air temperature, known as the Ritchie equation. The Ritchie equation finds extensive application in plant growth models and in research on managing agricultural water.

2.2.4. Makkink model

Makkink (1957) developed the MAK method based on the Penman-Monteith equation, and its validation was conducted using data from a lysimeter study on short grass in the Netherlands.

2.2.5. Jensen-Haise model

Jensen & Haise (1963) introduced the JH method, which utilizes global solar radiation and air temperature to develop irrigation plans based on comprehensive field data in arid and semi-arid conditions in the USA.

Allen et al. (1998) recommended calibrating empirical models locally by determining regression coefficients (a, b) using the FAO56-PM model as follows (Equation 2).

$$ET_0(PM) = a + bET_{0_{model}} \quad (2)$$

ET_0 (PM): the reference evapotranspiration computed by the FAO56-PM, ET_{0model}: the reference evapotranspiration estimated by other applied models, and a and b: regression coefficients.

Table 2- Empirical models to estimate ET_0

Model	***Reference***	***Meteorological inputs***	***Formula***
Hargreaves-Samani (HS)	Hargreaves & Samani (1985)	T_{max}, T_{min}	$ET_0 = 0.0023Ra(T_{max} - T_{min})^{0.5}(T + 17.8)$
Oudin	Oudin et al. (2005)	T_{max}, T_{min}	$ET_0 = \frac{R_a}{\lambda}(\frac{(T+5)}{100}\ If\ T + 5 > 0;\ ETo = 0\ otherwise$
Ritchie	Jones & Ritchie (1990)	T_{max}, T_{min}, R_s	$ET_0 = \alpha_1[3.87 \times 10^{-3} \times R_s(0.6T_{max} + 0.4T_{min} + 29)]$ $\alpha_1 = 1.1\ when\ 5°C < T_{max} < 35°C$ $\alpha_1 = 1.1 + 0.05(T_{max} - 35)\ when\ T_{max} > 35°C$ $\alpha_1 = 0.1exp[0.18(T_{max} + 20)]\ when\ T_{max} < 5°C$
Makkink (MAK)	Makkink (1957)	T_{max}, T_{min}, R_s	$ET_0 = 0.61\frac{SVPC}{SVPC + \gamma}\frac{R_s}{\lambda} - 0.12$
Jensen-Haise (JH)	Jensen & Haise (1963)	T_{max}, T_{min}, R_s	$ET_0 = 0.408R_s(0.0252T + 0.078)$

Notes: T= mean daily air temperature (°C), R_a = extraterrestrial radiation (MJ m^{-2} d^{-1}), λ = the latent heat flux (MJ kg^{-1}), α_1= coefficient, SVPC =slope vapour pressure curve (kPa $°C^{-1}$), γ = psychrometric constant (0.0672 kPa K^{-1})

2.3. Data-driven methods for ET_0 prediction

As per the NFLT (No Free Lunch Theorem), all optimization methods perform equally well on average. Hence, there is no agreement on whether a machine learning method can always provide better results compared to the others. Several aspects, like the size and structure of the dataset employed in the study, play a crucial role. Therefore, multiple methods are utilized and compared to obtain the most valuable or best prediction (Sterkenburg & Grünwald; 2021; Goldblum et al. 2023). The current study examines GEP, RF, ANN, MLR, and SVM models to predict the daily ET_0 using observed daily meteorological data set (T_{max}, T_{min}, U_2, and R_s) and FAO56-PM-targeted ET_0 values. For this purpose, as common in the literature (Irmak et al. 2003; Noi et al. 2017), this study employed 70% of the dataset (from April 1, 2000, to October 31, 2014) for training and calibration and the remaining 30% (from April 1, 2015, to October 31, 2021) for validating the data-driven and empirical models.

A correlation matrix was constructed to establish the connection between FAO56-PM ET_0 and several climate parameters (Figure 2). RH was found to have a very weak negative correlation (R = -0.165) with FAO56-PM ET_0. Therefore, RH was excluded when deciding on input combinations for the models. In this study, a total of four input combinations were devised to assess the effectiveness of data-driven models. Two of these combinations were the same as the input combination of the empirical equations utilized in the study: 1) T_{max} and T_{min} and 2) T_{max}, T_{min}, and R_s. The remaining combinations were defined as 3) T_{max}, T_{min}, and U_2 and 4) T_{max}, T_{min}, R_s, and U_2. Figure 3 provides a summary of the study's workflow.

Figure 2- Correlation between FAO56-PM ET_0 and various climate parameters

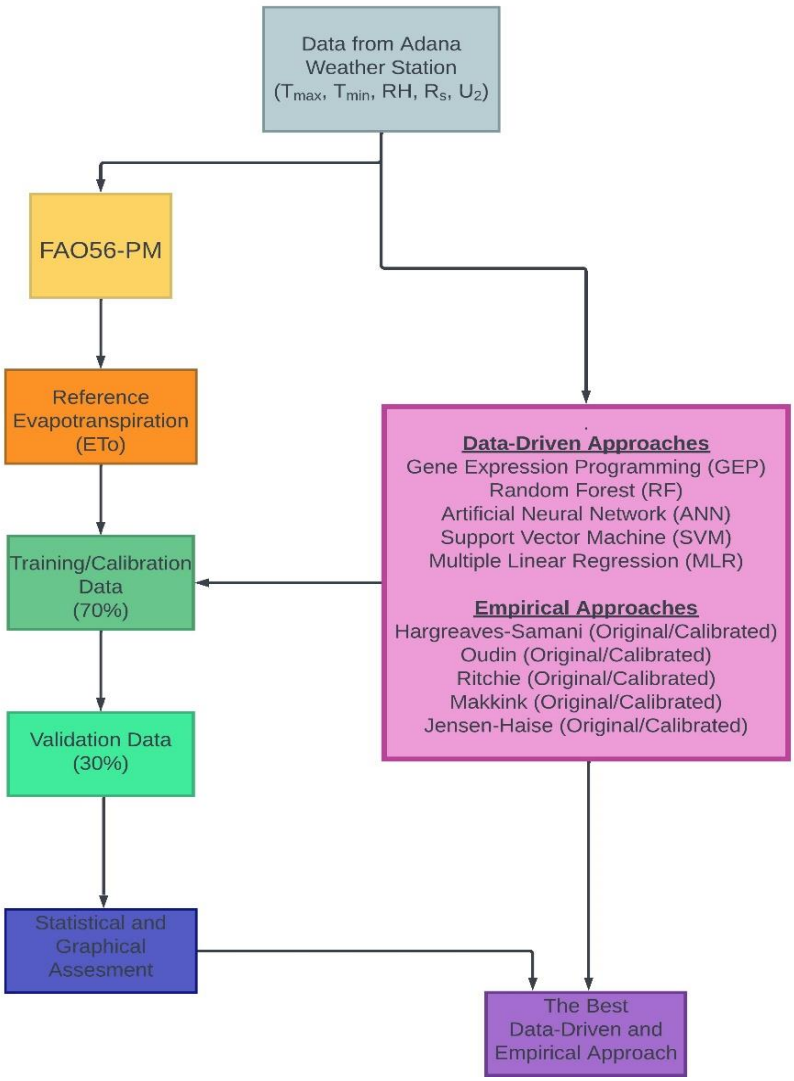


Figure 3- Flowchart outlining the research methodology used in the study

2.3.1. Gene expression programming (GEP)

GEP, invented by Ferreira in 2001, is a branch of evolutionary algorithms that models dynamic and non-linear processes. It is an advanced form of genetic algorithms and programming (Mehdizadeh et al. 2017). The sensitivity of GEP to the number of inputs is lower than the impact of the information content present in the data (Shiri, 2017). The GEP algorithm defines the solution to a problem using chromosomes that contain one or multiple genes. These genes can be combined to create chromosomes, which are then linked through a linking function (Traore & Guven 2013). Algebraic expressions for non-linear problems can be automatically generated by GEP. Moreover, GEP models have the ability to explicitly establish the connection between dependent and independent variables.

The first step in designing a GEP algorithm is to define the fitness function. The next step is to define the terminals and functions to be used. The third step is to determine the structure of chromosomes, including the number of generations, length, and number of genes. The fourth step is to determine the linking function. The fifth step is to determine the characteristics of the operators. Finally, the algorithm can be implemented. After running the program with different input combinations, the GEP model was developed once the model's accuracy stopped significantly improving (Wang et al. 2019).

2.3.2. Artificial neural networks (ANN)

ANNs can be described as mathematical models resembling biological neural networks. ANNs can learn from examples, identify patterns within the data, and adapt their solutions over time, making them efficient at processing large amounts of information (Jain et al. 2008). An ANN's internal architecture resembles a natural brain's structure, with several layers of fully interconnected nodes or neurons. A neural network is defined by its architecture, training algorithm, and activation function. The most common neural network architecture consists of three layers: the input layer for introducing data, one or more hidden layers for processing information, and the output layer for obtaining results. The type of ANN that we are referring to is known as a multilayer perceptron (MLP) (Fausett et al. 2006). The Backpropagation (BP) algorithm is used by the multilayer perceptron (MLP), which is the most widely used, persuasive, and effective neural network architecture (Choi et al. 2018). Kumar et al. (2011) provide further information about ANN. MLP neural networks with various combinations of climatic inputs and hidden layers were evaluated in this research to determine their usefulness in estimating ET_0.

2.3.3. Multiple linear regression (MLR)

Despite significant developments in data-driven modelling, using MLR for various modelling and model comparison purposes remains in demand (Izadifar 2010). MLR, a statistical modelling technique, uses multiple explanatory variables to anticipate the result of a response variable. MLR employs multiple explanatory variables, typically two or more, to approximate the result of a response variable by using a linear equation to establish a fitting linear model. The MLR model has a general form that can be expressed as follows (Equation 3):

$$Y = a_0 + a_1X_1 + a_2X_2 + \ldots + a_nX_n \quad (3)$$

Where Y: the predicted value of the dependent variable; $X_1,\ldots X_n$: independent variables; a_0: unknown intercept; $a_1,\ldots a_n$: estimated regression coefficients of the function.

2.3.4. Support vector machine (SVM)

The support vector machine (SVM) is a relatively new soft learning algorithm adopted for multiple applications in fields such as soft computing, hydrology, and environmental studies (Gocić et al. 2015). SVM is a well-known method in the field of machine learning that is based on the principles of classification and regression analysis theory. Its origin can be traced back to its developer, Vapnik (1995). Support vector regression (SVR) is commonly employed to elucidate the concept of regression in the SVM algorithm.

SVM uses the equation f(x) in regression analysis to establish the connection between a set of independent variables (x) and a dependent variable (y). This equation defines how the independent variables relate to the dependent variable (Equation 4).

$$f(x) = \hat{w}\cdot\phi(x) + þ \quad (4)$$

The weight vector is represented as 'ŵ'. '(x)' is a non-linear function that transforms the input space vector' x' into a high-dimensional feature space. The bias term is denoted as 'þ'.

SVM employs kernel functions such as linear (Lin), polynomial (Poly), and radial basis function (RBF) to transform input data into a feature space of high dimensionality. This study used the radial basis (RBF) as the kernel function since it is highly recommended in the literature (Tabari et al. 2012; Seifi & Riahi 2020). To achieve satisfactory performance, it is crucial to

appropriately set parameters C, ε, and γ when using the RBF kernel to train an SVM model. For this study, the parameters were chosen through a process of trial and error.

2.3.5. Random forest (RF)

The Random Forest (RF) model has recently become increasingly popular as an ensemble learning technique for classification and regression tasks. Its various advantages make it a preferred choice in classification and regression problems. These benefits include preventing overfitting, providing satisfactory performance, and enabling personalized parameter selection (Feng et al. 2017). Breiman (2001) first introduced random forest (RF), a machine-learning method that utilizes several decision trees to form an ensemble model. Ensemble learning is a technique that seeks to enhance the overall ability of machine-learning models to generalize by creating multiple base learners or combining multiple trees in their structure (Samadianfard et al. 2022). The RF method is one of the most effective and practical approaches to generating rules. This method is based on decision tree algorithms such as Classification and Regression Tree (CART). Compared to other decision tree ensembles, RF is known to be a more robust approach (Cutler et al. 2012). The Random Forest (RF) method can define the appropriate predictor without requiring the data to be re-scaled like other techniques. Conversely, regression trees that follow the traditional approach tend to over-fit on the training data set, which results in poor performance. Nevertheless, the RF method utilizes the characteristics of randomness to conquer this issue (Shirzad & Safari 2019).

Breiaman (2001) and Cutler et al. (2012) contain further information on RF. The current study utilized RF as the regression model for estimating ET_0. Parameter optimization was carried out with a focus on the number of trees to achieve the study's lowest possible error level.

2.4. Software and pseudo codes

The models in this study were developed using GeneXProTools 5.0 and the Orange 3.35 software suite. Due to limitations in the existing programs, the visual components were developed using the Python programming language (Phyton 3.10.12). The study utilized the following libraries: Matplotlib 3.7.1, Numpy 1.26.4, Pandas 2.1.4, and Seaborn 0.13.1. Applied model parameters and pseudo codes are provided in Appendix 1. Regarding the benefits of softwares, the following points can be articulated. Orange software is a user-friendly platform for developing data analysis and machine learning models. It has a drag-and-drop interface, making model building intuitive and not reliant on extensive technical expertise (Orange Data Mining 2024). GeneXProTools is a robust software application for modeling and analyzing data using genetic programming. It excels in automated exploration and optimization of complex mathematical models and is well-suited for data mining, modeling, optimization, and classification (Ferreira 2001).

2.5. Assessment of model performance

Table 3 demonstrates the use of multiple statistical indicators to evaluate the precision of the model's predictions (Willmott 1981; Karunanithi et al. 1994; Jacovides & Kontoyiannis 1995; Landeras et al. 2008). The current investigation assessed the performance of the models by employing various metrics such as RMSE, MAE, RaRMSE, RaMAE, and R, and applying criteria recommended by Corzo & Solomatine (2007) based on the RE values. As per criteria, an RE value of 15% or less indicates a small error, whereas an RE value between 15% and 35% indicates a moderate error, and finally, an RE value greater than or equal to 35% is considered a large error.

In the study, the conformance between the ET_0 derived from the FAO56-PM reference method and the modeled ET_0 values was assessed using a Taylor diagram, employing the Pearson correlation coefficient (R), RMSE, and standard deviation statistics. This method facilitated the identification of the most realistic model. Moreover, the research employed box plots to visually represent the distribution of numerical data. These box plots illustrate the central tendency (median), variability (interquartile range), and potential skewness through quartiles. Additionally, they help in identifying outliers and comparing the overall spread of different data sets (Mcgill et al. 1978; Taylor 2001).

Table 3- The statistical parameters employed to assess the performances of the models

Statistical indices	*Symbol*	*Equation*
Root mean square error	RMSE	$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2}$
Root mean square error ratio	RaRMSE	$\mathrm{RaRMSE} = 1 - \frac{\mathrm{RMSE}_{\text{calibrated model}}}{\mathrm{RMSE}_{\text{non-calibrated model}}}$
Relative error (%)	RE	$\mathrm{RE} = \frac{\mathrm{RMSE}}{\bar{O}} \times 100$
Mean absolute error	MAE	$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\lvert P_i - O_i \rvert$
Mean absolute error ratio	RaMAE	$\mathrm{RaMAE} = 1 - \frac{\mathrm{MAE}_{\text{calibrated model}}}{\mathrm{MAE}_{\text{non-calibrated model}}}$

Notes: P_i represents the estimated ET_0 values by the models, measured in mm d^{-1}; O_i represents the FAO56-PM ET_0 values, also measured in mm d^{-1}; n represents the total number of data; and $\bar{O}$ represents the mean of FAO56-PM ET_0 values, measured in mm d^{-1}.

3. Results and Discussion

3.1 Empirical models

Table 4 displays the statistical outcomes of both the original and calibrated empirical models that have been taken into account for this study. The linear regression parameters used for calibrating empirical models are presented in Table 5. The JH model (non-calibrated) demonstrated the poorest performance with the values of RMSE of 1.845 mm d^{-1}, MAE of 1.658 mm d^{-1}, and RE of 39.1% in the validation period. According to criteria suggested by Corzo & Solomatine (2007), the JH model showed high errors due to RE ≥ 35%. In contrast, during the validation period, the HS model (non-calibrated) demonstrated the best performance with an RMSE value of 0.742 mm d^{-1}, an MAE value of 0.548 mm d^{-1}, and a RE value of 15.7%, and this model showed medium due to RE values. Similarly, Sabziparvar & Tabari (2010) also showed that HS and Turc models estimated ET_0 more accurately than MAK and Priestley-Taylor (PT) models in semiarid climates. It can be argued that the varying efficacy of empirical approaches in estimating ET_0 stems from the fact that each method is tailored to distinct geographic areas and specific local climatic conditions. Initially developed in the United States for implementation in arid and semi-arid conditions, the JH equation based on T_{max}, T_{min}, and R_s may not provide accurate estimates for regions under different climatic influences. The HS equation based on T_{max} and T_{min}, endorsed globally by the Food and Agriculture Organization (FAO), yielded superior results compared to other empirical equations in our study. The Makkink equation based on T_{max}, T_{min}, and R_s, originally developed under the Netherlands' mild and humid climate conditions as an adaptation of the Penman-Monteith equation, exhibited limited efficacy in estimating ET_0 within our research area, characterized by a Mediterranean climate. The Oudin equation, tested in French basins, demonstrated comparable predictive performance to the HS model in estimating ET_0 within this research. The Ritchie equation, which incorporates coefficients adjusted for specific temperature ranges, did not yield satisfactory results in estimating ET_0.

This disparity suggests the need for region-specific adjustments to improve predictive accuracy. In this instance, it is necessary to ascertain calibration coefficients tailored to the specific local climatic conditions or to calibrate equations using linear regression as recommended by Allen et al. in 1998.

In the study, after calibration by linear regression, the ET_0 estimation performances of Jensen-Haise (JH) and Makkink (MAK) models significantly increased with a sharp decline in RMSE, MAE and RE values, as shown in Table 4. These models strongly correlated with the FAO56-PM method (R = 0.94 for JH; R = 0.93 for MAK, as shown in Table 5) and gave low errors due to RE ≤ 15%. This study calculated RaRMSE and RaMAE values to assess the calibrated models' accuracy. Figure 4 presents a graphical representation of the RaRMSE and RaMAE values for the linear regression-based calibrated models, compared to the non-calibrated (original) models during the validation period. Figure 4 shows that positive values represent progress, while negative values indicate a decline in the models. The Cal_HS model exhibited a 25.6% increase in RMSE and MAE values, as shown in Figure 4. In comparison, Cal_Oudin demonstrated a 13.7% increase in RMSE values and a 21.9% increase in MAE values. It's worth mentioning that the HS and Oudin models calibrated locally had worse statistics than their non-calibrated counterparts. Similarly, a study by Khodayvandie et al. (2022) reported that the locally calibrated HS model gave more inadequate statistics than the original HS model at two locations in Iran. Some studies have also stated that there was no improvement in the prediction performances of some empirical models after they were calibrated (Landeras et al. 2008; Valipour 2015; Djaman et al. 2016; Shiri, 2017; Farias et al. 2020). The Cal_HS and Cal_Oudin showed medium errors like their original versions, with 19.8 and 19.0 % RE values, respectively. Therefore, HS and Oudin models do not need to be calibrated by linear regression for the best performance in this study region. The Jensen-Haise (JH) model showed the most significant improvement

(decreases of 73 and 80% of RMSE and MAE in the validation period, respectively), ranked first among all models after calibration, and was followed by calibrated Makkink (MAK). Ritchie model's performance slightly improved after calibration, and the model's RMSE and MAE values have decreased by 16.3% and 13.7%, respectively. Similar to our research, studies by Djaman et al. (2016), Shiri (2017), Banda et al. (2018), Ferreira et al. (2019), Khodayvandie et al. (2022), and Gharehbaghi & Kaya (2022) confirmed that there was generally an improvement in the performance of empirical models after local calibration. As in our study, research conducted by Farzanpour et al. (2019) in semiarid regions of Iran found that the JH model had the highest performance improvement among the radiation-based models after the local calibration procedure. Similarly, Gharehbaghi & Kaya (2022) found that the calibrated JH model best estimated the ET_0 among the empirical models in the Kutahya province of Türkiye, which has dry, hot summers and snowy and cold winters in their study. It can be said that the calibration coefficients obtained by linear regression, as shown in Table 4 in the present study, are valid for the Adana Plain, which has a hot summer Mediterranean climate. Suppose the available data is insufficient or unreliable to solve the FAO-56 PM equation. In that case, it is possible to use calibration coefficients for JH and MAK models when dealing with irrigation practices in the studied area.

Table 4- The statistical summary of the empirical models

	Calibration period			*Validation period*		
Model	*RMSE* $mm\ d^{-1}$	*MAE* $mm\ d^{-1}$	*RE* %	*RMSE* $mm\ d^{-1}$	*MAE* $mm\ d^{-1}$	*RE* %
JH	1.846	1.654	44.3	1.845	1.658	39.1
MAK	0.793	0.687	19.0	1.089	0.970	23.1
Ritchie	1.296	1.014	31.1	1.379	1.054	29.3
HS	0.953	0.762	22.9	0.742	0.548	15.7
Oudin	0.743	0.566	17.8	0.789	0.575	16.7
Cal_ JH	0.390	0.288	9.4	0.491	0.340	10.4
Cal_ MAK	0.419	0.308	10.1	0.564	0.408	12.0
Cal_ Ritchie	0.956	0.741	22.9	1.154	0.910	24.5
Cal_HS	0.703	0.543	16.9	0.932	0.745	19.8
Cal_ Oudin	0.684	0.524	16.4	0.897	0.702	19.0

Notes: Calibrated models denoted as Cal

Table 5- Empirical and correlation coefficients used for calibrating empirical models

		$ET_0(PM) = a + bET_{0_{model}}$	
Model	***a***	***b***	***R***
JH	0.8334	0.5766	0.94
MAK	0.4183	1.0718	0.93
Ritchie	1.8943	0.4855	0.57
HS	0.2288	0.8238	0.80
Oudin	0.2685	0.8805	0.81

Notes: a represents intercept, b represents slope, and R represents correlation coefficient

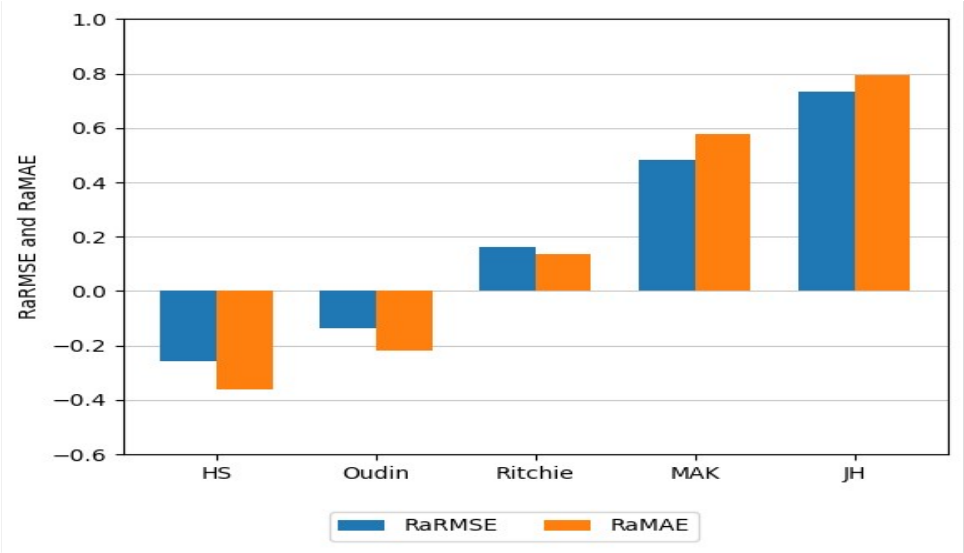


Figure 4- Root mean square error ratio (RaRMSE) and mean absolute error ratio (RaMAE) values of the linear regression-based calibrated empirical models for the validation period

3.2. Data-driven approaches

The statistics of the data-driven models created for the four combinations are shown in Table 6. Table 7 presents the GEP and MLR models' mathematical expressions, respectively. Table 7 shows that, generally, there is a minimal difference between the RMSE, RE, and R values of all models during the training and validation periods. However, the training period's statistical results were slightly better than the validation period for all models. The models with T_{max}, T_{min}, and R_s inputs (3rd combination) and T_{max}, T_{min}, R_s, and U_2 inputs (4th combination) of all data-driven methods successfully estimated ET_0, as evidenced by their high R and low RMSE, MAE and RE values (RE ≤ 15 %). 4th combination of all data-driven methods provided the most accurate predictions of ET_0 during both the training and validation periods. The RMSEs were found to be less than 0.414 mm per day, while the REs were not more than 9.1% in value. So, the findings from the present research coincide with the existing literature suggesting that incorporating more climate variables usually results in improved accuracy of model estimation (Shiri 2017; Mattar & Alazba 2019; Niaghi et al. 2021; Yamaç 2021; Yıldırım et al. 2023; Bayram & Çıtakoğlu 2023).

The RF models showed the lowest RMSE, MAE, and RE in the training and validation periods compared to the GEP, ANN, MLR, and SVM models (during the validation period, the RF1 model was the only exception). The RF4 model achieved the best results with the values of RMSE of 0.116 mm d^{-1}, MAE of 0.076 mm d^{-1}, RE of 2.0 %, and R of 0.995 in the training period, and values of RMSE of 0.224 mm d^{-1}, MAE of 0.151 mm d^{-1}, RE of 3.9 %, and R of 0.983 in the validation period. Overall, the RF and ANN models displayed the best performance among all data-driven models. GEP, MLR, and SVM models showed nearly similar performance considering T_{max}, T_{min}, U_2 and T_{max}, T_{min}, R_s, and U_2 combinations. However, SVM1 and SVM2 models displayed worse performance than their analogous among all data-driven models (RMSE = 1.505-1.454 mm d^{-1}, MAE =1.273-1.244 mm d^{-1}, RE = 30.4-29.6 %, respectively in the validation period). The hyperparameters for the SVM models were determined through trial and error during the study. Due to the limited options for hyperparameter selection provided by the Orange package, SVM1 and SVM2 models are considered not to reach the desired performance level. According to the data presented in Table 6, SVM3 and SVM4 demonstrated slightly better performance levels than those of GEP3 and GEP4, although somewhat inferior to other analogous data-driven models. The results obtained were consistent with previous studies conducted by Sayyadi et al. (2009); Rahimikhoob (2010); Traore et al. (2010), and Yurtseven & Serengil (2021). The studies mentioned utilized the ANN (MLP) method and found it more accurate than other methods when estimating ET_0 in different climates worldwide. Similar to our research, according to Wang et al. (2022), RF-based ET_0 models outperformed GEP-based ET_0 models in their study performed in southwest China's Pearl River basin, which has a tropical and subtropical humid climate. Also, Niaghi et al. (2021) found that in the Red River Valley in the USA, the RF model had a superior ET_0 performance compared to the GEP, SVM, and MLR approaches. In contrast, Yurtseven & Serengil (2021) found that ANN and SVM methods performed better than the RF method in estimating ET_0 in semiarid highland environments.

Studies in the existing literature have explored the impact of different data preprocessing techniques on estimation accuracy in ANNs. For instance, Katipoğlu et al. (2023) investigated the influence of various data preprocessing methods on the accuracy of evaporation estimation using ANN in Adana, which has a Mediterranean climate. Their findings revealed that the standard scalar optimization algorithm presented the highest level of accuracy, and the power transformer showed second-degree promising results. Recently, a novel method employed in evaporation modeling involves implementing hybrid models. Katipoğlu (2023) predicted evaporation with wavelet-based hyperparameter optimized k-nearest neighbors (KNN) and extreme gradient boosting (XGBoost) algorithms in a semi-arid environment. The study specifically investigated the effectiveness of different "mother wavelet" types in enhancing the accuracy of evaporation prediction models. Combining the biorthogonal 2.2 (rbio2.2) mother wavelet with the KNN algorithm yielded the most accurate evaporation predictions. Decomposing the data into sub-signals using wavelet transform generally improved the performance of both KNN and XGBoost models.

In our study, the results presented in Table 6 and Figure 5 indicate that T_{max}, T_{min}, U_2 input (2nd combination) models had higher accuracy than T_{max} and T_{min} input (1st combination) models, and T_{max}, T_{min}, and R_s input (3rd combination) models had higher accuracy than T_{max}, T_{min}, U_2 input (2nd combination) models in two studied periods for all data-driven methods. Additionally, adding U_2 and R_s into the first combination significantly improved the accuracy of the models. During the calibration and validation periods, the first and second model combinations had medium errors, while the third and fourth combinations had low errors based on RE values. The data presented in Figure 5 shows that the RE (relative error) values for the first and second combinations of GEP, ANN, MLR, and RF methods are less than 25%. On the other hand, for the first and second combinations of SVM, the RE values are around 30%. The RE values for the 3rd and 4th combinations of RF and ANN are below 10%, while for the 3rd and 4th combinations of GEP, MLR, and SVM, they are below 15%.

Including U_2 data in the RF1 model improved the accuracy of ET_0 estimation for the new model (RF2), reducing RMSE and MAE values by 20.0 and 19.5%, respectively, during the validation period. On the other hand, when R_s data were incorporated into the RF1 model, the accuracy of ET_0 estimation for the new model (RF3) significantly improved, with a decrease in RMSE and MAE values by 60.9 and 65.7%, respectively, during the validation period. The addition of U_2 and R_s data to the RF1 model led to significant improvement in the accuracy of the ET_0 estimation of the new model (RF4), as evidenced by the decrease in RMSE and MAE values by 78.7% and 81.9%, respectively, during the validation period. Other data-driven methods also demonstrated behaviour similar to that of the RF method. The research conducted by Traore and Guven (2012) and Citakoglu et al. (2014) revealed similar outcomes. It was observed that incorporating U_2 with other meteorological variables enhanced the

efficiency of the models. It is clear from here that the R_s variable has more effect on estimating ET_0 than the T_{max}, T_{min}, and U_2 variables. In Section 2.3, Figure 2 illustrates the strong positive correlation (R = 0.875) between R_s and ET_0, while indicating a notably weak negative correlation (R = -0.165) between RH and ET_0. It is noteworthy that the addition of RH to the GEP4 model (T_{max}, T_{min}, R_s, U_2) during the formulation of the GEP models did not prompt consideration of the RH parameter by the GEP software when establishing the equations. This observation suggests that RH holds minimal influence within the study area. Notably, all data-driven models constructed using the 4-input combination effectively predicted ET_0 without the inclusion of RH, exemplified by the RF model yielding RE = 2.0% and RMSE = 0.116 mm d^{-1} during the training period (Table 5). Consequently, RH was omitted in this study's formulation of input combinations. Pereira et al. (2015) state that R_s can be the dominant factor in predicting ET_0 during summer in humid and sub-humid climates. This is because, in these conditions, the impact of the radiation term is greater than that of the pressure deficit of water vapor and wind factor of the PM method. Studies conducted by Yamaç (2021), Dimitriadou and Nikolakopoulos (2022), Yıldırım et al. (2023), Bayram and Çıtakoğlu (2023) found similar results that R_s has more effect on estimating ET_0 than other variables. A study was conducted in the Peloponnese region of Greece, which has a Mediterranean climate (Csa) similar to our study area. Dimitriadou and Nikolakopoulos (2022) found that net radiation (R_n) and sunshine hours (n) had a more significant impact on ET_0 than other variables.

Table 6- Performance statistics of data-driven models in training and validation periods

		Training period				***Validation period***			
Model	***Input Parameters***	*RMSE mm d^{-1}*	*MAE mm d^{-1}*	*RE %*	*R*	*RMSE mm d^{-1}*	*MAE mm d^{-1}*	*RE %*	*R*
GEP1	T_{max}, T_{min}	0.921	0.759	22.3	0.613	1.094	0.921	23.3	0.405
GEP2	T_{max}, T_{min}, U_2	0.820	0.680	19.7	0.713	0.850	0.708	18.1	0.582
GEP3	T_{max}, T_{min}, R_s	0.438	0.335	10.7	0.928	0.591	0.411	13.1	0.858
GEP4	T_{max}, T_{min}, R_s, U_2	0.337	0.257	8.3	0.958	0.413	0.300	9.0	0.901
ANN1	T_{max}, T_{min}	0.968	0.794	21.8	0.631	0.983	0.807	21.7	0.595
ANN2	T_{max}, T_{min}, U_2	0.825	0.684	18.8	0.751	0.831	0.694	18.7	0.734
ANN3	T_{max}, T_{min}, R_s	0.420	0.303	8.1	0.942	0.422	0.296	7.7	0.939
ANN4	T_{max}, T_{min}, R_s, U_2	0.264	0.197	5.6	0.977	0.265	0.194	5.4	0.976
MLR1	T_{max}, T_{min}	0.920	0.754	22.1	0.615	1.107	0.937	23.5	0.633
MLR2	T_{max}, T_{min}, U_2	0.818	0.679	19.6	0.713	0.847	0.707	18.0	0.767
MLR3	T_{max}, T_{min}, R_s	0.383	0.280	9.2	0.944	0.518	0.355	11.0	0.941
MLR4	T_{max}, T_{min}, R_s, U_2	0.249	0.183	6.0	0.977	0.299	0.227	6.3	0.974
RF1	T_{max}, T_{min}	0.633	0.488	12.9	0.862	1.054	0.832	22.0	0.507
RF2	T_{max}, T_{min}, U_2	0.450	0.341	9.1	0.933	0.843	0.670	17.7	0.725
RF3	T_{max}, T_{min}, R_s	0.206	0.138	3.5	0.986	0.412	0.285	7.2	0.941
RF4	T_{max}, T_{min}, R_s, U_2	0.116	0.076	2.0	0.995	0.224	0.151	3.9	0.983
SVM1	T_{max}, T_{min}	1.503	1.264	30.9	-0.670	1.505	1.273	30.4	-0.718
SVM2	T_{max}, T_{min}, U_2	1.445	1.235	29.9	-0.583	1.454	1.244	29.6	-0.644
SVM3	T_{max}, T_{min}, R_s	0.565	0.455	11.8	0.944	0.573	0.456	11.5	0.883
SVM4	T_{max}, T_{min}, R_s, U_2	0.364	0.272	7.6	0.957	0.369	0.275	7.6	0.953

Table 7- The GEP and MLR models' mathematical expressions

Model	*Formula*
GEP1	$ET_0 = \sqrt{(3.41) + \frac{(3.26)\text{x}(-3.63) + T_{max}}{\sqrt{T_{max}}}} + \sqrt{\frac{\frac{T_{max}\ \text{x}\ T_{max}}{6.52\ \text{x}\ T_{min}}}{\sqrt{T_{max}^{\,6.52}}}} + \left(\frac{\frac{T_{max}\ \text{x}\ T_{max}}{(-9.58) - T_{min}}}{\sqrt{3.75^{4.28}}}\right)^2$
MLR1	$ET_0 = -0.391 + 0.129 \times T_{max} + 0.03 \times T_{min}$
GEP2	$ET_0 = U2 + \log(T_{max} + U2 + T_{min}\ \text{x}\ (-9.07) + T_{max}^{\,2} + T_{max}) + (-8.45) + \sqrt{T_{max} + U_2 + 1.57} + (-7.90)$
MLR2	$ET_0 = -1.908 + 0.170 \times T_{max} - 0.009\ \text{x}\ T_{min} + 1.127\ \text{x}\ U_2$
GEP3	$ET_0 = \max\left(\frac{\frac{R_s + 8.45}{2} + \frac{R_s^{\,2}}{2}}{2},\ \frac{1}{8.45}\ \text{x}\ (R_s + T_{min})\right) + \tanh(-0.14) + \frac{(-0.14)\ \text{x}\ R_s}{2} + \frac{(-1.72)}{R_s} + \frac{2}{\frac{T_{max} + T_{min}}{2} - \frac{(-4.34) + T_{max}}{2} + \frac{(-4.34) - T_{max}}{T_{max}}}$ while $\max(x,y) = \begin{cases} y, & x \le y \\ x, & x > y \end{cases}$
MLR3	$ET_0 = -1.336 + 0.020\ \text{x}\ T_{max} + 0.070\ \text{x}\ T_{min} + 0.179\ \text{x}\ R_s$
GEP4	$ET_0 = max\left(\left(tanh(R_s) + \frac{(-2.09 + U2)}{2}\right)^2, \left(\frac{R_s + 5.97}{2} + \frac{U2 + (-9.02)}{2}\right)\right) + log\left(min\left((T_{max} - U2 + T_{max}),\ \left((U2)^2 + \frac{(-1.72) + R_s}{2}\right)\right)\right) + arctan\left(\frac{U2 + arctan\left(\frac{R_s + T_{min}}{2}\right) - U2 - R_s + 0.78}{2}\right)$ $while\ max(x,y) = \begin{cases} y, & x \le y \\ x, & x > y \end{cases}\ and\ \ min(x,y) = \begin{cases} x, & x \le y \\ y, & x > y \end{cases}$
MLR4	$ET_0 = -2.350 + 0.054\ \text{x}\ T_{max} + 0.040\ \text{x}\ T_{min} + 0.792\ \text{x}\ U_2 + 0.169\ \text{x}\ R_s$

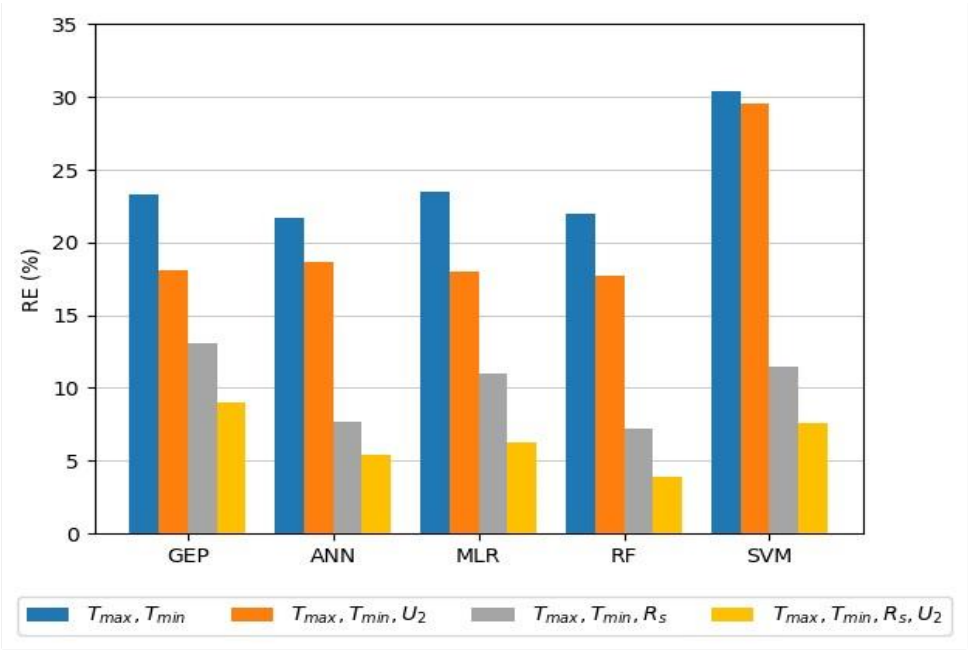


Figure 5- Performance of the data-driven models during the validation period

3.3. Comparison of empirical and data-driven approaches

This section of the study compares the performance of original (non-calibrated), locally calibrated empirical models and data-driven methods in estimating ET_0. The HS and Oudin models in this study correspond to the GEP1, ANN1, MLR1, RF1, and SVM1 models, while the JH, MAK, and Ritchie models align with the GEP3, ANN3, MLR3, RF3, and SVM3 models based on

their respective input combinations. As shown in Table 4 and Table 6, the models that used three or four inputs were more effective than the models that used only two inputs when it came to estimating ET_0, whereas the four-input data-driven models (T_{max}, T_{min}, R_s, U_2 input) outperformed all original and calibrated empirical models. ANN1 and RF1 models (T_{max} and T_{min} inputs) provided higher accuracy than their analogous data-driven models in training and validation periods. On the other hand, original (non-calibrated) HS and Oudin models (T_{max} and T_{min} input) performed better than their equivalent data-driven models in the validation period. The calibrated JH (Cal_JH) model predicted ET_0 slightly better than the GEP3, MLR3 and SVM3 models, while the ANN3 and RF3 models outperformed the Cal_JH model in the validation period. Cal_MAK model estimated somewhat better than the GEP3 model in the validation period, whereas the models ANN3, MLR3, RF3, and SVM3 were found to be more accurate than the Cal_MAK model when estimating ET_0 in the validation period. Similar to our research, Mehdizadeh et al. (2017) have found that some calibrated radiation-based empirical models performed similarly to soft computing approaches in their study.

On the other hand, HS and ANN1 models with T_{max} and T_{min} inputs presented more accurate results than their equivalent models, while among all the models created with T_{max} and T_{min} inputs, the HS model was the most accurate. Cal_JH and RF3 models created using T_{max}, T_{min}, and R_s inputs performed more accurately than their analogous models, whereas out of all the models that were tested with T_{max}, T_{min} and R_s input, the RF3 model displayed the highest level of accuracy. The RF4 model proved to be the most accurate in predicting the ET_0 in all models.

Figures 6, 7, and 8 show diagrams depicting the Taylor, the box plot, and the prediction-errors box plot, and Figure 9 shows time series graphics for the best empirical and data-driven model during the validation period, respectively. Table 8 enlists the descriptive statistics of prediction errors for the top input combinations. In Figure 6, it is evident that the RF4 model outperformed the other models. The RF4 model performed best with low RMSE and high correlation coefficient. The RF4 model provided estimates of ET_0 similar to those of the FAO56-PM method. Also, it is apparent from the Figure 6 that RF3 and Cal_JH models produce similar outcomes. The boxplot charts of the developed data-driven models are similar, as shown in Figure 7. Average ET_0 values estimated by data-driven methods range from 4.35 to 4.38 mm d^{-1}, while that of FAO56-PM is 4.37 mm d^{-1}. However, both HS and Cal_JH had a higher ET_0 compared to FAO56-PM. The average ET_0 of the HS was 4.80 mm d^{-1}, while the average ET_0 of Cal_JH was 4.47 mm d^{-1}. According to Table 8, the RF4 model followed the corresponding values with the lowest lower quartile (0.046) and upper quartile (0.1966), and the standard deviation value is less than that of other data-driven and empirical models. According to Figure 8, it can be inferred that RF4, characterized by the lowest median value and narrow interquartile range, emerges as the most effective model. According to the information depicted in Figure 9, it is apparent that data-driven models were more effective than empirical equations in forecasting peak ET_0 values. Among the models assessed, RF4 exhibited the highest level of accuracy in predicting peak ET_0 values. In contrast, the HS and Cal_JH equations demonstrated deficiencies in precisely forecasting peak ET_0 values. Based on Figure 9, it can be deduced that the Hargreaves-Samani equation generally overestimates ET_0, while the Cal_JH equation tends to underestimate ET_0.

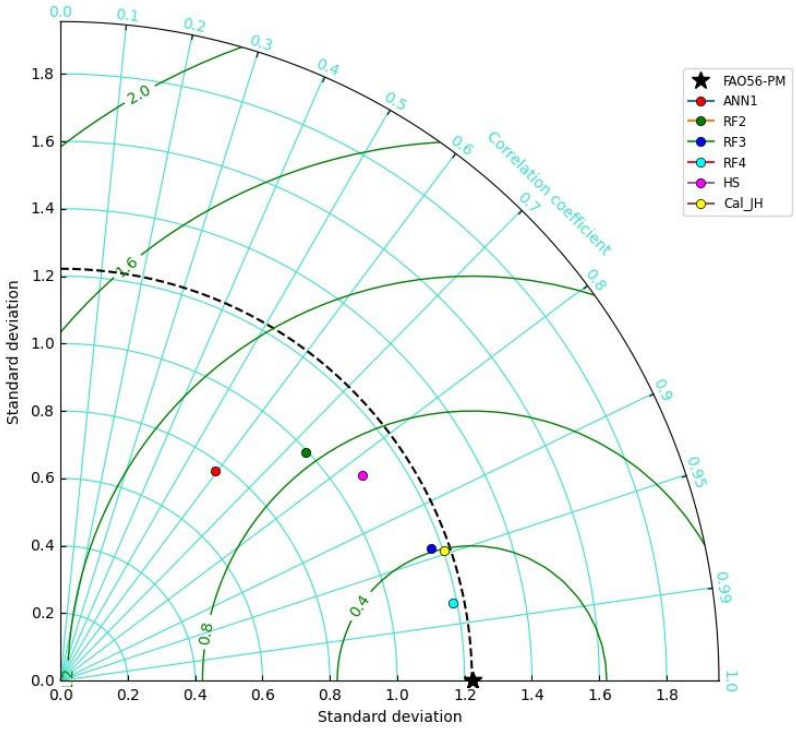


Figure 6- Taylor diagram for the best empirical and data-driven model according to each input combination

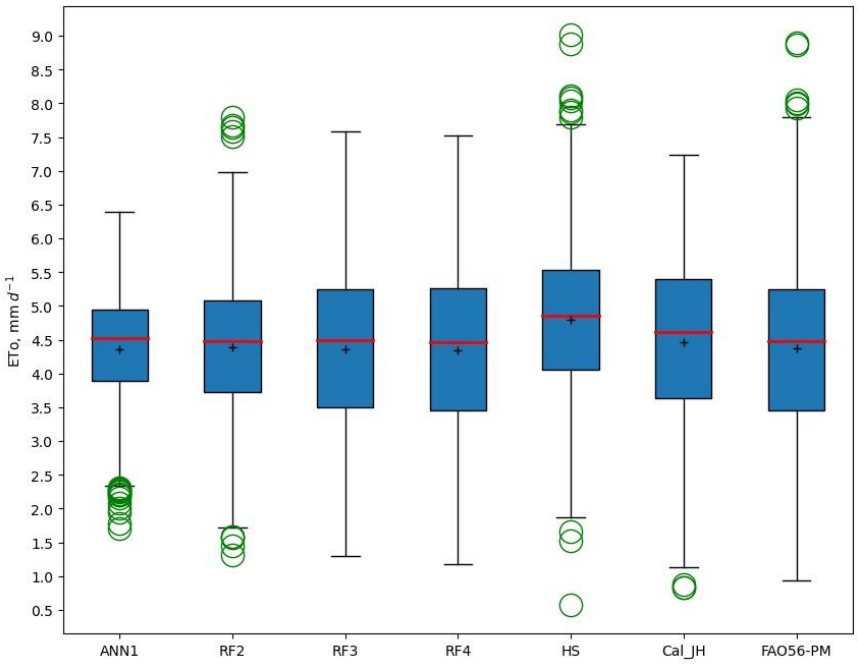


Figure 7- Box plot diagram for the best empirical and data-driven model according to each input combination

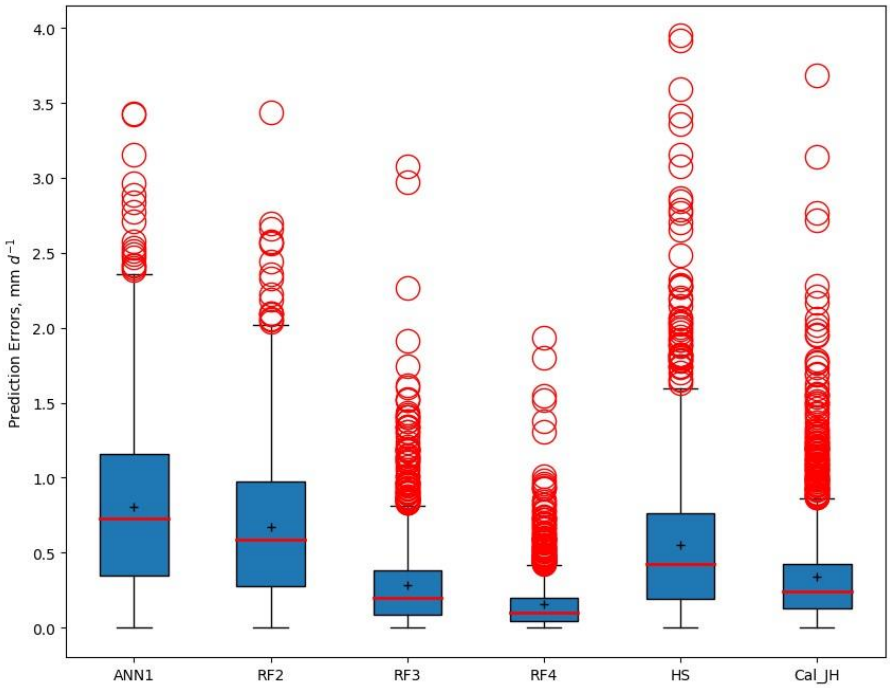
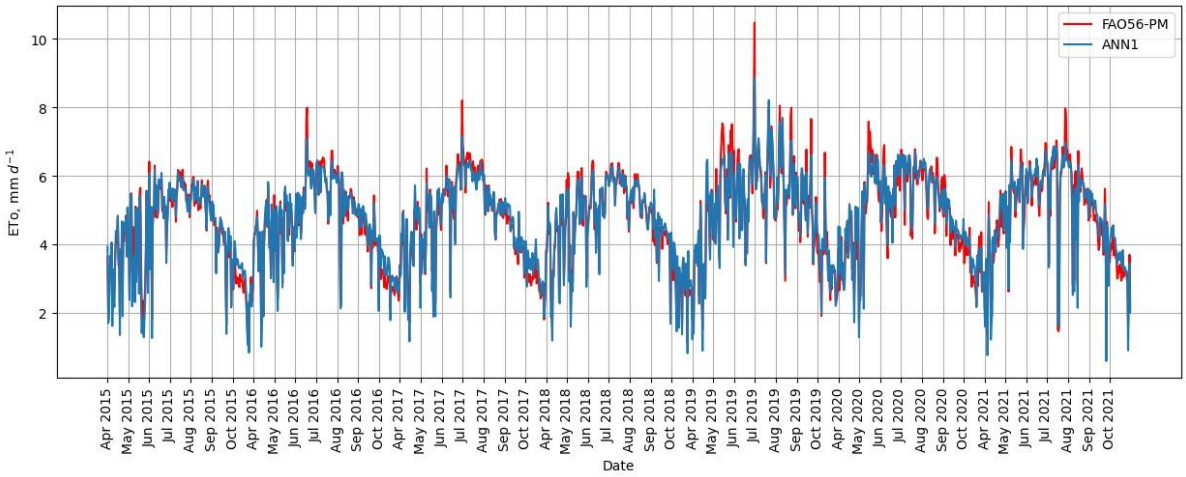
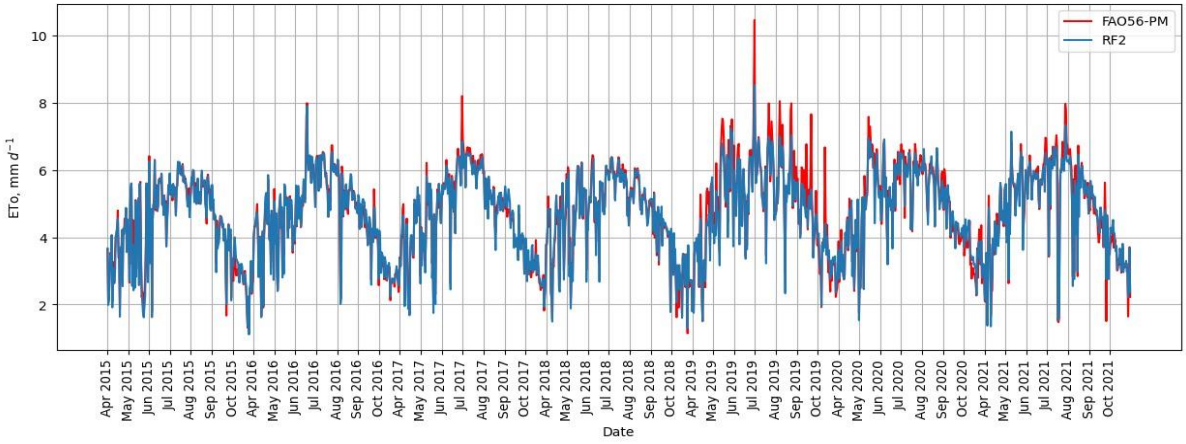
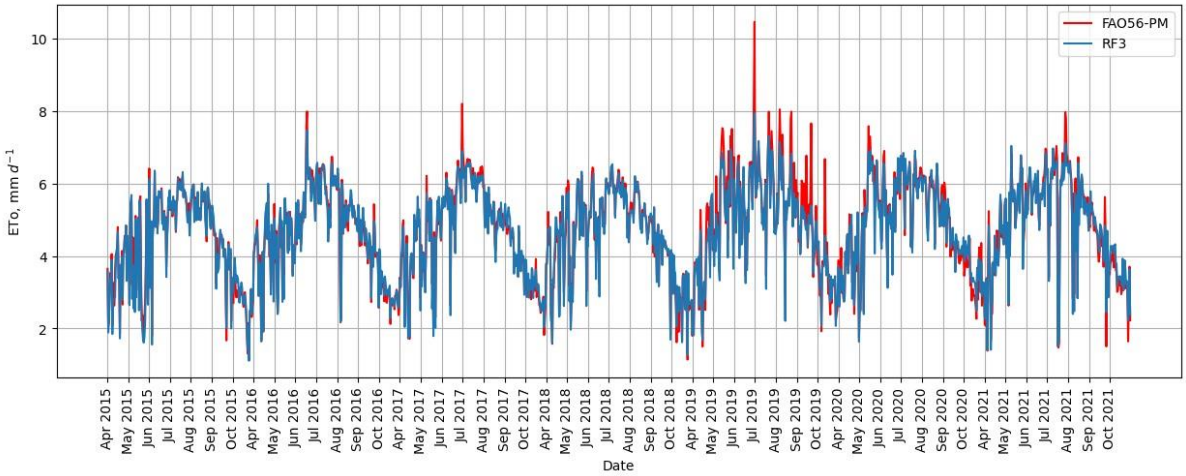
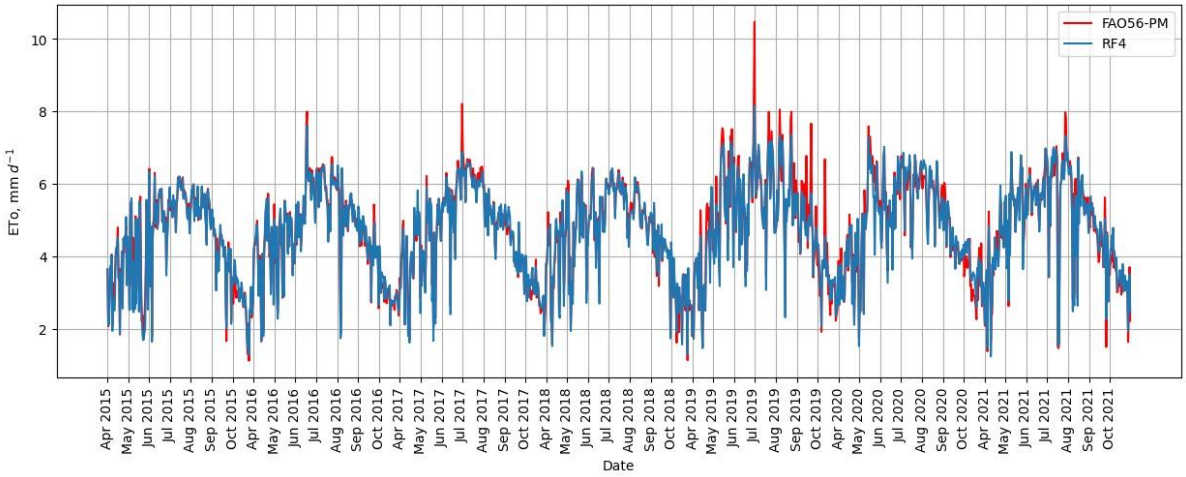


Figure 8- Prediction-errors box plot diagram for the best empirical and data-driven model according to each input combination


FAO56-PM
ANN1
ETo, mm d^{-1}
Date


FAO56-PM
RF2
ETo, mm d^{-1}
Date


FAO56-PM
RF3
ETo, mm d^{-1}
Date


FAO56-PM
RF4
ETo, mm d^{-1}
Date

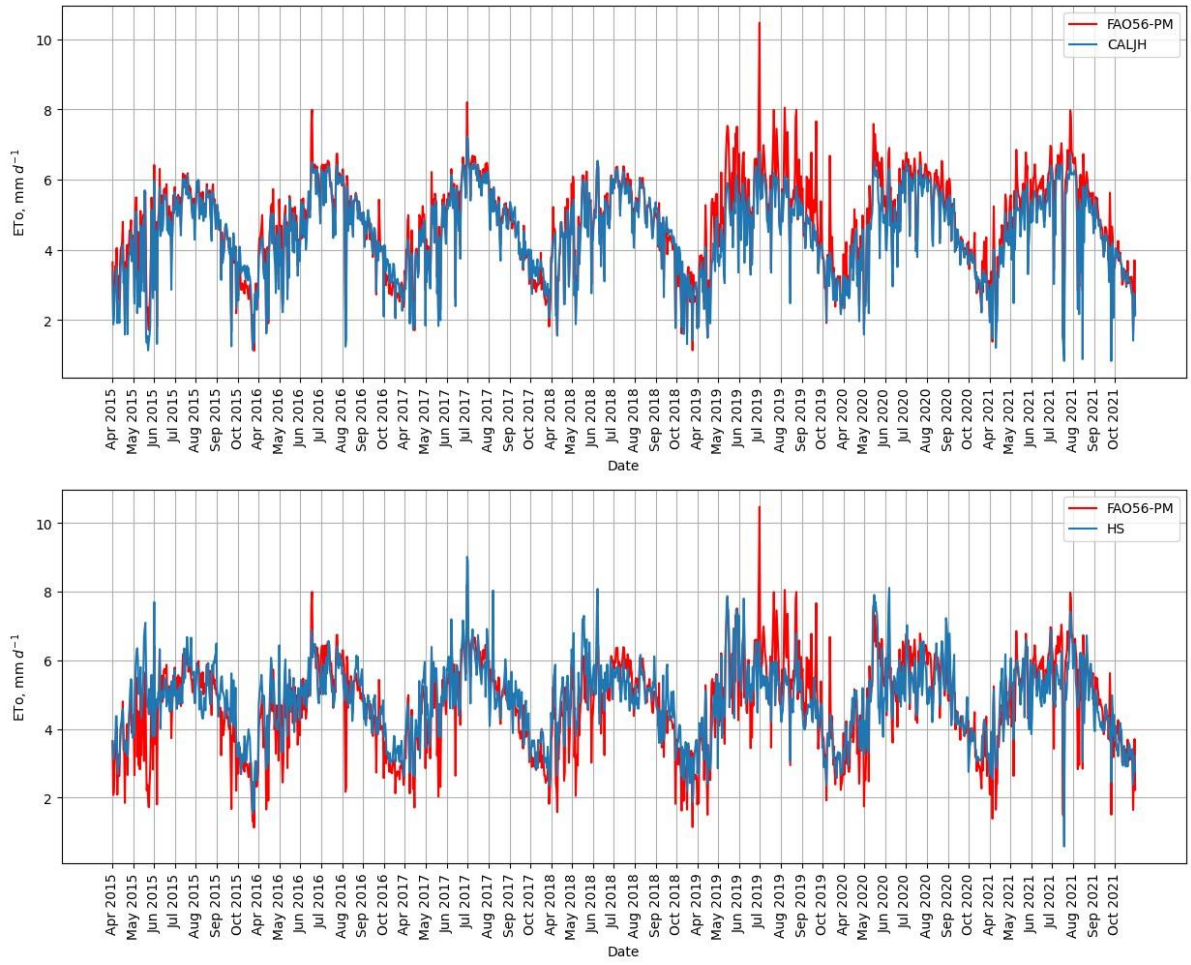


Figure 9- Time series graphics of daily FAO56-PM ETo and daily model-estimated ETo for the optimal data-driven and empirical model based on each input combination

Table 8- Box plot statistics for the optimal models based on each input combination

Model	Lower Quartile	Upper Quartile	Mean	Median	Standard Deviation
ANN1	0.3496	1.1593	0.8064	0.7282	0.5619
RF2	0.2762	0.9745	0.6730	0.5860	0.5005
RF3	0.0877	0.3829	0.2844	0.2006	0.2965
RF4	0.0466	0.1966	0.1544	0.0973	0.1799
HS	0.1902	0.7615	0.5482	0.4250	0.4999
Cal_JH	0.1259	0.4215	0.3398	0.2382	0.3542

Many studies show that data-driven models with all input parameters are superior to empirical and calibrated empirical models, as in our study (Benzaghta et al. 2012; Karimaldini et al. 2012; Tabari et al. 2013b). Benzaghta et al. (2012) found that using all input parameters in ANN leads to better evaporation estimation performance. Jain et al. (2008) highlighted that the accurate estimation of ET_0 requires temperature and radiation data as crucial inputs in Southwestern Idaho, USA, which has a semi-arid climate. In our study, the performance of equation methods was significantly impacted by the presence or absence of critical input. However, the performance values in data-driven models may vary despite having the same input set due to model dynamics.

In this study, the hyperparameter values used to create data-driven models in the Orange software were limited to the software's provided constraints. Consequently, precise determination of optimal model performance values through trial and error was not feasible in the SVM models. Nonetheless, utilizing software such as python enables the performance of machine learning model tuning and hyperparameter optimization using techniques like “Grid Search” and “Random Search”. We recommend employing these methods in future studies to enhance the accuracy of ET_0 predictions. In recent research, hybrid models have shown promise in enhancing the accuracy of estimating ET_0. These models have demonstrated improved estimation accuracy by leveraging the strengths of diverse algorithms. Furthermore, it is suggested that future studies incorporate the use of hybrid models and integrate deep learning models alongside traditional data-driven methods to optimize ET_0 estimation further.

4. Conclusions

The estimation of reference evapotranspiration (ET_0) is crucial for water resource management, irrigation practices, and agricultural and hydrometeorological research. Thus, the precise prediction of ET_0 is of the utmost importance.

Based this study, the key findings derived from this study are outlined as follows: The study conclusively establishes that the application of data-driven techniques yields satisfactory outcomes in estimating daily ET_0, even in cases where not all the climate parameters required for the reference method FAO56-PM ET_0 are available. The MLR and GEP equations acquired can have applications in agricultural irrigation within the researched area. The calibration process significantly improved the accuracy of the JH and MAK empirical equations in forecasting ET_0. We recommend utilizing the calibration coefficients derived from these equations to make precise ET_0 projections within the research area. The Orange software's limitations in selecting hyperparameters, particularly in this study, resulted in reduced ET_0 estimation performance from the SVM1 and SVM2 models compared to other equivalent model combinations. It is advisable to consider employing methodologies such as "grid search" and "random search" for the purpose of hyperparameter optimization in forthcoming research endeavors within the domain of machine learning.

In future research, the predictive accuracy of reference ET_0 could be investigated utilizing deep learning models, hybrid models, and methodologies incorporating lag values.

Funding

The authors did not receive support from any organization for the submitted work.

Data Availability

The data supporting this study's findings are available on request from the corresponding author.

Conflict of Interest

The authors declare that there is no conflict of interest.

References

Abrishami N, Sepaskhah A R & Shahrokhnia M H (2019). Estimating wheat and maize daily evapotranspiration using artificial neural network. Theoretical and Applied Climatology 135: 945-958. https://doi.org/10.1007/s00704-018-2418-4

Achite M, Jehanzaib M, Sattari M T, Toubal A K, Elshaboury N, Wałęga A, Krakauer N Y, Yoo J & Kim T (2022). Modern techniques to modeling reference evapotranspiration in a semiarid area based on ANN and GEP models. Water 14(8): 1210. https://doi.org/10.3390/w14081210

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

Almorox J & Grieser J (2016). Calibration of the Hargreaves–Samani method for the calculation of reference evapotranspiration in different Köppen climate classes. Hydrology Research 47(2): 521-531. https://doi.org/10.2166/nh.2015.091

Antonopoulos V Z & Antonopoulos A V (2017). Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. Computers and Electronics in Agriculture 132: 86-96. https://doi.org/10.1016/j.compag.2016.11.011

Banda P, Cemek B & Küçüktopcu E (2018). Estimation of daily reference evapotranspiration by neuro computing techniques using limited data in a semiarid environment. Archives of Agronomy and Soil Science 64(7): 916-929. https://doi.org/10.1080/03650340.2017.1414196

Bayram S & Çıtakoğlu H (2023). Modeling monthly reference evapotranspiration process in Turkey: application of machine learning methods. Environmental Monitoring and Assessment 195(1): 67. https://doi.org/10.1007/s10661-022-10662-z

Benzaghta M A, Mohammed T A, Ghazali A H, Amin M S M & Soom M A B M (2012). Prediction of evaporation in tropical climate using artificial neural network and climate based models. Scientific Research and Essays 7(36): 3133-3148. https://doi.org/10.5897/SRE11.1311

Breiman L (2001). Random forests. Machine Learning 45: 5-32

Chen Z, Zhu Z, Jiang H and Sun S (2020). Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. *Journal of Hydrology* 591(2020): 125286. https://doi.org/10.1016/j.jhydrol.2020.125286

Choi Y, Kim M, O'Shaughnessy S, Jean J, Kim Y & Song W J (2018). Comparison of artificial neural network and empirical models to determine daily reference evapotranspiration. *Journal of Korean Society of Agricultural Engineers* 60(6): 43-54. https://doi.org/10.5389/KSAE.2018.60.6.043

Citakoglu H, Cobaner M, Haktanir T & Kisi O (2014). Estimation of monthly mean reference evapotranspiration in Turkey. Water Resources Management 28: 99-113. https://doi.org/10.1007/s11269-013-0474-1

Cobaner M, Citakoglu H, Haktanir T & Kisi O (2017). Modifying Hargreaves–Samani equation with meteorological variables for estimation of reference evapotranspiration in Turkey. Hydrology Research 48: 480-497. https://doi.org/10.2166/nh.2016.217

Corzo G & Solomatine D (2007). Baseflow separation techniques for modular artificial neural network modelling in flow forecasting. *Hydrological Sciences Journal* 52(3): 91-507. https://doi.org/10.1623/hysj.52.3.491

CSB (2023). Urbanization and climate change. The governor of Adana, Ministry of Environment. https://adana.csb.gov.tr/ilim izi-taniyalim-i-1222, Accessed 16 February 2023.

Cutler A, Cutler D R & Stevens J R (2012). Random forests. Ensemble machine learning: Methods and applications, pp. 157-175. https://doi.org/10.1007/978-1-4419-9326-7_5.

Dimitriadou S & Nikolakopoulos K G (2022). Multiple linear regression models with limited data for the prediction of reference evapotranspiration of the Peloponnese, Greece. Hydrology 9(7): 124. https://doi.org/10.3390/hydrology9070124

Djaman K, Tabari H, Balde A B, Diop L, Futakuchi K & Irmak S (2016). Analyses, calibration and validation of evapotranspiration models to predict grass-reference evapotranspiration in the Senegal river delta. *Journal of Hydrology*: Regional Studies 8: 82-94. https://doi.org/10.1016/j.ejrh.2016.06.003

Dong J, Xing L, Cui N, Guo L, Liang C, Zhao L, Wang Z & Gong D (2024) Estimating reference crop evapotranspiration using optimized empirical methods with a novel improved Grey Wolf Algorithm in four climatic regions of China. Agricultural Water Management 291: 2024. https://doi.org/10.1016/j.agwat.2023.108620

Dou X & Yang Y (2018). Evapotranspiration estimation using four different machine learning approaches in different terrestrial ecosystems. Computers and Electronics in Agriculture 148: 95-106. https://doi.org/10.1016/j.compag.2018.03.010

Farias V D, Costa D L, Pinto J V, Souza P J, Souza E B & Ortega-Farias S (2020). Calibration of reference evapotranspiration models in Pará. Acta Scientiarum Agronomy 42. https://doi.org/10.4025/actasciagron.v42i1.42475

Farzanpour H, Shiri J, Sadraddini A A & Trajkovic S (2019). Global comparison of 20 reference evapotranspiration equations in a semiarid region of Iran. Hydrology Research 50(1): 282-300. https://doi.org/10.2166/nh.2018.174

Fauset L V (2006). Fundamentals of neural networks: architectures, algorithms and applications. Pearson Education India.

Feng Y, Cui N, Zhao L, Hu X & Gong D (2016). Comparison of ELM, GANN, WNN and empirical models for estimating reference evapotranspiration in humid region of Southwest China. *Journal of Hydrology* 536: 376-383. https://doi.org/10.1016/j.jhydrol.2016.02.053

Feng Y, Peng Y, Cui, N, Gong D & Zhang K (2017). Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. Computers and Electronics in Agriculture 136: 71-78. https://doi.org/10.1016/j.compag.2017.01.027

Ferreira C (2001). Gene expression programming: A new adaptive algorithm for solving problems. Complex Systems 13(2): 87-129

Ferreira L B, Cunha F F, Duarte A B, Sediyama G C & Cecon P R (2018). Calibration methods for the Hargreaves-Samani equation. Ciencia E Agrotecnologia 42: 104-114. http://dx.doi.org/10.1590/1413-70542018421017517

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

Gao X, Peng S, Xu J, Yang S & Wang W (2015). Proper methods and its calibration for estimating reference evapotranspiration using limited climatic data in Southwestern China. Archives of Agronomy and Soil Science 61(3): 415-426. https://doi.org/10.1080/03650340.2014.933810

Gavili S, Sanikhani H, Kisi O & Mahmoudi M H (2018). Evaluation of several soft computing methods in monthly evapotranspiration modelling. Meteorological Applications 25(1): 128-138. https://doi.org/10.1002/met.1676

Gharehbaghi A & Kaya B (2022). Calibration and evaluation of six popular evapotranspiration formula based on the Penman-Monteith model for continental climate in Turkey. Physics and Chemistry of the Earth, Parts A/B/C, 127, 103190. https://doi.org/10.1016/j.pce.2022.103190

Gocić M, Motamedi S, Shamshirband S, Petković D, Sudheer C, Roslan H & Muhammad A (2015). Soft computing approaches for forecasting reference evapotranspiration. Computers and Electronics in Agriculture 113: 164-173. https://doi.org/10.1016/j.compag.2015.02.010

Goldblum M, Finzi M, Rowan K & Wilson A G (2023). The no free lunch theorem, Kolmogorov complexity, and the role of inductive biases in machine learning. arXiv preprint arXiv:2304.05366. https://doi.org/10.48550/arXiv.2304.05366

Gomariz-Castillo F, Alonso-Sarria F & Cabezas-Calvo-Rubio F (2018). Calibration and spatial modelling of daily ET_0 in semiarid areas using Hargreaves equation. Earth Science Informatics 11(3): 325-340. https://doi.org/10.1007/s12145-017-0327-1

Hargreaves G H & Samani Z A (1985). Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture 1: 96-99.

Huo Z, Feng S, Kang S & Dai X (2012). Artificial neural network models for reference evapotranspiration in an arid area of northwest China. *Journal of Arid Environments* 82: 81-90. https://doi.org/10.1016/j.jaridenv.2012.01.016

Irmak S, Irmak A, Allen R G & Jones J W (2003). Solar and net radiation-based equations to estimate reference evapotranspiration in humid climates. *Journal of Irrigation and Drainage Engineering* 129(5): 336-347. https://doi.org/10.1061/(ASCE)0733-9437(2003)129:5(336)

Izadifar Z (2010). Modeling and analysis of actual evapotranspiration using data driven and wavelet techniques. Undergraduate Dissertation, University of Saskatchewan.

Jacovides C P & Kontoyiannis H (1995). Statistical procedures for the evaluation of evapotranspiration computing models. Agricultural Water Management 27: 365-371. https://doi.org/10.1016/0378-3774(95)01152-9

Jain S K, Nayak P C & Sudheer K P (2008). Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. Hydrological Processes 22(13): 2225-2234. https://doi.org/10.1002/hyp.6819

Jang J C, Sohn E H, Park K H & Lee S (2021). Estimation of daily potential evapotranspiration in real-time from GK2A/AMI data using artificial neural network for the Korean Peninsula. Hydrology 8(3): 129 https://doi.org/10.3390/hydrology8030129

Jensen M E & Haise H R (1963). Estimating evapotranspiration from solar radiation. *Journal of the Irrigation and Drainage Division* 89(4): 15-41. https://doi.org/10.1061/JRCEA4.000028

Jones J W & Ritchie J T (1990). Crop growth models. In: G. J. Hoffman, T. A. Howel, & K. H. Solomon (eds.) Management of farm irrigation systems. ASAE Monograph 9: 63–69

Kades C (2019). Economy of Adana in 2019 [in Turkish] https://www.adanato.org.tr/ WebDosyalar/V2/Dosyalar/2020/8/13/adana-ekonomisi-2019-16-53-18.pdf, Accessed 14 January 2023

Karimaldini F, Shui L T, Mohamed T A, Abdollahi M & Khalili N (2012). Daily evapotranspiration modeling from limited weather data by using neuro-fuzzy computing technique *Journal of Irrigation and Drainage Engineering* 138(1): 21-34. https://doi.org/10.1061/(ASCE)IR.1943-4774.000034

Karunanithi N, Grenney W J, Whitley L D & Bovee K (1994). Neural networks for river flow prediction. *Journal of Computing in Civil Engineering* 8(2): 201-220. https://doi.org/10.1061/(ASCE)0887-3801(1994)8:2(201)

Katipoğlu O M (2023). Evaporation prediction with wavelet-based hyperparameter optimized K-Nearest Neighbors and Extreme Gradient Boosting Algorithms in a semi-arid environment. Environmental Processes 10(4): 50-51. https://doi.org/10.1007/s40710-023-00669-0

Katipoğlu O M, Pekin M A & Akil S (2023). The Impact of preprocessing approaches on neural network performance: A case sudy on evaporation in Adana, a Mediterranean climate." *Indonesian Journal of Earth Sciences* 3(2): A821-A821 https://doi.org/10.52562/injoes.2023.821

Kaya Y Z, Zeleňáková M, Üneş F, Demirci M, Hlavatá H & Mésároš P (2021). Estimation of daily evapotranspiration in Košice City (Slovakia) using several soft computing techniques. Theoretical and Applied Climatology 144(2021): 287-298. https://doi.org/10.1007/s00704-021-03525-z

Khodayvandie B, Nazemi A H, Shiri J & Trajkovic S (2022). Adopting regional calibration scenarios for ensuring reliable ETo estimates in semiarid regions: assessing the ancillary data supply method. *ISH Journal of Hydraulic Engineering* 28(sup1): 299-309. https://doi.org/10.1080/09715010.2020.1784803

Kisi O (2016). Modeling reference evapotranspiration using three different heuristic regression approaches. Agricultural Water Management 169: 162-172. https://doi.org/10.1016/j.agwat.2016.02.026

Kumar M, Raghuwanshi N S & Singh R (2011). Artificial neural networks approach in evapotranspiration modeling: A review. Irrigation Science 29: 11-25. https://doi.org/10.1007/s00271-010-0230-8

Landeras G, Ortiz-Barredo A & López J J (2008). Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). Agricultural Water Management 95(5): 553-565. https://doi.org/10.1016/j.agwat.2007.12.011

Liu S & Xu Z (2018). Micrometeorological methods to determine evapotranspiration. In: Li, X., Vereecken, H. (eds) Observation and measurement. Ecohydrol. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-47871-4_7-2

Makkink G F (1957). Ekzameno de la formulo de Penman. Netherlands Journal of Agricultural Science 5: 290-305.

Mattar M A & Alazba A A (2019). GEP and MLR approaches for the prediction of reference evapotranspiration. Neural Computing and Applications 31: 5843-5855. https://doi.org/10.1007/s00521-018-3410-8

Mcgill R, Tukey J W & Larsen W A (1978). Variations of box plots. The American Statistician 32(1): 12–16. https://doi.org/10.1080/00031305.1978.10479236

Mehdizadeh S, Behmanes J & Khalili K (2017). Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. Computers and Electronics in Agriculture 139: 103-114. https://doi.org/10.1016/j.compag.2017.05.002

Mohsin S & Lone M A (2021). Modeling of reference evapotranspiration for temperate Kashmir Valley using linear regression. Modeling Earth Systems and Environment 7: 495-502. https://doi.org/10.1007/s40808-020-00921-8

Negm A, Minacapilli M & Provenzano G (2018). Downscaling of American National Aeronautics and Space Administration (NASA) daily air temperature in Sicily, Italy, and effects on crop reference evapotranspiration. Agricultural Water Management 209: 151-162. https://doi.org/10.1016/j.agwat.2018.07.016

Noi P T, Degener J & Kappas M (2017). Comparison of multiple linear regression, cubist regression, and random forest algorithms to estimate daily air surface temperature from dynamic combinations of MODIS LST Data. Remote Sensing 9(5): 398. https://doi.org/10.3390/rs9050398

Orange Data Mining (2024). Orange (Version 3.34) [Software]. https://orangedatamining.com/ Accessed 15 January 2024

Oudin L, Hervieu F, Michel C, Perrin C, Andréassian V, Anctil F & Loumagne C (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model? Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. Journal of Hydrology 303(1-4): 290-306. https://doi.org/10.1016/j.jhydrol.2004.08.026

Pereira L S, Allen R G, Smith M & Raes D (2015). Crop evapotranspiration estimation with FAO56: past and future. Agricultural Water Management 147: 4-20. https://doi.org/10.1016/j.agwat.2014.07.031

Prasad R, Deo R C, Li Y & Maraseni T (2017). Input selection and performance optimization of ANN-based streamflow forecasts in the drought-prone Murray Darling Basin region using IIS and MODWT algorithm. Atmospheric Research 197 (2017): 42-63. https://doi.org/10.1016/j.atmosres.2017.06.014, 2017

Rahimikhoob A (2010). Estimation of evapotranspiration based on only air temperature data using artificial neural networks for a subtropical climate in Iran. Theoretical and Applied Climatology 101: 83-91. https://doi.org/10.1007/s00704-009-0204-z

Niaghi A R, Hassanijalilian O & Shiri J (2021). Estimation of reference evapotranspiration using spatial and temporal machine learning approaches. Hydrology 8(1): 25. https://doi.org/10.3390/hydrology8010025

Reis M M, Silva A J, Junior J Z, Santos L D, Azevedo A M & Lopes E M (2019). Empirical and learning machine approaches to estimating reference evapotranspiration based on temperature data. Computers and Electronics in Agriculture 165: 104937. https://doi.org/10.1016/j.compag.2019.104937

Sabziparvar A A & Tabari H (2010). Regional estimation of reference evapotranspiration in arid and semiarid regions. Journal of Irrigation and Drainage Engineering 136: 724-731. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000242

Samadianfard S, Kargar K, Shadkani S, Hashemi S, Abbaspour A & Safari M J (2022). Hybrid models for suspended sediment prediction: Optimized random forest and multilayer perceptron through genetic algorithm and stochastic gradient descent methods. Neural Computing and Applications 34: 3033–3051. https://doi.org/10.1007/s00521-021-06550-1

Sarıgöl M & Katipoğlu O M (2024). Estimation of monthly evaporation values using gradient boosting machines and mode decomposition techniques in the Southeast Anatolia Project (GAP) area in Turkey. Acta Geophysica 2023: 1–18. https://doi.org/10.1007/s11600-023-01067-8

Sayyadi H, Oladghaffari A, Faalian A & Sadraddini AA (2009). Comparison of RBF and MLP neural networks performance for estimation of reference crop evapotranspiration. Water and Soil Science 19(1): 1-12

Seifi A & Riahi H (2020). Estimating daily reference evapotranspiration using hybrid gamma test-least square support vector machine, gamma test-ANN, and gamma test-ANFIS models in an arid area of Iran. *Journal of Water and Climate Change* 11(1): 217-240. https://doi.org/10.2166/wcc.2018.003

Shiri J, Kisi O, Landeras G, Lopez J J, Nazemi A H & Stuyt L (2012). Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain). *Journal of Hydrology* 414: 302-316. https://doi.org/10.1016/j.jhydrol.2011.11.004

Shiri J (2017). Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in hyper-arid regions of Iran. Agricultural Water Management 188: 101-114. https://doi.org/10.1016/j.agwat.2017.04.009

Shiri J (2018). Improving the performance of the mass transfer-based reference evapotranspiration estimation approaches through a coupled wavelet-random forest methodology. *Journal of Hydrology* 561: 737-750. https://doi.org/10.1016/j.jhydrol.2018.04.042

Shirzad A & Safari M J S (2019). Pipe failure rate prediction in water distribution networks using multivariate adaptive regression splines and random forest techniques. *Urban Water Journal* 16(9): 653-661. https://doi.org/10.1080/1573062X.2020.1713384

Srivastava A, Sahoo B, Raghuwanshi N S & Chatterjee C (2018). Modelling the dynamics of evapotranspiration using Variable Infiltration Capacity model and regionally calibrated Hargreaves approach. Irrigation Science 36(4-5): 289-300. https://doi.org/10.1007/s00271-018-0583-y

Sterkenburg T F & Grünwald P D (2021). The no-free-lunch theorems of supervised learning. Synthese 199: 9979-10015. https://doi.org/10.1007/s11229-021-03233-1

Tabari H, Kisi O, Ezani A & Talaee P H (2012). SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semiarid highland environment. *Journal of Hydrology* 444: 78-89. https://doi.org/10.1016/j.jhydrol.2012.04.007

Tabari H, Grismer M E & Trajkovic S (2013a). Comparative analysis of 31 reference evapotranspiration methods under humid conditions. Irrigation Science 31: 107–117. https://doi.org/10.1007/s00271-011-0295-z

Tabari H, Martinez C, Ezani A & Talaee P H (2013b). Applicability of support vector machines and adaptive neurofuzzy inference system for modeling potato crop evapotranspiration. Irrigation Science 31(4): 575-588. https://doi.org/10.1007/s00271-012-0332-6

Taylor K E (2001), Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research* 106(D7): 7183–7192, https://doi:10.1029/2000JD900719

Traore S, Wang Y M & Kerh T (2010). Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone. Agricultural Water Management 97(5): 707-714. https://doi.org/10.1016/j.agwat.2010.01.002

Traore S & Guven A (2012). Regional-specific numerical models of evapotranspiration using gene-expression programming interface in Sahel. Water Resource. Management 26: 4367-4380. https://doi.org/10.1007/s11269-012-0149-3

Traore S & Guven A (2013). New algebraic formulations of evapotranspiration extracted from gene-expression programming in the tropical seasonally dry regions of West Africa. Irrigation Science 31: 1-10. https://doi.org/10.1007/s00271-011-0288-y

TSMS (2022). General directorate of state meteorology affairs. https://mevbis.mgm.gov.tr/ mevbis/ui/index.html#Workspace, accessed 15 March 2022

TSMS (2023). General directorate of state meteorology affairs. https://www.mgm.gov.tr/ veridegerlendirme/il-ve-ilceler-istatistik. aspx?k=unde- fined&m=ADANA, accessed 15 January 2024

TSMS (2024). General directorate of state meteorology affairs. https://www.mgm.gov.tr/FILES/iklim/iklim_siniflandirmalari/koppen.pdf, Accessed 12 January 2024.

Üneş F, Kaya Y Z & Mamak M (2020. Daily reference evapotranspiration prediction based on climatic conditions applying different data mining techniques and empirical equations. Theoretical and Applied Climatology 141(2020): 763-773. https://doi.org/10.1007/s00704-020-03225-0

Valipour M (2015). Investigation of valiantzas' evapotranspiration equation in Iran. Theoretical and Applied Climatology 121: 267-278. https://doi.org/10.1007/s00704-014-1240-x

Vapnik V (1995). Support-vector networks. Machine Learning 20: 273-297. https://doi.org/10.1007/BF00994018

Wang S, Lian J, Peng Y, Hu B & Chen H (2019). Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi, China. Agricultural Water Management 221: 220-230. https://doi.org/10.1016/j.agwat.2019.03.027

Wang J, Raza A, Hu Y, Buttar N A, Shoaib M, Saber K… & Ray R L (2022). Development of monthly reference evapotranspiration machine learning models and mapping of Pakistan-A comparative study. Water 14(10): 1666. https://doi.org/10.3390/w14101666

Willmott C J (1981). On the validation of models. Physical Geography 2: 184-194. https://doi.org/10.1080/02723646.1981.10642213

Yamaç S S (2021). Reference evapotranspiration estimation with kNN and ANN models using different climate input combinations in the semiarid environment. *Journal of Agricultural Sciences* 27 (2): 129-137. https://doi.org/10.15832/ankutbd.630303

Yassin M A, Alazba A A & Mattar M A (2016). Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. Agricultural Water Management 163: 110-124. https://doi.org/10.1016/j.agwat.2015.09.009

Yildirim D, Küçüktopcu E, Cemek B & Simsek H (2023). Comparison of machine learning techniques and spatial distribution of daily reference evapotranspiration in Türkiye. Applied Water Science 13(4): 107. https://doi.org/10.1007/s13201-023-01912-7

Yirga S A (2019). Modelling reference evapotranspiration for Megecha catchment by multiple linear regression. Modeling Earth Systems and Environment 5: 471-477. https://doi.org/10.1007/s40808-019-00574-2

Yurtseven I & Serengil Y (2021). Comparison of different empirical methods and data-driven models for estimating reference evapotranspiration in semi-arid Central Anatolian Region of Turkey. *Arabian Journal of Geosciences* 14: 1-28. https://doi.org/10.1007/s12517-021-08150-8

Appendix 1- Applied model parameters and pseudo codes

Model Parameters for GEP	Model Parameters for RF	Model Parameters for ANN
Number of chromosomes: 30	Number of trees: 10	Hidden layers: 100
Head size: 7-8	Maximal number of considered features: unlimited	Activation: ReLu
Genes: 3	Maximal tree depth: unlimited	Solver: L-BFGS-B
Function set: +, -, *, /, Exp, Ln, X2, 3Rt, Inv	Stop splitting nodes with maximum instances: 5	Alpha: 0.0001
Linking function: addition	Replicable training: yes	Max iterations: 200
Fitness function: RMSE		Replicable training: yes
Mutation: 0.00138		
Inversion: 0.00546		
IS transposition: 0.00546	**Model Parameters for SVM**	**Model Parameters for MLR**
RIS transposition: 0.00546	SVM type: SVM, c=1.3, Ɛ = 0.5	Parameters: Fit intercept (unchecking it fixes it to zero)
Gene transposition: 0.00277	Kernel: RBF, exp(-auto|x-y|^2)	Regularization: No regularization
One-point recombination: 0.00277	Numerical tolerance: 0.001	
Two-point recombination: 0.00277	Iteration limit: 200	
Gene recombination: 0.00277		

Box Plot

1. Initialize necessary variables and parameters: - Define a list of column identifiers to read from the Excel file (pandas). - Define a list of labels corresponding to each dataset. - Initialize dictionary (data structure) to store the data and calculated statistics.
2. Read the Excel data: - Loop through the list of column identifiers: - For each identifier, read the corresponding column from the Excel file into a variable (pandas). - Store this data in a dictionary with the corresponding label as the key.
3. Convert the data to NumPy arrays: - Loop through the labels in the dictionary: - Convert each dataset to a NumPy array and store it back in the dictionary (NumPy).
4. Define a function to calculate statistics: - The function should take a dataset as input and return the following (NumPy): - Lower Quartile - Upper Quartile - Mean - Median - Standard Deviation
5. Calculate statistics for each dataset: - Loop through the labels in the dictionary: - Apply the statistics function to each dataset. - Store the results in a dictionary with the corresponding label as the key.
6. Print the calculated statistics: - Loop through the labels in the statistics dictionary: - For each label, print the corresponding statistics.
7. Plot the boxplots: - Initialize the plotting parameters (e.g., median line style, mean marker, etc.) (matplotlib). - Prepare the data for plotting by extracting each dataset from the dictionary. - Use the prepared data to create boxplots with labels for each dataset (matplotlib). - Set the y-axis label and range (matplotlib).
- Display the plot (matplotlib).

Error Box Plot

1. Read data from the Excel file (pandas):- Load data from columns into variables: ANN1, RF2, RF3, RF4, HS, CALJH
2. Convert the loaded data into NumPy arrays (NumPy): - For each dataset (ANN1, RF2, RF3, RF4, HS, CALJH), convert the DataFrame to a flattened NumPy array.
3. Define the properties for the boxplot (standard Python dictionary): - Set the median line properties with red color and a specific line style. - Define the properties for mean markers (black edge, blue face). - Set the properties for outlier markers (red edge, specific marker size).
4. Calculate statistics for each dataset (NumPy): - For each dataset (ANN1, RF2, RF3, RF4, HS, CALJH), calculate the lower quartile, upper quartile, mean, median, and standard deviation using the `calculate_statistics` function.
5. Plot the boxplots (matplotlib): - Set the figure size for the plot. - Define the y-axis ticks from 0 to 4.1 with intervals of 0.5. - Create a boxplot for each dataset (ANN1, RF2, RF3, RF4, HS, CALJH) with appropriate labels and the defined properties (whiskers, median, mean). - Label the y-axis and range (matlplotlib).

Taylor Diagram

1. Calculate correlation coefficients between datasets (NumPy): - Compute the correlation between ANN1 and FAO56-PM. - Compute the correlation between RF2 and FAO56-PM. - Compute the correlation between RF3 and FAO56-PM. - Compute the correlation between RF4 and FAO56-PM. - Compute the correlation between HS and FAO56-PM. - Compute the correlation between CALJH and FAO56-PM.
2. Extract the relevant correlation values from the correlation matrices (NumPy): - Extract the correlation value from the ANN1-FAO56-PM matrix. - Extract the correlation value from the RF2-FAO56-PM matrix. - Extract the correlation value from the RF3-FAO56-PM matrix. - Extract the correlation value from the RF4-FAO56-PM matrix. - Extract the correlation value from the HS- FAO56-PM matrix. - Extract the correlation value from the CALJH- FAO56-PM matrix.
3. Store the extracted correlation values in a list (standard Python list): - Create a list to hold the correlation values for ANN1, RF2, RF3, RF4, HS, and CALJH.
4. Define the `TaylorDiagram` class: - Initialize the class with standard deviation (STD), figure, rectangular plot area, and label parameters. - Set up the polar transform for correlation angles (NumPy and matplotlib).
- Define correlation labels and convert them to polar angles (NumPy). - Set up the grid for the diagram (matplotlib). - Create a subplot within the figure (matplotlib). - Configure the axis labels and directions:
- Set up the top axis for the correlation coefficient with turquoise color. - Set up the left and right axes for standard deviation. - Hide the bottom axis. - Draw grid lines on the subplot. - Add reference points and standard deviation contours to the plot (matplotlib). - Initialize a list to collect sample points.

5. Define `add_sample` method in `TaylorDiagram` class: - Plot a sample point on the Taylor diagram using its standard deviation and correlation coefficient. - Append the plotted sample point to the list of sample points.
6. Define `add_contours` method in `TaylorDiagram` class: - Create a meshgrid of standard deviation and polar angles (NumPy). - Calculate RMSE (Root Mean Square Error) values (NumPy). - Draw contour lines on the Taylor diagram to represent RMSE levels (matplotlib).
7. Define `add_samples` function: - For each dataset, plot a sample point on the Taylor diagram using its standard deviation, correlation coefficient, and specified color. - Customize the marker appearance (matplotlib).
8. Define `srl` function: - Create a new figure for the Taylor diagram (matplotlib). - Initialize the `TaylorDiagram` class with observed standard deviation and other parameters. - Add contour lines to the diagram with specified colors and linewidths. - Add sample points to the diagram using the `add_samples` function. - Create a legend for the plotted points. - Display the Taylor diagram (matplotlib).
9. Define the data for the diagram: - Set the observed standard deviation using statistics from the FAO56-PM dataset. - Define lists for the standard deviations, correlation coefficients, and labels for each dataset (ANN1, RF2, RF3, RF4, HS, Cal_JH). - Specify colors for each data point.
10. Call the `srl` function: - Pass the observed standard deviation, standard deviation list, correlation coefficient list, label list, color list, and file name to the function.

Correlation Matrix

1. Read data from the Excel file (pandas): - Load data from columns into variables Tmax, Tmin, RH, Rs, U2, FAO56-PM.
2. Convert the loaded data into NumPy arrays and flatten them (NumPy): - For each dataset (Tmax, Tmin, RH, Rs, U2, FAO56-PM), convert the DataFrame to a flattened NumPy array.
3. Combine the flattened arrays into a single 2D array (NumPy): - Stack the NumPy arrays column-wise into a single 2D array.
4. Calculate the correlation matrix for the combined data (NumPy): - Compute the correlation matrix using the combined 2D array, treating each column as a variable.
5. Loop through the correlation matrix to process each pair of correlations (standard Python loop): - For each pair of variables (i, j) in the correlation matrix
6. Create a figure for the heatmap (matplotlib): Set the figure size to 8 by 6 inches size.
7. Generate a heatmap to visualize the correlation matrix (seaborn): - Plot the correlation matrix as a heatmap using the following parameters: - Annotate each cell with the correlation coefficient value, formatted to 3 decimal places. - Use the "Pink-Green" color map for the heatmap - Set annotation font size to 15.
- Define x-axis and y-axis tick labels with specific variables. - Add a label to the color bar with the text 'Correlation Coefficient'.
8. Adjust the font size and rotation of the tick labels on the axes (matplotlib):
- Set the x-axis tick labels with no rotation and a font size of 10. - Set the y-axis tick labels with a 90-degree rotation and a font size of 10.
9. Customize the color bar (matplotlib): - Access the color bar from the heatmap and adjust its label font size to 12. - Adjust the tick label font size on the color bar to 12.
10. Display the heatmap (matplotlib): - Show the generated heatmap.

Time Series

1. Set the file path for the Excel file.
2. Read data from specified columns in the Excel file (pandas): - Columns into FAO56-PM, ANN1, RF2, RF3, RF4, HS, CALJH. - Columns into DAY, MONTH, YEAR.
3. Convert the data to flattened NumPy arrays (NumPy): - Flatten DAY, MONTH, YEAR into nDAY, nMONTH, nYEAR. - Flatten ANN1, RF2, RF3, RF4 into nANN1_num, nRF2_num, nRF3_num, nRF4_num. - Flatten HS, CALJH, FAO56-PM into nHS_num, nCALJH_num, nFAO56_PM_num.
4. Define the `DataRecord` class: - Initialize the class with attributes: value, day, month, and year. - Implement a `__repr__` method to return a string representation of the object, displaying the value, day, month, and year.
5. Create lists of `DataRecord` objects for each dataset: - For each dataset (ANN1, RF2, RF3, RF4, FAO56_PM, HS, CALJH): - Use list comprehension to create `DataRecord` objects by zipping together the corresponding values, days, months, and years. - Store the resulting list of `DataRecord` objects in the appropriate variable (e.g., `ANN1_records`, `RF2_records`, etc.).
6. Define the `plot_data` function: - Accept parameters: a list of `DataRecord` objects (`records`), a plot title (`title`), and an optional step for x-axis labels.
7. Sort the records by date: - Convert `day`, `month`, and `year` attributes of each record into a `datetime` object. - Sort the records by the `datetime` object. - Extract the formatted dates as strings and the corresponding values into separate lists.
8. Set up the plot (matplotlib): - Create a figure with specified dimensions. - Plot the sorted dates against the values with a blue line. - Set the plot title, x-axis label, and y-axis label.
9. Customize x-axis labels: - Identify dates that are the 1st of each month. - Format these dates as "Month Year" for x-axis labels. - Set the x-axis labels to the formatted dates with a 90-degree rotation.
10. Display the grid and the plot (matplotlib).
11. Use `plot_data` function to visualize different datasets: - Call `plot_data` for each dataset (`ANN1_records`, `RF2_records`, `RF3_records`, `RF4_records`, `FAO56_PM_records`, `HS_records`, `CALJH_records`) with appropriate titles and step values.