Comparative analysis of Kernel-based versus BFGS-ANN and deep learning methods in monthly reference evapotranspiration estimation

Mohammad Taghi SATTARI [1,2,3], Halit APAYDIN[3], Shahab S. BAND[4], Amir MOSAVI[5,6,7], Ramendra PRASAD[8]

[1]Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz 51666, Iran
[2]Institute of Research and Development, Duy Tan University, Danang 550000, Vietnam.
[3]Department of Agricultural Engineering, Faculty of Agriculture, Ankara University, Ankara 06110, Turkey
[4]Future Technology Research Center, National Yunlin University of Science and Technology, Douliou, Yunlin 64002, Taiwan
[5]Faculty of Civil Engineering, Technische Universität Dresden, 01069 Dresden, Germany
[6]Thuringian Institute of Sustainability and Climate Protection, 07743 Jena, Germany
[7]John von Neumann Faculty of Informatics, Obuda University, 1034 Budapest, Hungary
[8]Department of Science, School of Science and Technology, The University of Fiji, Lautoka, Fiji

Correspondence to: Mohammad Taghi SATTARI (mtsattar@tabrizu.ac.ir and mohammadtaghisattari@duytan.edu.vn), Shahab S. Band (shamshirbands@yuntech.edu.tw)

Abstract. Proper estimation of the magnitude of reference evapotranspiration (ET_0) ~~amount~~ is an indispensable matter for agricultural water management ~~in~~ for ~~the efficient use of water~~efficient water use. ~~The~~ This ~~aim of study is to estimate the amount of ET_0 with a different machine and deep learning methods by using minimum meteorological parameters in the Corum region~~study aims to estimate the amount of ET_0 with ~~a different~~ machine and deep learning ~~methods~~approaches by using minimum meteorological parameters in the Corum region, which ~~is~~ has an arid and semi-arid climate ~~with an~~and regardes as an important agricultural center of Turkey. In this context, monthly averages of meteorological variables ~~of average~~ *i.e.*, ~~,~~ maximum and minimum temperature, sunshine duration, wind speed, average, maximum, and minimum relative humidity are used as input ~~sdata monthly~~. Two different kernel-based (Gaussian Process Regression (GPR) and Support Vector Regression (SVR)) methods, BFGS-ANN, and Long short-term memory ~~models~~ (LSTM) models were used to estimate ET_0 amounts in 10 different combinations. ~~According to t~~The results ~~obtained,~~showed that all four methods ~~used~~ predicted ET_0 amounts ~~with~~in acceptable accuracy and error levels. BFGS-ANN model showed higher success (R^2:0.9781) than the others. In kernel-based GPR and SVR methods, Pearson VII function-based universal kernel was the most successful (R^2:0.9771) ~~kernel function~~. Besides, the scenario ~~that is~~ related to temperature in all scenarios used, including average temperature, maximum and minimum temperature, and sunshine duration, gave the best results. The second-best scenario was ~~the one that covers~~ with only the sunshine duration as the input ~~. In this case, the ANN (~~to the BFGS-ANN~~)~~ which ~~model, which is optimized with the~~

BFGS method that uses only the sunshine duration, can be estimated ET_0 with thehaving 0.971 correlation coefficient of 0.971. ET0 Conclusively, this study shows the better efficacy of the BFGS in ANN for enhanced performance of the ANN model applied in ET_0 estimation for arid and semi-arid drought-prone regions.without the need for other meteorological parameters.

1 Introduction

It can be noted that water is one of the most important elements on earth that is used by every organism. Without water, it will become almost impossible to survive. It may easy to survive for a few days without getting enough meal but it is completely unable to survive without drinking water. For plants, water is the main requirement for its growth. It means that it is not easy to grow plants without giving them proper water. Furthermore, one of the most important facts is that water is helping the living organisms to maintain their body temperature. The next fact is that water also helps to absorb the nutrients efficiently. It is also used in dissolving nutrients. Another reason is that water is responsible for refraining from different kinds of diseases. With the help of water, it will become easy to maintain the body shape more effectively without any sort of problem.

Accurate estimation of reference crop evapotranspiration (ET_0) and crop water consumption (ET) is an important essential process for the management ofmanaging water in the agricultural sector in arid and semi-arid climatic conditions where water is scarce and valuable. Although ET_0 is a complex element of the hydrological cycle, it is also an important component of hydro-ecological applications and the water management in the agricultural sector. ET_0 is important essential in increasing irrigation efficiency, irrigation planning, and reuse of water.The ET_0 estimation of ET_0i isan essential tool for critical tool in the forcible management of irrigation and hydro-meteorological studies on the basin and national scale (Pereira et al. 1999, Xu and Singh 2001, Anli 2014), since knowledge of ET_0 would allow for reduced water wastage, increased irrigation efficiency, proper irrigation planning, and reuse of water.

In general, the equations that calculate ET_0 values are very complex, nonlinear, and contains randomness and all in all has a number of underlying assumptions. The results obtained from these equations differ greatly with the measured values. ET_0 is considered a complex and nonlinear phenomenon that interacts with water, agriculture, and climate sciences. It is difficult to calculateemulate such a phenomenon by experimental and classical mathematical methods. There are about twenty well-known methods for predictingestimating ET_0 based on different meteorological variables and assumptions. Many of these methods commonly require input and output datasets. The Penman-Monteith (FAO56PM) method proposed by FAO is recommended to estimate ET_0, as it usually gives the bestusable results in different climatic conditions (Hargreaves and Samani 2013, Rana and Katerji 2000, Feng et al. 2016, Nema et al. 2017). Cobaner et al. (2016) modified the Hargreaves-Samani (HS) equation used in the determination of ET_0. Solving these equations and finding the correct value of the parameters requires sophisticated programs for employment of differential equations, which require rigorous optimization methods together with broad range of spatio-temporal good quality and accurate input data with knowledge of initial conditions (Prasad et al. , 2017).

On the other hand, to be used can be difficult. However, the developments in artificial intelligence (AI) methods and the increase in the accuracy of the estimation results have increased the desire for these AI methods. The AI models offer a number

of advantages including; their ease of development compared to physically based models; does not require underlying boundary conditions or other assumptions or initial forcings; and has the ability to operate at localized positions (Prasad et al. , 2020). Consequently, ~~Ma~~many studies have been reported to ~~introduce~~ have applied AI approaches for ET_0 estimations ~~and to test these methods~~. ~~ET_0 is considered a complex and nonlinear phenomenon that interacts with water, agriculture, and climate sciences. It is difficult to calculate such a phenomenon by experimental and classical mathematical methods. There are about twenty well known methods for predicting ET_0 based on different meteorological variables and assumptions. Many of these methods commonly require input and output datasets. The Penman-Monteith (FAO56PM) method proposed by FAO is recommended to estimate ET_0, as it usually gives the best results in different climatic conditions (Hargreaves and Samani 2013, Rana and Katerji 2000, Feng et al. 2016, Nema et al. 2017).~~

~~Today, it is seen that a~~Artificial intelligence techniques based on machine learning (ML) ~~can~~ has been successfully utilized in predicting complex and nonlinear~~the~~ processes ~~showing a complex and nonlinear structure. It is witnessed that scientists frequently use artificial intelligence methods based on excess data~~ in natural sciences, especially hydrology (Koch et al. , 2019, Prasad, Deo, 2017, Solomatine, 2002, Solomatine and Dual, 2003, Yaseen et al. , 2016, Young et al. , 2017). Thus, methods such as ML, deep learning, and artificial intelligence have gained popularity in ~~calculating~~ estimating and predicting ET_0~~, as can be seen from the literature search~~.

The artificial neural networks (ANN)– has been the widely used ML model till date. Sattari et al. (2013) used the backpropagation algorithm of ~~artificial neural network (~~ANN~~)~~ and tree-based M5 model to estimate the monthly ET_0 amount by employing a climate dataset (air temperature, total sunshine duration, relative humidity, precipitation and wind speed) in the Ankara region and compared ~~. As the input parameter, the monthly air temperature, total sunshine duration, relative humidity, precipitation, wind speed, and monthly time index were used, while. In contrast,~~ the estimated ET_0 with ~~computed by~~ FAO56PM computations~~was used as output for both approaches~~. The results revealed that the ~~neural network~~ANN approach gives better results ~~with this data set~~. In another study, ~~Citakoglu et al. (2014) predicted the monthly average ET_0 using the ANN and adaptive network based fuzzy inference system (ANFIS) techniques. Different combinations of long term average monthly climate data of wind speed, air temperature, relative humidity, and solar radiation were recorded at the stations used as input data. According to the results of the study; Although ANFIS is slightly more successful than ANN, it has been found that both methods can be used in estimating the monthly mean ET_0.~~ Pandey et al. (2017) in their study, ML techniques for ET_0 estimation using limited meteorological data; evaluated evolutionary regression (ER), ANN, multiple nonlinear regression (MLNR), and SVM and found ~~. According to the results of the study,~~ the ANN FAO56PM model performed ~~ing~~ better. In their study, Nema et al. (2017) studied the possibilities of using ANN to increase monthly evapotranspiration prediction performance in the humid area of Dehradun. They developed different ANN models, including combinations of various training functions and neuron numbers, and compared them with ET_0 calculated with FAO56PM. They found that the ANN trained by the Levenberg-Marquardt algorithm with 9 neurons in a single hidden layer made the best estimation performance in their case. ~~Reis et al. (2019) also used Hargreaves Samani,~~ The ANN, with multiple linear regression (MLR), ~~and~~ ELM and Hargreaves Samani models were tested by Reis et al. (2019) to predict ET_0 in the presence of temperature data

in the Verde Grande River basin, ~~in southeastern~~ Brazil. The ~~results~~study revealed that ~~artificial intelligence~~AI methods have superior performance~~ity~~ over other models. Abrishami et al. (2019) estimated the amount of daily ET_0~~ET~~ for wheat and corn ~~by the~~using ANN ~~method. Leaf area index, daily climate data, and plant height were employed as the input parameters. The results indicated the~~and found proper ~~capacity~~ and acceptable performance of ANNs with two hidden layers. However, some studies showed slightly better performance of other models. Citakoglu et al. (2014) predicted monthly average ET_0 using the ANN and adaptive network-based fuzzy inference system (ANFIS) techniques using combinations of long-term average monthly climate data such as wind speed, air temperature, relative humidity, and solar radiation as inputs. They found ANFIS to be slightly better than ANN, yet found that both methods can be successfully used in estimating the monthly mean ET_0. Likewise, ANN and ANFIS models by employing the Cuckoo search algorithm (CSA) was applied by Shamshirband et al. (2016) using data from twelve meteorological stations in Serbia. The results showed that the hybrid ANFIS-CSA could be employed for high-reliability ET_0 estimation.

Despite ANNs being universal approximators having the ability to approximate any linear or nonlinear system without being constrained to a specific form, it has some inherent disadvantages. Slow learning speed, over-fitting and constrained in local minima with relatively tedious to determine key parameters, such as training algorithms, activation functions and hidden neurons. ~~Generally, ANN methods have disadvantages.~~ These inherent ~~disadvantages are sometimes due to the~~ structural problems ~~of the method and~~ sometimes makes ~~it~~s difficulty in adopting for application~~s~~. However, despite all the disadvantages, it is still a preferred method in all branches of science and especially in hydrology. Having said that, in this study, the ANN is benchmarked with other comparative models. One such model is support vector machine (SVM) developed by Vapnik (2013). SVMs has good generalization ability since it utilizes the concept of structural risk minimization hypothesis in ~~minimiinge~~minimising both empirical risk and the confidence interval of the learning algorithm. Due to the fact that the solid mathematical foundation of statistical learning theory giving it an advantage, the SVMs have been preferred in a number of studies and produced highly competitive performances in real-world applications (Quej et al. , 2017). ~~In general, the capability of different ANN methods is discussed in terms of various disadvantages such as the large number of hidden layers, slow learning speed, overfitting and stuck to local minimums.~~ Subsequently, ~~At the same time, machine learning methods have some deficiencies such as sometimes difficult to use, high memory requirement and large amount of computation time in the learning process.~~Wen et al. (2015) predicted ~~the~~ daily ET_0 ~~amount by the~~via SVM ~~method~~, using a limited climate dataset in Ejina basin, ~~the region of~~ China~~'s~~ ~~extremely arid Ejina basin. In the study, the~~using the highest and lowest air temperatures, daily solar radiation and wind speed values ~~were used~~ as model input~~s~~ and FAO56PM results as model output. ~~Besides, t~~The ~~performance of the SVM method~~SVM method's performance was compared to ANN and empirical techniques, including Hargreaves, Priestley-Taylor, and Ritchie, which revealed that ~~. The results indicated that among these models,~~ the SVM ~~provided the highest~~recorded better performance. ~~Gocic et al. (2015) determined ET_0 using the FAO56PM method using data from 1980-2010. They used four different computational intelligence methods: Genetic programming (GP), SVM-Firefly algorithm (SVM-FFA), ANN, and SVM-Wavelet. While this study shows that the hybrid SVM-Wavelet provides the highest accuracy for the estimation of ET_0; It is determined that SVM-Wavelet followed by the hybrid SVM-FFA provides the higher~~

correlation coefficient compared to ANN and GP.Zhang et al. (2019) examined SVM's success in ET_0 estimation. The results were compared with Hargreaves, FAO-24, Priestley-Taylor, McCloud, and Makkink. SVM was determined to be the most successful model. However, SVM also has several drawbacks, such as high computational memory requirement as well being computational exhaustive as a large amount of computing time during learning process is necessary.

In order to overcome the disadvantages of these two widely accepted approaches (ANN and SVM), many new modeling techniques have been proposed in recent years. For instance, the two state-of-the-art machine learning techniques, namely Gauss Process Regression (GPR) and long short-term memory (LSTM) are also being recently trialled in the hydrologic time series modeling and forecasting applications. Following the newer developments, Cobaner et al. (2016) modified the Hargreaves Samani (HS) equation used in the determination of ET_0. They used measured data on 275 meteorological observation stations in seven different regions of Turkey. It was concluded that in regions where meteorological measurements are insufficient, it can be used more successfully than the original HS equation for the prediction of ET_0. Feng et al. (2016) used a trained extreme learning machine (ELM), the integration of genetic algorithm ANN (GA-ANN), and wavelet neural networks (WNN) to predict ET_0 in a humid basin of Southeast China. The results were compared with empirical ET_0 (temperature based i.e., Hargreaves and modified Hargreaves and radiation based i.e., Makkink, Priestley Taylor, and Ritchie). Results found that ELM and GANN are better than WNN models. It has been determined that temperature based ELM and GANN have higher accuracy than Hargreaves and modified Hargreaves, and radiation based ELM and GANN have higher performance than Makkink. Shamshirband et al. (2016) estimated ET_0 with ANN and ANFIS models by employing the Cuckoo search algorithm (CSA). In the study, the models of monthly climate datanthly climate data models from the twelve meteorological stations in Serbia were used as input in the 1983-2010 period. For the study, they chose the FAO56PM equation as the ET_0 equation. The results showed that the hybrid ANFIS-CSA could be employed for high reliability ET_0 estimation. Pandey et al. (2017) in their study, ML techniques for ET_0 estimation using limited meteorological data; evaluated evolutionary regression (ER), ANN, multiple nonlinear regression (MLNR), and SVM. According to the results of the study, the ANN FAO56PM model performed better. In their study, Nema et al. (2017) studied the possibilities of using ANN to increase the performance of monthly evaporation predictionmonthly evapotranspiration prediction performance in the humid area of Dehradun. They developed different ANN models, including combinations of various educational functions and neuron numbers, and compared them with ET_0 calculated with FAO56PM. According to the results, the ANN trained by the Levenberg-Marquardt algorithm with nine neurons in a single hidden layer made the best estimation performance. Feng et al. (2017) estimated the daily ET_0 using temperature data by employing ELM and generalized regression neural network (GRNN) at six meteorological stations in the southwestern Sichuan basin in China. In the study, ET_0 was determined by FAO56PM and Hargreaves (HG) model for comparison. Two different scenarios were evaluated for ET_0 estimation. ELM and GRNN models gave acceptable results. Fan et al. (2018) used four models based on SVM, ELM, and decision tree to predict daily reference evapotranspiration, using a limited number of meteorological data in various climates of ChinaChina's various climates. Considering the complexity level and estimation performance, Extreme Gradient Boosting (XGBoost) and Gradient Boosting Decision Tree (GBDT) models found to be suitable methods for the prediction ofwere suitable methods for predicting daily

ET_0. Banda et al. (2018) estimated the daily ET_0 for the Bulawayo Goetz region from climate data using neurocomputing methods and compared them with experimental methods. According to the results of the studystudy results, neurocomputing calculation techniques were found to be superior to empirical methods. Wu et al. (2019) estimated ET_0 using ANN, Random Forest (RF), GBDT, XGBoost, Multivariable Adaptive Regression Spline (MARS), Support Vector Regression (SVR) in the presence of temperature data. The results of the employed models were compared with those of the temperature-based Hargreaves-Samani equation. The research found that the success of ML techniques differs depending on the scenarios. Tree-based models (RF, GBDT, and XGBoost) provided higher prediction performance than the rest of the models. Reis et al. (2019) also used Hargreaves-Samani, ANN, multiple linear regression (MLR), and ELM models to predict ET_0 in the presence of temperature data in the Verde Grande River basin in southeastern Brazil. The results revealed that artificial intelligence methods have superiority over other models. Wang (2019) estimated the amount of ET_0 using PSO and least squares (LSSVM) models, based on the Hilbert Huang transformation, using daily data from 2000-2009 of China's Hetian Xinjiang meteorological station. Under the condition of the working areaworking area's condition, the accuracy of the ML techniques was found to be higher than those of the empirical models. Abrishami et al. (2019) estimated the amount of daily ET for wheat and corn by the ANN method. Leaf area index, daily climate data, and plant height were employed as the input parameters. The results indicated the proper capacity and acceptable performance of ANNs with two hidden layers. Zhang et al. (2019) examined SVM's success in ET_0 estimation. The results were compared with Hargreaves, FAO-24, Priestley-Taylor, McCloud, and Makkink. SVM was determined to be the most successful model. Saggi et al. (2019) used deep learning and ML techniques in Punjab's Hoshiarpur and Patiala regions to determine daily ET_0. In the study, supervised learning algorithms, including deep learning-multilayer sensors (DL), random forest (RF), generalized linear model (GLM) and gradient boosting machine (GBM) were tested. In the study, it was determined that the DL model offers high performance for daily ET_0 modeling. Shabani et al. (2020) used ML methods, including Gauss Process Regression (GPR), random forest (RF), and SVR, in the presence ofwith meteorological data inputs related to Iran to estimate evaporation (PE) in Iran and compare it with the measured values. The results showed that PE can be estimated usingand found that ML methods with ahad high performance and even with a small number of meteorological parameters that can be easily measured. In a recent study, deep learning and ML techniques determine daily ET_0 have been explored in Punjab's Hoshiarpur and Patiala regions, India (Saggi et al., 2019). They found that supervised learning algorithms such as deep learning-multilayer sensors (DL) model offers high performance for daily ET0 modeling. However, to the best of author's knowledge there have been very few attempts to test the practicability and ability of these two advanced approaches (LSTM and GPR) for ET_0 modeling and prediction. In addition, many studies included solar radiation in the modelling process, yet did not include sunshine hours in the modelling and will be dealt with in this study. With recent developments in ML methods with the use of deep learning techniques such as LSTM in water engineering together with technical developments in computers and the emergence of relatively comfortable coding languages this study explores the application of different deep learning (LSTM) and other machine learning methods (ANN, SVM and GPR) in estimation of ET_0 with the aim to shed light on future research and to determine effective modelling approaches relevant to this field. ET_0 is one of the essential elements in water, agriculture, hydrology, and meteorology studies, and its accurate estimation has been

an open area of research due to ET_0 being a complex and nonlinear phenomenon. Hence, robust deep learning and ML approaches including LSTM, ANN, SVM and GPR methods need to be aptly tested. As a result this study has three important goals; i) to estimate the amount of ET_0 using deep learning and machine learning methods, i.e., GPR, SVR, ANN employing Broyden–Fletcher–Goldfarb–Shanno (BFGS-ANN) learning algorithm, and LSTM n Corum conditions with a total annual rainfall of 427 mm classed as arid and semi-arid climatic region; ii) to investigate the effect of different kernel functions of the SVR and GPR models on the performance of ET_0 estimation and; iii) to determine the model that provides the highest performance with the least meteorological variable requirement for the study. A ~~Feng et al. (2017) estimated the daily ET_0 using temperature data by employing ELM and generalized regression neural network (GRNN) at six meteorological stations in the southwestern Sichuan basin in China. In the study, ET_0 was determined by FAO56PM and Hargreaves (HG) model for comparison. Two different scenarios were evaluated for ET_0 estimation. ELM and GRNN models gave acceptable results.Wu et al. (2019) estimated ET_0 using ANN, Random Forest (RF), GBDT, XGBoost, Multivariable Adaptive Regression Spline (MARS), Support Vector Regression (SVR) in the presence of temperature data. The results of the employed models were compared with those of the temperature based Hargreaves Samani equation. The research found that the success of ML techniques differs depending on the scenarios. Tree based models (RF, GBDT, and XGBoost) provided higher prediction performance than the rest of the models.~~

~~Corum region, with a total annual rainfall of 427 mm,~~With a total annual rainfall of 427 mm, the Corum region ~~can be classified as an arid and semi-arid region of Turkey.~~ Pproper prediction of reference evapotranspiration ~~is~~ would be ~~important~~ vital ~~to manage~~in managing limited water resources for optimum agricultural production. ~~ET_0 is one of the most important~~ essential ~~elements in water, agriculture, hydrology, and meteorology studies, and its determination is a very important~~fundamental ~~process. ET_0 modeling is called a complex and nonlinear phenomenon according to the interactions of climate factors. For this reason, ML, deep learning, and artificial intelligence methods have gained importance and popularity in this field.~~

~~Generally, ANN methods have disadvantages. These disadvantages are sometimes due to the structural problems of the method and sometimes its difficulty in application. However, despite all the disadvantages, it is still a preferred method in all branches of science and especially in hydrology. In general, the capability of different ANN methods is discussed in terms of various disadvantages such as the large number of hidden layers, slow learning speed, overfitting and stuck to local minimums. At the same time, machine learning methods have some deficiencies such as sometimes difficult to use, high memory requirement and large amount of computation time in the learning process.~~ ~~In such a case, some recent developments are seen in ANN methods and the use of deep learning techniques such as LSTM in water engineering. However, technical developments in computers and the emergence of relatively comfortable coding languages such as Phyton have enabled some deficiencies to be overcome. In this study, it is thouht that using different deep learning, machine learning and ANN methods in estimation of ET_0 can shed light on future research and help in determining more effective models in this field.~~

~~Recently, there have been large scale developments in both computer software and hardware and data collection. Due~~Thanks ~~to these technologies, large amounts of data can be easily measured, trained with the help of powerful computers and software, and can make more accurate predictions than conventional methods. Therefore, the demand for these AI methods is increasing~~

day by day and in all branches of sciencescience branches. This study has three important goals. In Corum conditions, which are is arid and semi-arid climate, the amount of ET_0 was estimated by four machines, and deep learning methods, i.e., (GPR, SVR, hybrid ANN optimized with the Broyden – Fletcher – Goldfarb – Shanno approach (BFGS-ANN), and the long short-term memory (LSTM)) method and the methods used were compared for the first time. Secondly, the effect of different kernel functions of the SVR and GPR models on the performance of ET_0 estimation is investigated. Finally, determining the model that provides the highest performance with the least meteorological variable requirement for the study.

2 Study area and dataset used Material and method2.1 Study area and dataset used

The surface area of CorumCorum's surfaceencompasses an area isof 1 278 381 ha, of which 553 011 ha, or 43%, is agricultural land (Figure 1). Its population is 525 180 and 27% of it lives in rural areas. The water resource potential of the citycity's water resource potential is 4 916 hm^3/year and 84 988 ha of agricultural land is being irrigated. The main agricultural products are wheat, paddy, chickpeas, onions, walnuts, and green lentils. This study was conducted using the monthly meteorological values data including highest and lowest temperature, sunshine duration, wind speed, average, highest, and lowest relative humidity measured from January 1993 - December 2018 in Corum province (Anonymous, 2017) as model inputs leading to 312 months. Data used monthly are average, highest and lowest temperature, sunshine duration, wind speed, average, highest, and lowest relative humidity. 200 Two hundred of the total 312 months200 monthspoints of data were used for training, and the remaining 112 were used for testing. Statistics of the data used are given in Table 1. During the training period, the daily average, highest, and lowest temperature averages are 10.80, 18.27, and 4.02^{0}C, respectively. AThe average sunshine duration in the region is 6.29 hours, wind speed is 1.72 m/s, and mean humidity is 70.41-%. The lowest skewness coefficient was found in RHmax with -0.64 and the highest in RHmin parameter with 0.35. The lowest kurtosis coefficient has Tmean with -1.24 and the highest with 1.12 by RHmax parameter. The highest variation was observed in RHmin with 140.40 and the lowest in sunshine duration with 0.18. SmilarlySimilarly in the test period, the daily average, highest, and lowest temperature averages are 11.44^{0}C, 18.60^{0}C and 4.89^{0}C, respectively. AThe average sunshine duration in the region is 5.74 hours, wind speed is 1.64 m/s, and mean humidity is 68.08-%. The lowest skewness coefficient was found in RHmax with -0.53 and the highest in RHmin parameter with 0.75. The lowest kurtosis coefficient has Tmean with -1.25 and the highest with -0.37 by RHmax and RHmin parameters. The highest variation was observed in RHmin with 202.50 and the lowest in sunshine duration with 0.16. The skewness and kurtosis coefficients in the train and the test period are similar in all parameters except the maximum relative humidity. The frequency distributionsies of meteorological data of the study area are given in Figure 12 which conforms with the distribution statistics. As it is understood from the figure, the dependent variable ET_0 values do not conform to the normal distribution.

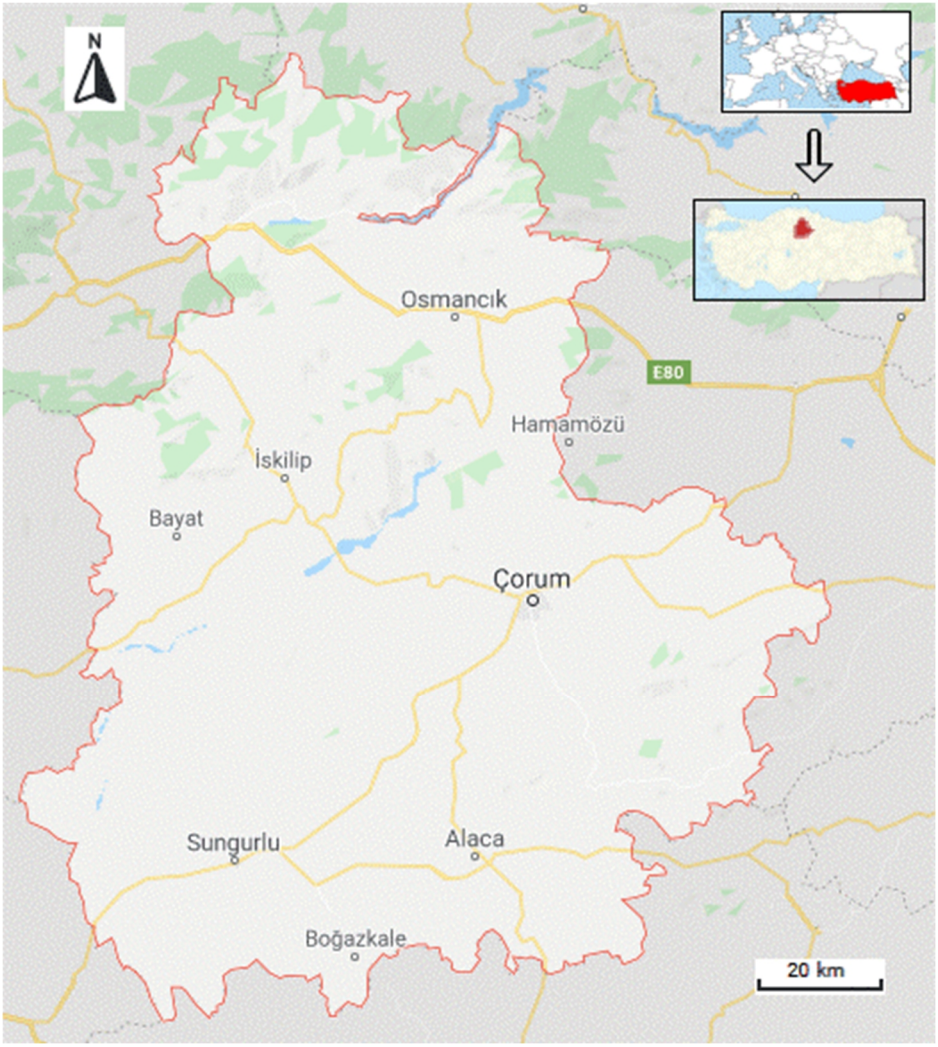


Figure 1. Location of study area, Corum Province, Turkey (maps.google.com)

Table 1. Basic statistics of the data used in the study during the training and testing periods

Period	Statistic	Tmean (^{0}Ce)	Tmax (^{0}Ce)	Tmin (^{0}Ce)	*n* (hr)	*U* (m/s)	RHmean (%)	RHmax (%)	RHmin (%)	ET_0 (mm/month)
Training data set	Minimum	-6.18	-1.27	-11.3	1	0.95	51.6	66.87	21.51	11.76
	Maximum	25.06	35.44	14.75	11.97	2.69	94.74	98.93	82.83	185.59
	Mean	10.80	18.27	4.02	6.29	1.72	70.41	87.76	47.48	79.15
	Stdev	8.00	9.32	6.34	2.96	0.42	8.02	5.39	11.88	52.64

	Skewness	-0.09	-0.15	-0.13	0.06	0.13	0.16	-0.64	0.35	0.34
	Kurtosis	-1.24	-1.21	-1.06	-1.25	-0.85	-0.37	1.12	-0.48	-1.29
	Coefficient of variation	63.75	86.50	40.02	8.72	0.18	63.97	28.86	140.40	2756.72
	Number of records	200	200	200	200	200	200	200	200	200
Testing data set	Minimum	-4.25	1.08	-9.21	0.83	0.7	45.8	72.06	19.03	13.99
	Maximum	25.06	34.85	15.63	10.87	2.45	94.07	99.83	80.12	180.53
	Mean	11.44	18.60	4.89	5.74	1.64	68.08	90.09	40.53	79.21
	Stdev	7.82	9.17	6.23	2.92	0.39	11.23	6.21	14.17	53.02
	Skewness	-0.04	-0.15	-0.03	0.08	0.08	0.25	-0.53	0.75	0.36
	Kurtosis	-1.25	-1.20	-1.12	-1.23	-0.65	-0.74	-0.37	-0.37	-1.27
	Coefficient of variation	61.68	84.89	39.20	8.60	0.16	127.17	38.90	202.50	2836.65
	Number of records	112	112	112	112	112	112	112	112	112

NB: T: Temperature, n: Sunshine duration, *U*: Wind speed, RH: Relative humidity

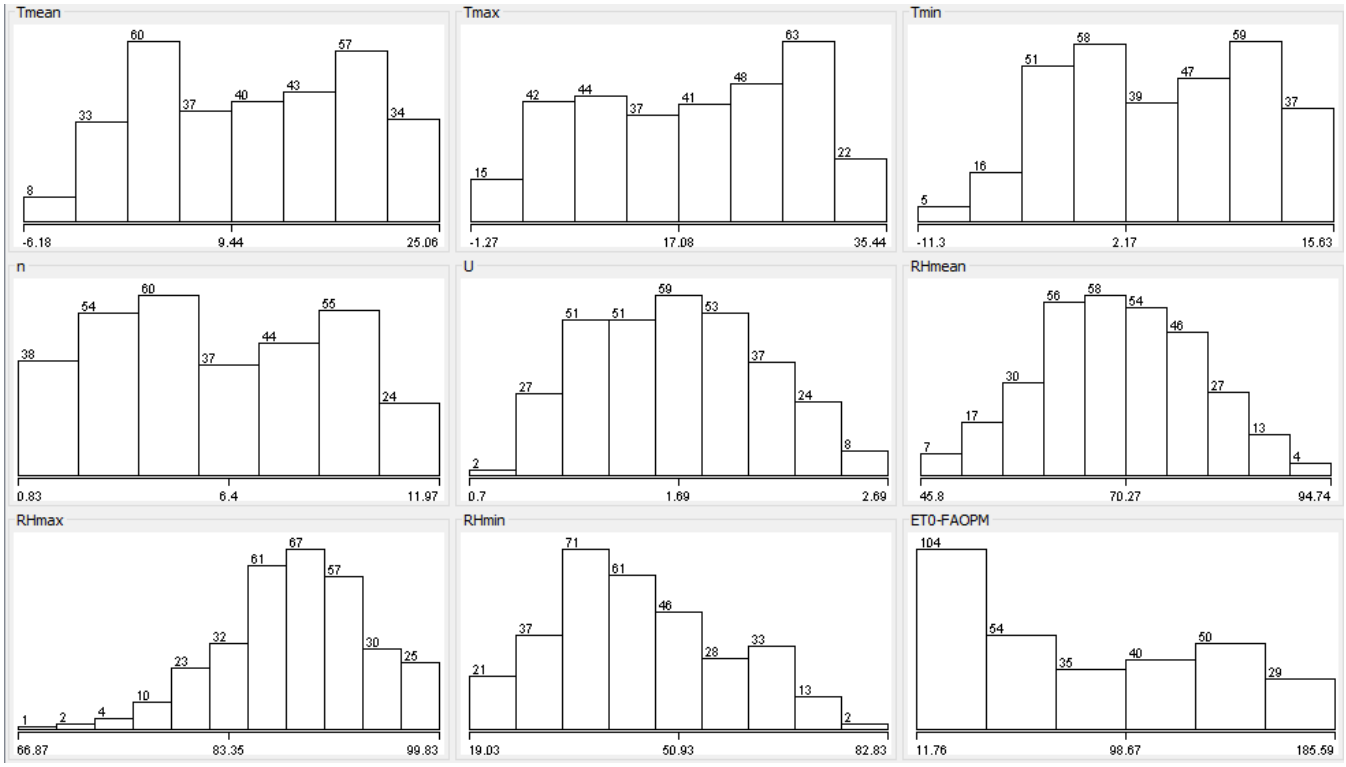


Figure ~~1~~2. ~~Frequencies of the variables used~~Frequency distributions of meteorological input data sets conforming with the distribution statistics.

To determine the meteorological factors employed in the model and the formation of scenarios, the relationship between ET_0 and other variables were calculated as~~is~~ given in Figure ~~2~~3. Input determination is an essential component of model development as irrelevant inputs are likely to worsen the model performances (Hejazi and Cai, 2009, Maier and Dandy, 2000, Maier et al. , 2010), while a set of carefully selected inputs could ease the model training process and increase the physical representation whilst providing a better understanding of the system (Bowden et al. , 2005). The Sunshine duration in this study was very highly correlated with ET_0 (R^2 = 0.92) together with the variables Tmean, Tmax and Tmin were all highly correlated (R^2 > 0.8). The RH mean was the least correlated variable (R^2 = 0.24) in this study. As can be understood visually, the meteorological variables associated with temperature and especially the sunshine duration have a high correlation with ET_0. Considering these relationships, ~~10~~ ten different input scenarios were created, and the effect of meteorological variables on ET_0 estimation was evaluated. Table 2 gives the meteorological variables used in each scenario. While all parameters were taken into account in the first scenario, the ones that could affect ET_0 more in the following scenarios were added ~~to the~~in the respective scenarios.

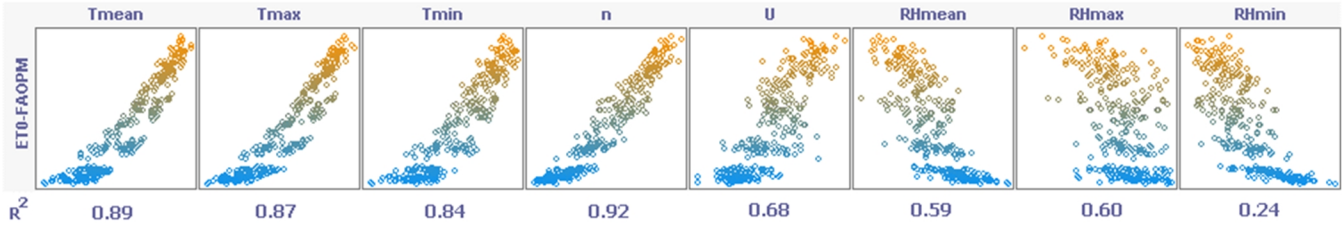


Figure ~~2~~3. Scatter plot showing the ~~C~~correlation between ET_0 and the independent variable. Coefficient of determination has been added for clarity.~~s~~

Table 2. ~~Variables used in~~ Illustrates the scenarios developed in this study with respective inputs in respective scenarios.

Scenario	Inputs
1 (All Variables)	TMean, TMax, TMin, *n*, *U*, RHMax, RHMin, RHMean
2	TMean, *n*, *U*, RHMean
3	TMax, *n*, RHMax
4	TMax, *n* , *U*
5	TMean, TMax, TMin, n
6	*n*, *U*, RHMax
7	*n*, RHMax
8 (Highest R^2)	*N*
9	TMin
10	TMax

3 Methods

3.1 Calculation of ET_0

The United Nations, Food and Agriculture Organization (FAO) ~~have~~ recommend~~ed~~ Penman-Monteith (PM) equation (Eq.1) to calculate evapotranspiration of reference crop (Doorenbos and Pruitt, 1977). Although the PM equation is much more complex than the other equations, it has been formally explained by FAO. The equation has two main features: (1) It can be

used in any weather conditions without local calibration, and (2) the performance of the equation is based on the lysimetric data in an approved spherical range (Allen et al. 1989). ~~Requirement~~ The requirement for ~~the~~ many ~~of~~ meteorological factors can be defined as the main problem. However, ~~in many countries, there is still no equipment to record these parameters correctly~~there is still no equipment to record these parameters correctly in many countries, or data is not regularly recorded (Gavili et al., 2018).

$$ET_0 = \frac{0.408\,\Delta\,(R_n - G) + \gamma \frac{900}{T+273}\,u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \qquad \text{(Eq.1)}$$

Where

ET_0 refers to the reference evapotranspiration [mm day^{-1}],

G refers the soil heat flux density [MJ m^{-2} day^{-1}],

u_2 refers to the wind speed at 2 m [m s^{-1}],

e_a refers to the actual vapour pressure [kPa],

e_s refers to the saturation vapour pressure [kPa],

e_s-e_a refers to the saturation vapour pressure deficit [kPa],

T refers to the mean daily air temperature at 2 m [°~~o~~C],

Rn refers to the net radiation at the crop surface [MJ m^{-2} day^{-1}],

γ refers to the psychrometric constant [kPa °~~o~~C^{-1}],

Δ refers to the slope vapour pressure curve [kPa °~~o~~C^{-1}].

3.2 Broyden– Fletcher – Goldfarb – Shanno Artificial Neural Networks (BFGS-ANN)

McCulloch and Pitts (1943) pioneered the original idea of neural networks. ANN is essentially a black-box modelling approach that does not identify the training algorithm explicitly, yet the modellers often trial several algorithms to attain an optimal model (Deo and Şahin, 2015). In this study, ~~Chauhan and Shrivastava (2008) and Kumar et al (2002) compared ANN models with climate based methods and stated that ANN models performed better. In this study, to go one step further, the hybrid neural network method (BFGS-ANN) optimized with~~ the Broyden – Fletcher – Goldfarb – Shanno (BFGS) training algorithm ~~approach wh~~as been used to estimate ET_0 amounts. In optimization studies~~, the Broyden–Fletcher–Goldfarb–Shanno (~~ BFGS~~)~~ method is a repetitious approach for solving unlimited nonlinear optimization problems (Fletcher, 1987). The hybrid BFGS-ANN technique trains a multilayer perceptron ANN with one hidden layer by reducing the given cost function plus a quadratic penalty ~~with the (~~using the BFGS ~~)~~ technique. The BFGS approach includes Quasi-Newton methods. For such problems, the required condition for reaching an optimal level ~~accures~~ occurs when the gradient is zero. Newton and the BFGS methods

cannot be guaranteed to converge unless the function has a quadratic Taylor expansion near an optimum. However, BFGS can have a high accuracy even for non-smooth optimization instances (Curtis et al. 2015).

Quasi-Newton methods do not compute the Hessian matrix of second derivatives. Instead, the Hessian matrix is drawn by updates specified by gradient evaluations. Quasi-Newton methods are extensions of the secant method to reach the basis of the first derivative for multi-dimensional problems. ~~In multi-dimensional problems, the secant equation does not specify a certain solution~~The secant equation does not specify a specific solution in multi-dimensional problems, and Quasi-Newton methods differ in ~~how to limit~~limiting the solution. The BFGS method is one of the frequently used members of this class (Nocedal and Wright 2006). In the BFGS-ANN method application, all attributes, including the target attribute (meteorological variables and ET_0) are standardized. In the output layer, the sigmoid function is employed for classification. In approximation, the sigmoidal function can be specified for both hidden and output layers. For regression, the activation function can be employed as the identity function in the output layer. This method was implemented on the basis of radial basis function networks trained in a fully supervised manner using WEKA's Optimization class by minimizing squared error with the BFGS method. İn this method, all attributes are normalized into the [0,1] scale (Frank 2014).

3.3 Support Vector Machine (SVR)

The statistical learning theory is the basis of the SVM. The optimum hyperplane theory and kernel functions and nonlinear classifiers were added as linear classifiers (Vapnik, 2013). Models of the SVM are separated into two main categories: (a) The classifier SVM and (b) the regression (SVR) model. An SVM is employed ~~for the classification of~~to classify data in various classes, and the SVR is employed for estimation problems. Regression is used to take a hyperplane suitable for the data used. The distance to any point in this hyperplane shows the error of that point. The best technique proposed for linear regression is the least-squares (LS) method. However, it may be ~~completely impossible to~~entirely impossible to use LS estimator in the presence of outliers. In this case, a robust predictor ~~have~~ has to be developed that will not be sensitive to minor changes, as the processor will perform poorly. Three kernel functions were used including Polynomial, Pearson VII function-based universal, and radial basis function with the level of Gaussian Noise Parameters added to the diagonal of the covariance matrix and the random number of seed to be used (equal to 1.0); the most suitable kernel function in each scenario was determined by trial and error (Frank, 2014) and the description is provided in Section 3.6.

~~In this study, during SVR and GPR modeling, three kernel functions were used including Polynomial, Pearson VII function-based universal, and radial basis function with the level of Gaussian Noise Parameters added to the diagonal of the covariance matrix and the random number of seed to be used (equal to 1.0); the most suitable kernel function in each scenario was determined by trial and error (Frank, 2014).~~

3.4 Gauss process regression (GPR)

The Gauss process (GP)R or GP is defined by Rasmussen and Williams (2005) as follows: . A GP is a complex set of random variables , any limited number of which have a joint Gaussian distribution. Kernel-based methods such as SVM and GPs can work together to solve flexible and applicable problems. The GP is generally explained by two functions: Average and covariance functions (Eq. 2). The average function is a vector; the covariance function is a matrix. The GP model is a possibly a nonparametric black box technique.

$$f \approx \mathrm{GP}\ (m, k) \qquad \text{(Eq. 2)}$$

Where f refers to Gauss distribution, m refers to a mean function and k refers to covariance function.

The value of covariance expresses the correlation between the individual outputs concerning the inputs. The covariance value determines the correlation between individual outputs and inputs. The covariance function produces a matrix of two parts (Eq.3).

$$\mathrm{Cov}\ (x_p) = C_f\ (x_p) + C_n\ (x_p) \qquad \text{(Eq. 3)}$$

Here C_f represents the functional part, but defines the unknown part of the modeling system, while C_n represents the noise part of the system, C_f represents the functional part, but defines the unknown part of the modeling system, while C_n represents the system's noise part. A Gaussian process (GP) is closely related to SVM, and both are part of the kernel machine area in ML models. Kernel methods are sample-based learners. Instead of learning a fixed parameter, the kernels memorize the training data sample and assign a certain weight to it.

3.5 Long short-term memory (LSTM)

LSTM is a high-quality evolution of Recurrent Neural Networks (RNN). This neural network is presented for addressing the problems that were present in RNN and this had been done throughto address the problems presentthat existed in RNN and this had beenis done by adding more interactions per cell. These systems are also a special one because it contains remembering information from a longn extended period. Moreover, it also contains includes four basic essential interacting layers, and all of them contain include the different methods of communicationcommunication methods. Its main structure can be explained with the help of Figure 3 (Le et al., 2019, Brownlee, 2020).

Figure 3. LSTM Architecture (Brownlee, 2020)

The next thing is that its complete network consists of a memory block. These blocks are also called ~~as~~ cells. The information is stored in one cell and then transferred into the next one with the help of gate controls. Through the help of these gates, it ~~will~~ becomes ~~extremely easy~~straightforward to analyze the information accurately. All of these gates are extremely important, and they are called forget gates. This ~~can be~~is explained ~~with the help of~~in Eq. 4.

$$f_t = \sigma\left(W_f[h_{t-1}, X_t] + b_f\right) \quad \text{(Eq. 4)}$$

LSTM units or blocks are part of the repetitive neural network structure. Repetitive neural networks are made to use some artificial memory processes that can help these AI programs mimic human thinking more effectively.

3.6 Kernel functions

~~In this study, as given in Table 3, f~~Four different kernel functions ~~that~~ are frequently used as depicted in ~~included in the~~ literature including the ~~were used. The~~ polynomial, radial-based function, Pearson VII function (PUK), and normalized polynomial kernels used and their formulas and parameters are tabulated in Table 3. As is clear from Table 3, some parameters must be determined by the user for each kernel function. While the number of parameters to be determined for PUK kernel is two, it requires ~~the determination of~~determining a parameter in the model formation that will be the basis for classification for other functions. When kernel functions are compared, it is seen that polynomial and radial based kernels are more plain and understandable. Although it may seem mathematically simple, the increase in the degree of the polynomial makes the algorithm complex. This significantly increases processing time and decreases the classification accuracy after a point. In contrast, changes in the ~~parameter (γ) of the radial-based function~~radial-based function parameter (γ), expressed as the kernel size, were less effective on classification performance (Hsu et al., 2010). The normalized polynomial function was proposed by Arnulf et al. (2001) in order to normalize the mathematical expression of the polynomial kernel instead of normalizing the data set. ~~It can be said that t~~The normalized polynomial kernel is a generalized version of the polynomial kernel. On the other hand, PUK kernel has a more complex mathematical structure than other kernel functions with its two parameters (σ, ω) known as Pearson width. These two parameters affect classification accuracy and ~~it is~~these parameters are not known in advance ~~which parameter pair will perform best~~. For this reason, determining the most suitable parameter pair in the use of PUK kernel is an important step.

Table 3. Basic kernel functions in the study and parameters ~~used in SVM~~that needs to be determined

Kernel functions	Mathematical Expression	Parameter
Polynomial kernel	$K(x, y) = \left((x.y) + 1\right)^d$	Polynomial degree (d)

Radial Based Function Kernel	$K(x,y) = e^{-\gamma\|(x-x_i)\|^2}$	Kernel size (γ)
PUK	$K(x,y) = \frac{1}{\left[1 + \left(\frac{2.\sqrt{\|x-y\|^2}\ \sqrt{2^{(1/\omega)}-1}}{\sigma}\right)^2\right]^{\omega}}$	Pearson width parameters (σ, ω)

The ~~editing parameter C must be determined by the user~~user must determine the editing parameter C for all SVM during runtime. If values that are too small or too large for this parameter are selected, the optimum hyperplane cannot be determined correctly~~, t~~. Therefore there will be a serious decrease in classification accuracy. On the other hand, if C equal to infinity, the SVM model becomes suitable only for datasets that can be separated linearly. As can be seen from here, the selection of appropriate values for the parameters ~~is a variable that~~ directly affects the accuracy of the SVM classifier. Although a trial and error strategy is generally used, the cross-validation approach enables successful results. The purpose of the cross-validation approach is to determine the performance of the classification model created. For this purpose, the data is separated into two categories where ~~. While~~ the first ~~part~~ is used as training ~~data in a~~the model ~~form that is the basis for classification~~and, the second part is processed as test data to determine the ~~performance of the model~~model's performance. As a result of applying the model created with the training set to the test data set, the number of samples classified correctly indicates the ~~performance of the classifier~~classifier's performance. Therefore, by using the cross-validation method, ~~the model that will be the basis for~~the classification and determination of the best kernel parameters ~~, in which the best classification performance is~~were obtained ~~, was created~~ (Kavzoglu and Golkesen, 2010).

In this study, during SVR and GPR modeling, the three kernel functions ~~were used including Polynomial, Pearson VII function-based universal, and radial basis function with the level of Gaussian Noise Parameters added to the diagonal of the covariance matrix and the random number of seed to be used (equal to 1.0);~~as in Table 3 were used and the most suitable kernel function in each scenario was determined by trial and error (Frank, 2014).

~~In this study,~~ For the BFGS-ANN, SVR, and GPR methods in the Weka software were used, while ~~. Weka is basically a data mining program developed in Java and distributed as open source by Waikato University, where functions such as data pre-processing and machine learning algorithms are presented together. P~~python language ~~were~~ as used for the LSTM method ~~during modelling~~.

3.7 Model Evaluation

~~Some common~~The statistical parameters ~~were~~ used in the selection and comparison of the models ~~used~~ in the study included ~~. Differences between ET_0 data determined by~~ the root mean square error (RMSE), mean absolute error (MAE), and correlation fit (R) ~~via~~ as shown in Eq. 5-7. Here, X_i and Y_i are the observed and predicted values, and N is the number of data.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|X_i - Y_i| \quad \text{(Eq. 5.)}$$

$$RMSE = \sqrt{\frac{1}{N}\Sigma_{i=1}^{N}(X_i - Y_i)^2} \quad \text{(Eq. 6.)}$$

$$R = \frac{N\sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{N(\sum X_i^2) - (\sum X_i)^2}\sqrt{N(\sum Y_i^2) - (\sum Y_i)^2}} \quad \text{(Eq. 7.)}$$

~~Also,~~In addition, Taylor diagrams ~~applied~~ were developed to check the performance of the ~~used~~ models, which illustrates ~~. In this diagram,~~ the ~~observed~~ experimental and statistical parameters ~~can be detected~~ simultaneously. ~~Moreover, various points on the polar graph are used to study the adaption between measured and predicted values in the Taylor diagrams. Besides, CC and normalized standard deviation are indicated by the azimuth angle, and radial distances from the base point, respectively (Taylor 2001).~~

4 Results

In this study, 10 different alternative scenarios were created by using combinations or input variables, i.e., monthly average, highest and lowest temperature, sunshine duration~~me~~, wind speed, average, highest, and lowest relative humidity data. ET_0 amounts were estimated with the help of kernel-based GPR and SVR methods, BFGS-ANN, and one of the deep learning methods LSTM models. ET_0 estimation results obtained from different scenarios according to the GPR method are summarized in Table 4. As can be seen from the table, the 5th scenario containing four meteorological variables including TMax, TMin, TMean and *n* ~~containing four meteorological variables~~ with the GPR method PUK function gave the best result (Train period: R^2 = 0.9667, MAE = 9.1279 mm/month, RMSE = 11.067 mm/month; Test period: R^2 = 0.96438~~2~~, MAE = 9.1947 mm/month, RMSE = 11.2109 mm/month) ~~gave the best result~~. However, the 8th scenario with only one meteorological variable (sunshine duration) ~~and an appropriate result~~registered quite well results with ~~(T~~training period: R^2 = 0.9472, MAE = 10.1629 mm/month, RMSE = 13.2694 mm/month and ~~h; T~~testing period: R^2 = 0.9392~~691~~, MAE = 11.8473 mm/month, RMSE = 15.8719 mm/month~~)~~. Since the scenario with the least input parameters and with an acceptable level of accuracy ~~can be~~is largely preferred, scenario 8 was chosen as the optimum scenario.

The scatter plot and time series plots of the test phase for scenario 5 and 8 are given in Figures 4 and 5. As can be seen from the figure, a relative agreement has been achieved between the FAO56PM ET_0 values and the ET_0 values modeled. When the time series graphs are examined, minimum points in estimated ET_0 values are more in harmony with FAO56PM values than maximum points.

Table 4. ~~Results obtained with the~~Outcomes of the GPR modelling approach ~~method~~from different kernel functions based on R^2, MAE, and RMSE. (Bold represents the best results; Yellow highlighted represents the optimally selected model)

Scenario No	Kernel functions	Train			Test		
		R^2	MAE	RMSE	R^2	MAE (mm/month)	RMSE (mm/month)
1	Polynomial	0.9084	13.1238	16.0365	0.8451	17.8013	21.4952
	PUK	0.9732	6.8024	8.9055	0.9506	10.5906	13.4330
	Radial basis function	0.9357	22.3706	25.3578	0.9220	22.3353	25.4332
2	Polynomial	0.8825	15.0049	18.3607	0.8332	19.4655	23.9183
	PUK	0.9666	7.2041	9.4750	0.9639	**8.9058**	11.5185
	Radial basis function	0.9450	27.7700	31.2897	0.9366	27.5940	31.2150
3	Polynomial	0.8697	15.7587	19.2936	0.7807	21.2623	26.2083
	PUK	0.9436	9.5556	12.6058	0.9335	12.2152	15.0187
	Radial basis function	0.9251	31.3045	35.4426	0.9073	31.9344	36.1935
4	Polynomial	0.7002	37.824	43.417	0.7105	36.6604	41.2745
	PUK	0.9637	7.7384	10.153	0.9629	9.3003	12.4647
	Radial basis function	0.9374	29.1996	32.9582	0.9491	29.7864	33.6709
5	Polynomial	0.6312	35.1424	40.3818	0.6030	33.8278	38.3742
	PUK	**0.9667**	**9.1279**	**11.067**	**0.9643**	**9.1947**	**11.2109**
	Radial basis function	0.9239	25.6568	29.4976	0.9239	26.2766	30.0768
6	Polynomial	0.8703	15.6789	19.3039	0.7841	21.5210	27.2959
	PUK	0.9569	8.5950	11.1225	0.9401	12.1685	15.8165
	Radial basis function	0.9229	33.0011	36.9189	0.8991	33.4845	37.9140
7	Polynomial	0.8599	16.6129	20.0640	0.7852	21.7258	26.9480
	PUK	0.9349	10.3820	13.5482	0.9310	12.9590	16.5650
	Radial basis function	0.9086	36.4501	40.9667	0.8746	36.9353	41.6716
8	Polynomial	0.9203	41.2839	46.5019	0.9281	40.4306	45.9593
	PUK	0.9472	10.1629	13.2694	0.9392	11.8473	15.8719
	Radial basis function	0.9283	37.0877	41.8535	0.9281	37.6298	42.3803

9	Polynomial	0.8394	44.0191	49.2989	0.8380	43.9357	50.0790
	PUK	0.8759	15.0361	18.5984	0.8634	16.2747	20.1854
	Radial basis function	0.8398	39.0547	44.3349	0.8380	40.0566	44.8850
10	Polynomial	0.8677	43.1716	48.2151	0.8746	42.6604	48.7584
	PUK	0.9027	13.3821	16.4932	0.9130	13.0145	15.8309
	Radial basis function	0.8679	38.2998	43.4373	0.8748	39.1677	43.9253

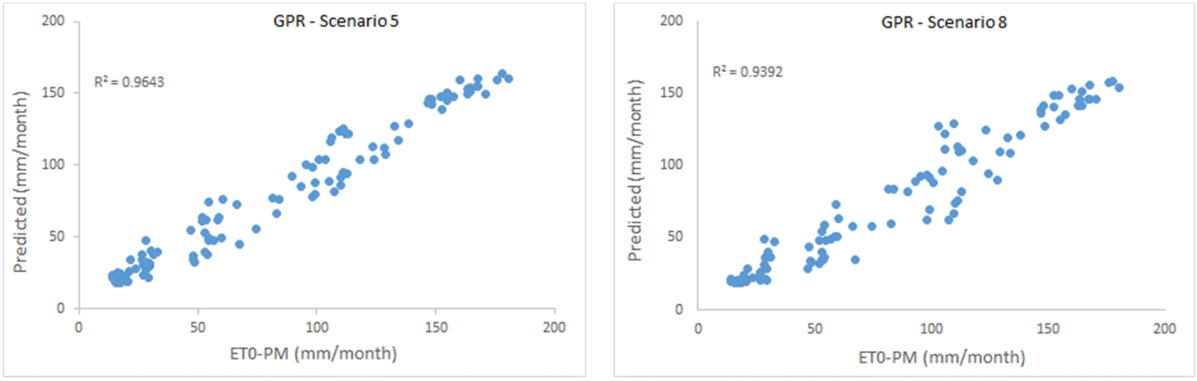


Figure 4. Scatter plots comparing GPR estimated and FAO56PM estimated ET_0 in~~of~~ scenarios 5 and 8

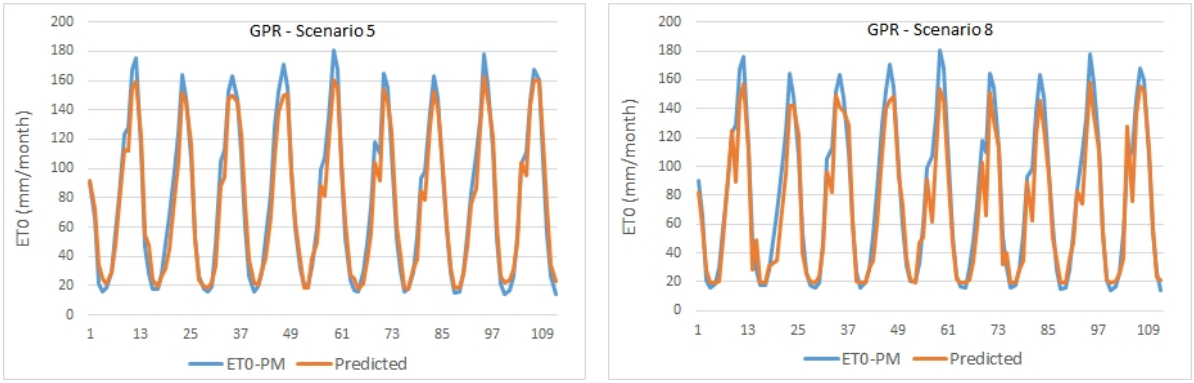


Figure 5. Time series graphics of GPR estimated and FAO56PM estimated ET_0 in scenarios 5 and 8

~~Within the~~For SVR model, again 3 different kernel functions were evaluated in scenarios under the same conditions, and the results ~~obtained~~ displayed ~~are given~~ in Table 5.

As can be seen here, scenarios 5 and 8 have yielded the best and most appropriate results according to the PUK function. According to the results, the 5th scenario including TMean, TMin, TMax and *n* meteorological variables gave the best result (Train period: R^2 = 0.9838, MAE = 6.0500 mm/month, RMSE = 8.5733 mm/month; Test period: R^2 = 0.9771~~885~~, MAE =

7.07 mm/month, RMSE = 9.3259 mm/month). However, scenario 8 gave the most appropriate result (Train period: R^2 = 0.9398, MAE = 9.7984 mm/month, RMSE = 13.0830 mm/month; Test period: R^2 = 0.9392~~691~~, MAE = 11.2408 mm/month, RMSE = 15.5611 mm/month) only with the meteorological variable of sunshine duration (*n*). Although the accuracy rate of the 8th scenario is somewhat lower than the 5th scenario, it provides convenience and is preferred in terms of application and calculation since it requires a single parameter. The sunshine duration can be measured easily and without the need for ~~excessive~~ high ~~costs~~ cost equipment and personnel. Consequently, by using ~~, so that with just~~only one parameter, the amount of ET_0 is estimated within acceptable accuracy limits.

Table 5. Outcomes of the SVR modelling approach from different kernel functions based on R^2, MAE, and RMSE. (Bold represents the best results; Yellow highlighted represents the optimally selected model)~~Results obtained with the SVR method for all scenarios~~

Scenario No	Kernel function	Train			Test		
		R^2	MAE	RMSE	R^2	MAE	RMSE
1	Polynomial	0.9667	7.6671	9.6167	0.9655	11.0033	13.5740
	PUK	0.9790	1.3130	2.9310	0.9683	8.70480	11.1693
	Radial basis function	0.9446	10.3256	12.5561	0.9366	11.1203	13.4468
2	Polynomial	0.9587	9.8445	12.0674	0.9526	10.1138	11.6124
	PUK	0.9775	4.3655	8.0208	0.9742	8.88250	11.6469
	Radial basis function	0.9487	11.0557	12.8207	0.9456	11.4313	13.5386
3	Polynomial	0.9392	10.088	13.468	0.9160	13.5919	15.903
	PUK	0.9608	7.1018	7.1018	0.9249	12.0206	15.6733
	Radial basis function	0.9401	12.1973	14.4483	0.9107	15.1051	18.4364
4	Polynomial	0.9491	10.5076	12.7585	0.9485	11.8516	14.1386
	PUK	0.9732	5.5868	8.6784	0.9604	9.2452	12.5707
	Radial basis function	0.9593	12.7177	14.8832	0.9500	12.6226	16.1700
5	Polynomial	0.9743	8.9452	11.5497	0.9657	8.5349	10.2108
	PUK	**0.9838**	**6.0500**	**8.5733**	**0.9771**	**7.0700**	**9.3259**
	Radial basis function	0.9414	11.8017	15.1588	0.9318	11.8607	14.4412
6	Polynomial	0.9399	10.3413	12.9082	0.9281	14.5901	17.9626
	PUK	0.9698	6.1970	9.1435	0.9497	11.2859	14.7455
	Radial basis function	0.9299	13.9103	17.0013	0.9120	16.7198	22.2031
7	Polynomial	0.9214	11.9563	14.8277	0.9214	14.7185	17.6297
	PUK	0.9426	9.1560	12.6111	0.9407	12.0180	15.5924

	Radial basis function	0.9164	17.8134	21.4555	0.8951	19.4352	25.7907
8	Polynomial	0.9283	12.0330	14.9227	0.9281	13.7164	16.4672
	PUK	0.9398	9.7984	13.0830	0.9392	11.2408	15.5611
	Radial basis function	0.9283	18.6912	22.9160	0.9281	19.1426	25.6111
9	Polynomial	0.8394	17.2037	21.1520	0.8380	17.9619	22.8538
	PUK	0.8755	14.3397	18.8555	0.8623	16.2552	20.9296
	Radial basis function	0.8398	25.6982	31.0532	0.8380	26.4915	31.2574
10	Polynomial	0.8777	14.7758	19.8128	0.8746	15.2039	19.8289
	PUK	0.9087	12.2525	17.3738	0.9084	12.0109	16.8281
	Radial basis function	0.8779	23.2745	28.4086	0.8748	23.7460	28.7051

The scatter plot and time series graph drawn for the SVR model ~~is given in Figure~~are given in Figures 6 and 7, which shows that ~~. As can be seen from the figure, except for the less frequent endpoints, the~~ all ~~other~~ points are compatible with FAO56PM - ET_0 values and ET_0 values estimated from the model, except for the less frequent endpoints. The R^2 values were also very high ($R^2 > 0.939$).

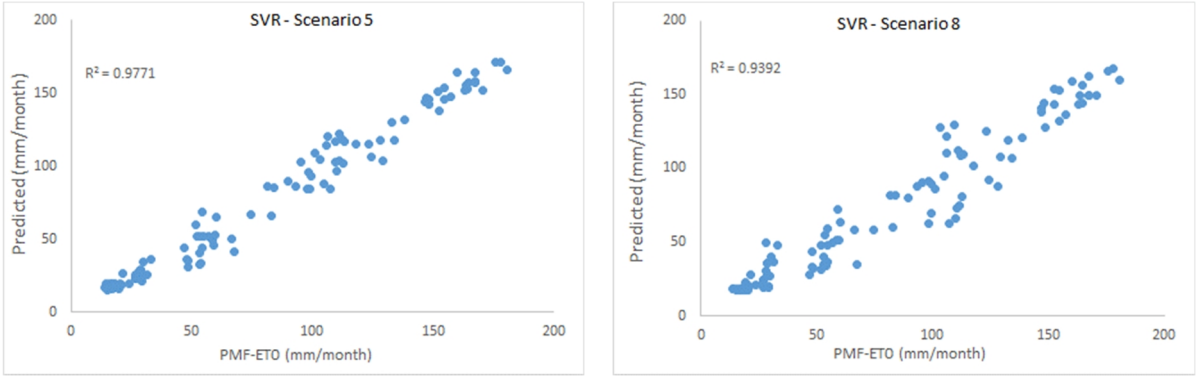


Figure 6. Scatter plots comparing SVR estimated and FAO56PM estimated ET_0 ~~of~~ in scenarios 5 and 8

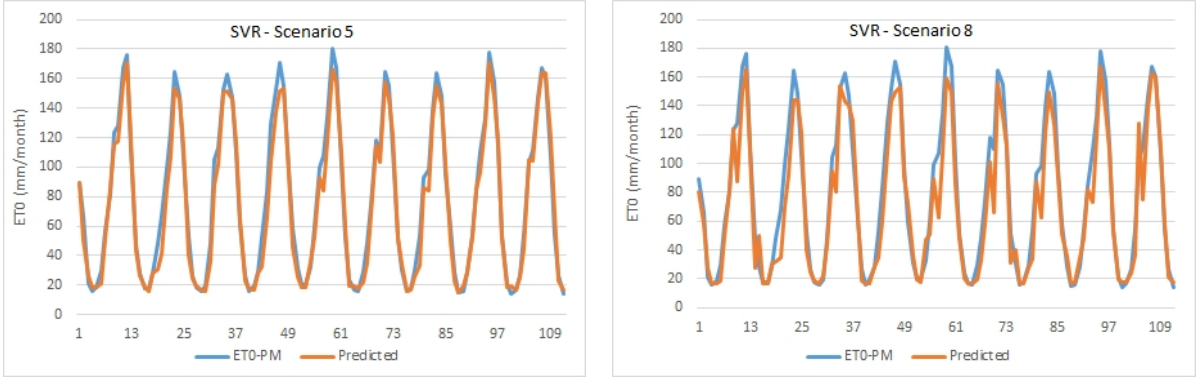


Figure 7. Time series graphics of SVR estimated and FAO56PM estimated ET_0 in~~of~~ scenarios 5 and 8

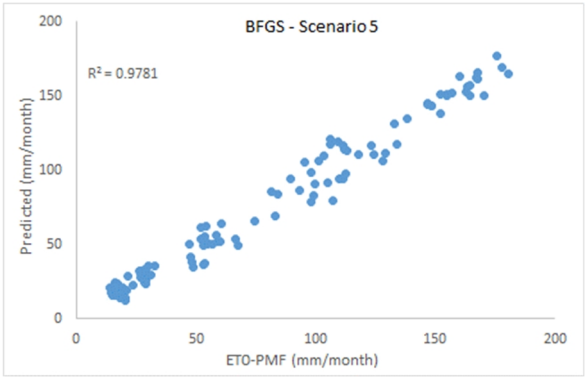
In this study, BFGS training algorithm was specifically used to train the ANN ~~was optimized with the BFGS method~~architecture and ET_0 amounts were estimated for all scenarios and the results are given in Table 6. In ~~the implementation of~~implementing the BFGS-ANN method, all features, including target feature (meteorological variables and ET_0) ~~,~~ are standardized. In the hidden and output layer, the sigmoid function is $f(x) = 1 / (1 + e^{-x})$ used for classification.

As can be seen here, scenarios 5 and 8 have given the best and most ~~appropriate~~ relevant results. According to the results, the 5th scenario including TMean, TMin, TMax and *n* meteorological variables again ~~gave~~ produced the best result (Train period: R^2 = 0.9843, MAE = 8.0025 mm/month, RMSE = 9.9407 mm/month; Test period: R^2 = 0.97818~~9~~, MAE = 6.7885 mm/month RMSE = 8.8991 mm/month). However, Scenario 8 gave the most appropriate result (Train period: R^2 = 0.9474, MAE = 10.1139 mm/month, RMSE = 13.1608 mm/month; Test period: R^2 = 0.94287~~1~~, MAE = 11.4761 mm/month, RMSE = 15.6399 mm/month) with only the sunshine duration (n) meteorological variable, hence been the optimally selected BFGS-ANN model. Although the ~~accuracy rate of the 8th scenario~~8th scenario's accuracy rate is marginally less than the 5th scenario, it is easy and practical in terms of application and calculation since it consists of only one parameter. The scatter plot and time series graph drawn for the BFGS-ANN model~~, is~~ given in Figures 8 and 9 concurs with the statistical metrics of Table 6. As can be seen, the BFGS-ANN method predicted ET_0 amounts with a high success rate, and a high level of agreement was achieved between the estimates obtained from the model and FAO56PM- ET_0 values. The R^2 values were also very high ($R^2 > 0.942$).

Table 6. Outcomes of the BFGS-ANN modelling approach for different Scenarios based on R^2, MAE, and RMSE. (Bold represents the best results; Yellow highlighted represents the optimally selected model)~~Results obtained with BFGS-ANN method~~

Scenario No	Train			Test		
	R^2	MAE	RMSE	R^2	MAE	RMSE
1	0.9778	6.7017	8.6972	0.9769	**6.6346**	**8.6243**
2	0.9763	7.2683	9.6751	0.9700	7.5305	10.3722
3	0.9450	9.2810	12.3463	0.9423	11.2870	14.3732
4	0.9670	7.8325	10.4035	0.9659	9.1159	12.4740
5	**0.9843**	**8.0025**	**9.9407**	**0.9781**	**6.7885**	**8.8991**
6	0.9536	8.9027	11.3546	0.9522	11.5089	14.7687
7	0.9466	10.2246	13.2535	0.9417	11.9444	15.7787
8	0.9474	10.1139	13.1608	0.9428	11.4761	15.6399
9	0.8768	14.8765	18.4766	0.8709	15.9139	19.8957

10	0.9158	13.0161	16.2424	0.9149	12.4874	15.5428



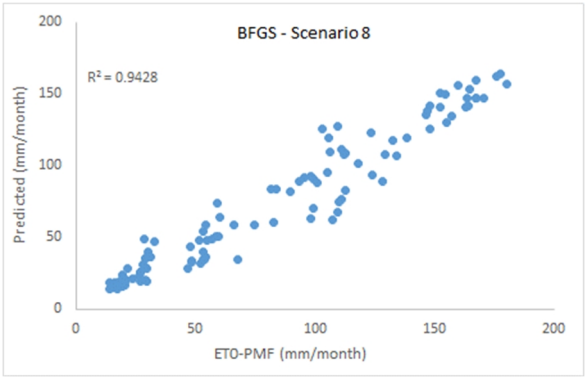
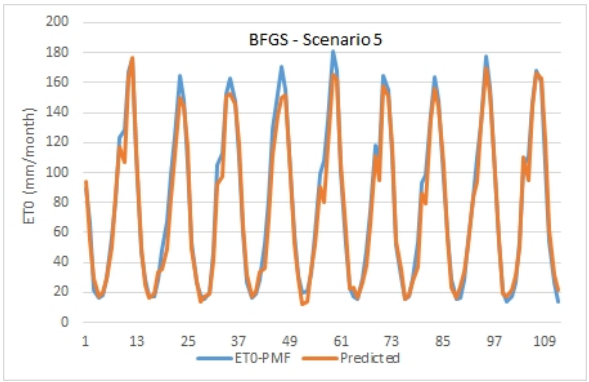


Figure 8. Scatter plots comparing BFGS-ANN estimated and FAO56PM estimated ET_0 in~~of~~ scenarios 5 and 8



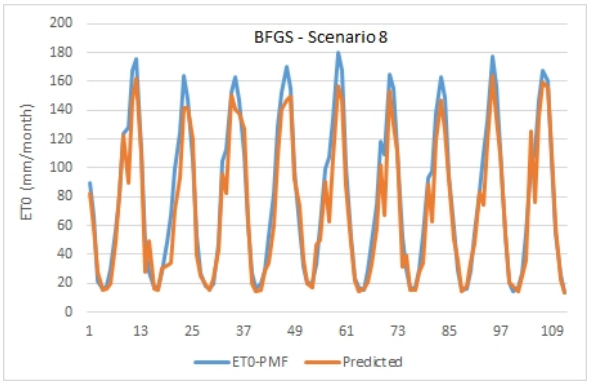


Figure 9. Time series graphics of BFGS-ANN estimated and FAO56PM estimated ET_0 in ~~of~~ scenarios 5 and 8

~~In this study, ET_0 amounts were estimated with~~Finally, ~~the~~ the LSTM method, which is ~~included in~~a deep learning technique was used to estimate the ET_0~~s~~ ~~,~~under the same 10 Scenarios. ~~In the LSTM model used in this study~~is study's LSTM model, ~~t~~Two hidden layers with 200 and 150 neurons were utilized in LSTM with~~,~~ the rectified linear unit (ReLU) ~~relu~~ activation function, and Adam optimization ~~swere used~~. The other parameters: Learning rate alternatives from $1e^{-1}$ to $1e^{-9}$, Decay as $1e^{-1}$ to $1e^{-9}$, and 500-750-1000 as epoch have been tried. The best results obtained for 10 different scenarios at the modeling stage, according to the LSTM method, are given in Table 7.

Table 7. Outcomes of the LSTM modelling approach for different Scenarios based on R^2, MAE, and RMSE. (Bold represents the best results; Yellow highlighted represents the optimally selected model)~~Results obtained with LSTM method~~

Scenario No	Train			Test		
	R^2	MAE	RMSE	R^2	MAE	RMSE

1	0.9825	7.0178	9.3020	0.9769	8.6232	11.4663
2	0.9618	9.0678	12.4321	0.9604	8.5703	11.7467
3	0.9403	13.841	16.3260	0.9345	14.8644	17.1128
4	0.9499	10.375	12.3748	0.9393	11.5043	13.7417
5	**0.9835**	**4.9405**	**6.8687**	**0.9759**	**6.2907**	**8.5897**
6	0.9694	11.532	15.7447	0.9602	8.1580	10.6059
7	0.9382	10.962	14.8716	0.9366	10.1113	13.6070
8	0.9461	12.461	15.7539	0.9384	11.6711	14.4864
9	0.8807	14.479	18.2882	0.8664	15.2565	19.4120
10	0.9231	14.195	17.1729	0.9220	13.7034	16.1857

As in other methods, the 5th and 8th scenarios of the LSTM model ~~gave~~ registered the best and most appropriate results. ~~According to the results,~~In the 5th scenario ~~including~~ TMean, TMin, TMax and *n* ~~meteorological variables~~as the input variables gave the best result (Train period: R^2 = 0.9835, MAE = 4.9405 mm/month, RMSE = 6.8687 mm/month; Test period: R^2 = 0.9759~~879~~, MAE = 6.2907 mm/month RMSE = 8.5897 mm/month). However, scenario 8 only gave the most appropriate result (Train period: R^2 = 0.9461, MAE = 12.461 mm/month, RMSE = 15.7539 mm/month; Test period: R^2 = 0.9384~~687~~, MAE = 11.6711 mm/month, RMSE = 14.4864 mm/month) with the sunshine duration (*n*) meteorological variable.

Scatter plot and time series graphs of observed and LSTM predicted ET_0 are ~~drawn for the LSTM model are~~ given in Figures 10 and 11, where again ~~. As can be seen, the LSTM method predicted ET_0 amounts with~~ a high success rate, and a high level of agreement was achieved between the estimates obtained from the model and FAO56PM- ET_0 values.

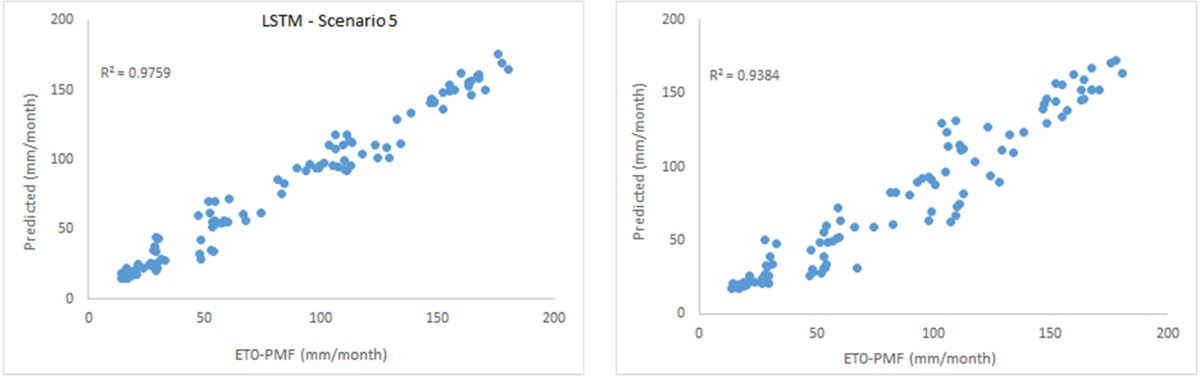


Figure 10. Scatter plots comparing LSTM estimated and FAO56PM estimated ET_0 in~~of~~ scenarios 5 and 8

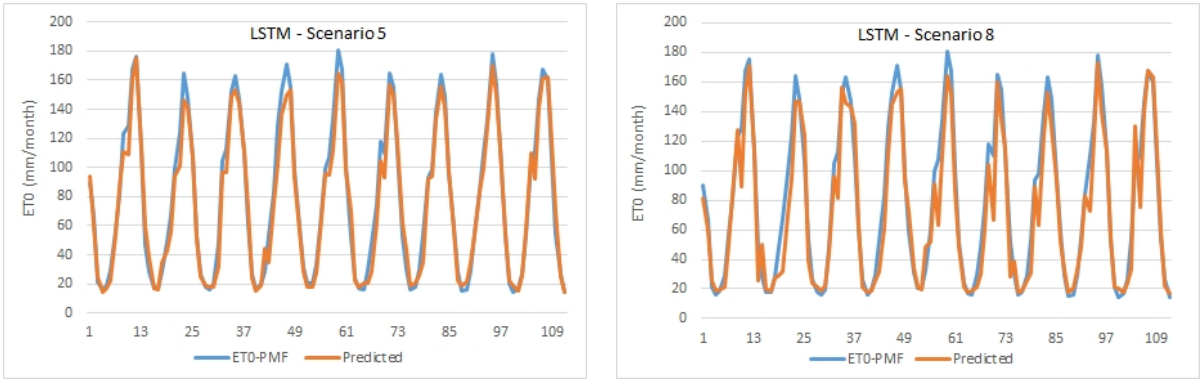


Figure 11. Time series graphics of LSTM estimated and FAO56PM estimated ET_0 in ~~of~~ scenarios 5 and 8

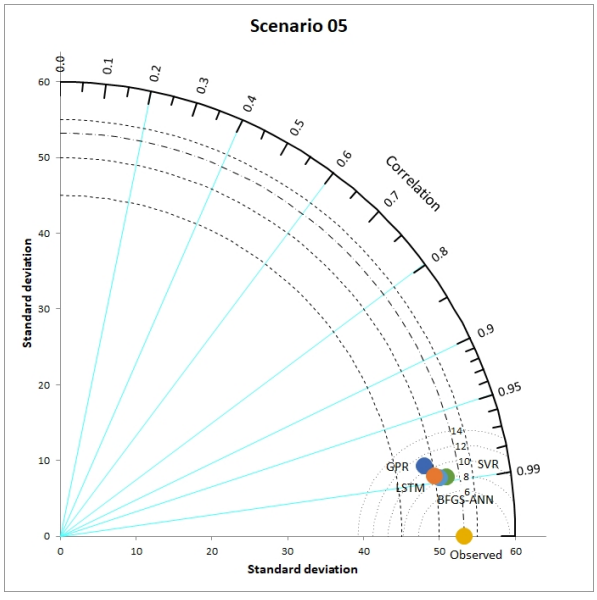
In order to compare and evaluate ~~and compare~~ the models used in this study, statistical values for the test phase are given in both FAO56PM- ET_0 and from the respective models in Table 8. The lowest skewness coefficient was found in scenario 5 in both GPR and SVR methods with 0.39 and the highest in LSTM scenario 8 with 0.52. The lowest kurtosis coefficient has Tmean with -1.23 and the highest with 0.36 by RHmean parameter. The highest variation was observed in RHmin with 174.19 and the lowest in U parameter with 0.17.

Table 8. Statistical values of the test phase for selected scenarios

	GPR		SVR		BFGS-ANN		LSTM		
Statistic	**Scenario 5**	**Scenario 8**	**Scenario 5**	**Scenario 8**	**Scenario 5**	**Scenario 8**	**Scenario 5**	**Scenario 8**	**ET_0PM**
Minimum	17.687	19.1090	15.1900	17.1520	12.2480	**13.9060**	14.2971	16.9787	13.99
Maximum	163.440	158.557	**180.530**	167.527	176.765	164.100	175.613	172.767	180.53
Mean	**75.8818**	71.3861	74.5771	71.2124	75.8644	70.7299	75.6023	72.3210	79.21
Stdev	48.8941	47.6359	51.5342	48.9192	50.6812	48.2539	50.0143	**50.2075**	53.26
Correlation	0.9820	0.9691	0.9885	0.9691	**0.9890**	0.9710	0.9879	0.9687	1
Skewness	0.39	0.47	0.39	0.51	0.41	0.46	**0.36**	0.52	0.36
Kurtosis	-1.29	**-1.27**	-1.32	-1.16	-1.24	-1.21	-1.21	-1.16	-1.27
Coefficient of variation	2344.09	2226.93	**2655.77**	2393.09	2568.59	2328.44	2501.43	2520.80	2836.65
Number of records	112	112	112	112	112	112	112	112	112

As can be seen from ~~this t~~Table 8, the closest value to the FAO56PM-ET_0 minimum value (13.99 mm/month) is 8th scenario in the BFGS-ANN method (13.906 mm/month~~);~~). Furthermore, the FAO56PM-ET_0 maximum value (180.53 mm/month) has been reached in the 5th scenario (180.53 mm/month) in the SVR method which is the closest and even the same value. The value closest to the mean value of FAO56PM-ET_0 (79.21 mm/month) belongs to the 5th scenario (75.8818 mm/month) in the GPR method; the value closest to the FAO56PM-ET_0 Stdev value (53.26 mm/month) is the value of the 5th scenario (51.5342 mm/month) in the SVR method. As shown from Table 8, all methods have estimated the ET_0 amounts within acceptable, yet disparate results are attained when comparing the statistics. Having said that, ~~If the~~when models are ranked according to the correlation coefficient, the best results were BFGS-ANN, SVR, LSTM, and GPR in the 5th scenario and BFGS-ANN, GPR, SVR, and LSTM in the 8th scenario. ~~As can be seeshown from the table, all methods have estimated the ET_0 amounts to be acceptable, all with little difference.~~

~~Furthermore~~Furthermore, ~~T~~to have a precise model comparative evaluat~~e~~ions ~~the models with the help of diagrams~~ besides the tables, the Taylor diagram for the 5th and 8th scenarios ~~is given in~~were plotted as in Figure 12. The points on the polar Taylor graph are used to study the adaption between measured and predicted values in the Taylor diagram. The correlation coefficient and normalized standard deviation are also indicated by the azimuth angle, and radial distances from the base point, respectively (Taylor 2001). As ~~can be understood~~displayed ~~from~~ in the figure, all four ~~methods used follow one another with a slight difference~~slightly,performed quite well but BFGS-ANN ~~seems~~ seemed to achieve higher success than others. ~~When looking at the~~As stated earlier in Figure 1- histogram ~~graph (Figure 1)~~, it is seen that FAO56PM- ET_0 values do not conform to normal distribution. This mismatch is considered to be ~~the reason why the~~the reason for poor performances of the GPR method ~~is relatively less successful than other~~over comparative ~~methods~~models.



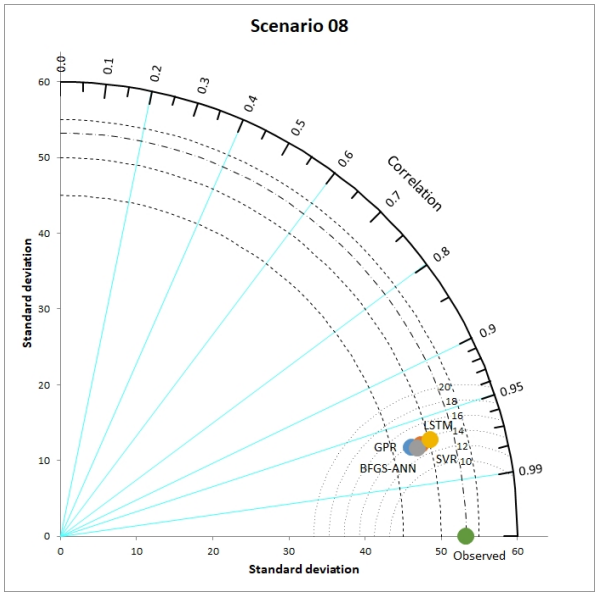


Figure 12. Taylor diagrams of scenarios 5 and 8

~~According to t~~The results of Figure 12 also show that models performances were higher in Scenario 5, however, ~~obtained in this study, it is seen that by~~ using ~~fewer~~the least input parameters to develop the most parsimonious model was the key target of the ~~in the~~ study ~~area~~, and was achieved by the Scenario 8 whereby ~~-~~ET_0 values ~~a~~were estimated correctly at a relatively appropriate and acceptable level~~s~~. Therefore, ~~it is thought that~~ these methods ~~can give~~produced trustworthy results and has the potential to make correct estimations in climates similar to the study area.

5 Conclusion

The amount of ET_0 can be calculated with many experimental equations. However, these equations can generally differ spatially and also contain many parameters. Since ET_0 ~~contains~~ includes a complex and nonlinear structure, it can~~not~~ be easily estimated with the previously measured data without requiring ~~a lot of~~~~many~~numerous parameters ~~with using ML methods~~. In this study, ~~the estimation of~~estimating the ~~amount of~~ ET_0 with ~~a~~ different machine learning~~es~~ and deep learning methods was made ~~by using the least meteorological variable in the Corum region, which arid and semi-arid climate with Turkey has~~using the least meteorological variable in Turkey's Corum region, with an arid and semi-arid climate ~~-~~as a~~n~~ ~~important~~ strategic ~~place in terms of~~ agricultural regione. In this context, firstly, ET_0 amounts were calculated with the Penman-Monteith method and taken into consideration as the output of the models ~~used~~. Then, ~~10~~ 10~~ten~~ different scenarios were ~~created~~ then created using different combinations of~~,~~ ~~including the average,~~ meteorological variables ~~highest and lowest temperatures, sunshine duration, wind speed, average, highest, and lowest~~ most insufficient ~~relative humidity meteorological variables~~. Consequently, Kernel-based GPR and SVR methods, BFGS-ANN, and LSTM models ~~are used~~were developed for monthly ~~, and~~ ET_0 amount ~~s are~~ estimation~~sed~~. ~~According to t~~The results ~~obtained,~~revealed better performance of the BFGS-ANN model ~~in performed better than~~comparison other models under study~~s~~, although all four methods ~~used~~ predicted ET_0 amounts within acceptable accuracy and error levels. In kernel-based methods (GPR and SVR), PUK was the most successful kernel function~~.~~. ~~Besides, t~~The 5th scenario, which is related to temperature and includes four meteorological variables ~~such (as average,~~ highest and lowest temperature averages, and sunshine duration~~),~~ gave the best results in all the scenarios used. ~~After the 5th scenario, the~~Scenario 8~~th, scenario,~~ which ~~includes~~ included only the sunshine duration, was determined as the most suitable and parsimonious scenario. In this case, the ET_0 amount was estimated using only sunshine duration without the need for other meteorological parameters for the study area. The Corum region is described as arid and semi-arid ~~as the~~with low rainfalls ~~are low. In a sense, this means that the~~and cloudiness ~~is low~~ and ~~the~~ longer sunshine duration ~~s~~~~is long~~, hence sunshine hours is the key driving factor of ET_0 in the region which is ~~clealy~~clearly highlighted by high model ~~performanes~~performances with sunshine hours as the only input. ~~. The reason for the models~~models' ~~reason to give quite good results by using only sunshine duration value is a long time in the work area.~~ Continuous measurement of meteorological variables in large farmland is a costly process that requires expert personnel, time, or good equipment. Simultaneously, some equations used for ET_0 calculations are not preferred

by specialists because they contain many parameters. In this case, it is very advantageous for water resources managers to estimate ET_0 amounts only with sunshine duration time, which is easy to measure and requires no extra cost. ~~In the follow-up study, it is aimed~~The follow-up study aims to evaluate the performance of GPR and LSTM models in a larger area on a daily time scale and with data to be obtained from more meteorology stations.

Funding:

This work was supported by Fellowships for Visiting Scientists and Scientists on the Sabbatical Leave Programme (2221) of The Scientific and Technological Research Council of Turkey (TUBITAK).

Acknowledgement:

We acknowledge the 'Open Access Funding by the Publication Fund of the TU Dresden'.

Author Contributions: Conceptualization, M~~.T.S.~~TS, and H~~.A.~~A; Data curation, M~~.T.S.~~TS, H~~.A.~~A; Formal analysis, S~~.S.~~S; Funding acquisition, H~~.A.~~A, M~~.T.S.~~TS; Investigation, M~~.T.S.~~TS, H~~.A.~~A and S~~.S.~~S; Methodology, M~~.T.S.~~TS, H~~.A.~~A and S~~.S.~~S; Project administration, H~~.A.~~A; Resources, M~~.T.S.~~TS and H~~.A.~~A; Software, M~~.T.S.~~TS A.M. and H~~.A.~~A; Supervision, M~~.T.~~TS, H~~.~~A.; Validation, A.M. and S~~.S.~~S; Visualization, M~~.T.S.~~TS, H~~.A.~~A and S~~.S.~~S; Writing—original draft, M~~.T.S.~~TS, H~~.A.~~A and S~~.S.~~S; Writing—review and editing, RP M~~.T.S.~~TS A.M. and S~~.~~S. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Code/Data availability: Data are available on request due to privacy or other restrictions.

References

Abrishami, N., Sepaskhah, A.R., Shahrokhnia, M.H. 2019. Estimating wheat and maize daily evapotranspiration using artificial neural network. Theor. Appl. Climatol. 135, 945–958. https://doi.org/10.1007/s00704-018-2418-4

Allen, Rg, Jensen, Me, Wright, Jl Burman, Rd. 1989. Operational estimates of reference evapotranspiration. Agronomy Journal, 81(4):650-662. Doi: 10.2134/Agronj1989.00021962008100040019x

Anonymous. 2017. Agricultural data of Corum province (In Turkish). Corum province Food Agriculture Livestock Directorate

Anli AS. 2014. Temporal Variation of Reference Evapotranspiration (ET0) in Southeastern Anatolia Region and Meteorological Drought Analysis through RDI (Reconnaissance Drought Index) Method. Journal of Agricultural Sciences Tarim Bilimleri Dergisi 20(3):248-260. https://doi.org/10.15832/tbd.82527

Arnulf B. A. Graf, S. Borer. 2001. Normalization in Support Vector Machines, Lecture Notes in Computer Science, 2191: 277-282.

Banda, P., Cemek, B., Kucukktopcu, E. 2018. Estimation of daily reference evapotranspiration by neuro computing techniques using limited data in a semi-arid environment. Arch. Agron. Soil Sci. 64, 916–929. https://doi.org/10.1080/03650340.2017.1414196

Bowden GJ, Dandy GC, Maier HR. Input determination for neural network models in water resources applications. Part 1—background and methodology. Journal of Hydrology. 2005;301:75-92.

Brownlee, J. 2020. How to Develop LSTM Models for Time Series Forecasting. Retrieved from https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/

Chauhan, S., and Shrivastava, R.K. 2008. "Performance evaluation of reference evapotranspiration estimation using climate based methods and artificial neural networks." Water Resour. Manage., 23,825–837. doi:10.1007/s11269-008-9301-5

Citakoglu, H., Cobaner, M., Haktanir, T., Kisi, O. 2014. Estimation of Monthly Mean Reference Evapotranspiration in Turkey. Water Resour. Manag. 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(2), 480-497.

Curtis, F.E., Que, X., 2015. A quasi-Newton algorithm for nonconvex, nonsmooth optimization with global convergence guarantees. Math. Program. Comput. 7, 399–428. https://doi.org/10.1007/s12532-015-0086-2

Deo RC, Şahin M. Application of the Artificial Neural Network model for prediction of monthly Standardized Precipitation and Evapotranspiration Index using hydrometeorological parameters and climate indices in eastern Australia. Atmos Res. 2015;161-162:65-81.

Doorenbos, J. and Pruitt, W.O. 1977. Crop Water Requirements. FAO Irrigation and Drainage Paper 24, FAO, Rome, 144 p.

Fan, J., Yue, W., Wu, L., Zhang, F., Cai, H., Wang, X., Lu, X., Xiang, Y., 2018. Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China. Agric. For. Meteorol. 263, 225–241. https://doi.org/10.1016/j.agrformet.2018.08.019

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. J. Hydrol. 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. Comput. Electron. Agric. 136, 71–78. https://doi.org/10.1016/j.compag.2017.01.027

Fletcher R. 1987. Practical methods of optimization (2nd ed.), New York: John Wiley & Sons, ISBN 978-0-471-91547-8, Wiley.

Rasmussen, CE. and Williams, CKI. 2005. Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning series).

Gavili, S., Sanikhani, H., Kisi, O., Mahmoudi, M.H., 2018. Evaluation of several soft computing methods in monthly evapotranspiration modelling. Meteorol. Appl. 25, 128–138. https://doi.org/10.1002/met.1676

Gocić, M., Motamedi, S., Shamshirband, S., Petković, D., Ch, S., Hashim, R., Arif, M., 2015. Soft computing approaches for forecasting reference evapotranspiration. Comput. Electron. Agric. 113, 164–173. https://doi.org/10.1016/j.compag.2015.02.010

Hargreaves GH, Samani ZA. 2013. Reference Crop Evapotranspiration from Temperature. Applied Engineering in Agriculture. 1(2): 96-99. (doi: 10.13031/2013.26773)

Hejazi MI, Cai X. Input variable selection for water resources systems using a modified minimum redundancy maximum relevance (mMRMR) algorithm. Advances in Water Resources. 2009;32:582-93.

Hsu, C.W., Chang, C.C., Lin, C.J., 2010. A Practical Guide to Support Vector classification, http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.

Kavzoglu, T. and Colkesen, I. 2010. Investigation of the Effects of Kernel Functions in Satellite Image Classification Using Support Vector Machines. Harita Dergisi Temmuz 2010 Sayi 144

Koch J, Berger H, Henriksen HJ, Sonnenborg TO. Modelling of the shallow water table at high spatial resolution using Random Forests. Hydrology and Earth System Sciences Discussions. 2019:1-26.

Kumar, M., Raghuwanshi, N.S., Singh, R., Wallender, W.W., and Pruitt, W.O. 2002. Estimating evapotranspiration using artificial neural network. J. Irrig. Drain. Eng., 128(4), 224–233. doi:10.1061/(ASCE)0733-9437(2002)128:4(224)

Le, X.-H., Ho, H. V., Giha, L., & Sungho, J. 2019. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. Water. https://doi.org/10.3390/w11071387

Nema, M.K., Khare, D., Chandniha, S.K., 2017. Application of artificial intelligence to estimate the reference evapotranspiration in sub-humid Doon valley. Appl. Water Sci. 7, 3903–3910. https://doi.org/10.1007/s13201-017-0543-3

Nocedal, J. & Wright, S. 2006. Numerical Optimization. 664 p. Springer. DOI: 10.1007/978-0-387-40065-5

Maier HR, Dandy GC. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental Modelling & Software. 2000;15:101-24.

Maier HR, Jain A, Dandy GC, Sudheer KP. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. Environmental Modelling & Software. 2010;25:891-909.

McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics. 1943;5:115-33.

Pandey, P.K., Nyori, T., Pandey, V., 2017. Estimation of reference evapotranspiration using data driven techniques under limited data conditions. Model. Earth Syst. Environ. 3, 1449–1461. https://doi.org/10.1007/s40808-017-0367-z

Pereira LS, Perrier A. & Allen R.G. 1999. Evapotranspiration: concepts and future trends. J. Irrig. and Drain. Engrg., ASCE 125(2): 45–51.

Prasad R, Deo RC, Li Y, Maraseni T. 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. 2017;197:42-63.

Prasad R, Joseph L, Deo RC. Modeling and Forecasting Renewable Energy Resources for Sustainable Power Generation: Basic Concepts and Predictive Model Results. 2020;68:59-79.

Rana, G, Katerji, N. 2000. Measurement and estimation of actual evapotranspiration in the field under Mediterranean climate: A review. European Journal Of Agronomy, 13(2-3):125-153.

Reis, M.M., da Silva, A.J., Zullo Junior, J., Tuffi Santos, L.D., Azevedo, A.M., Lopes, É.M.G., 2019. Empirical and learning machine approaches to estimating reference evapotranspiration based on temperature data. Comput. Electron. Agric. 165, 104937. https://doi.org/10.1016/j.compag.2019.104937

Quej VH, Almorox J, Arnaldo JA, Saito L. ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment. Journal of Atmospheric and Solar-Terrestrial Physics. 2017;155:62-70.

Saggi, M.K., Jain, S., 2019. Reference evapotranspiration estimation and modeling of the Punjab Northern India using deep learning. Comput. Electron. Agric. 156, 387–398. https://doi.org/10.1016/j.compag.2018.11.031

Sattari, M.T., Pal, M., Yurekli, K., Unlukara, A., 2013. M5 model trees and neural network based modelling of ET0 in Ankara, Turkey. Turkish J. Eng. Environ. Sci. 37, 211–220. https://doi.org/10.3906/muh-1212-5

Shabani, S., Samadianfard, S., Sattari, M.T., Mosavi, A., Shamshirband, S., Kmet, T., Várkonyi-Kóczy, A.R., 2020. Modeling Pan Evaporation Using Gaussian Process Regression K-Nearest Neighbors Random Forest and Support Vector Machines; Comparative Analysis. Atmosphere (Basel). 11, 66. https://doi.org/10.3390/atmos11010066

Shamshirband, S., Amirmojahedi, M., Gocić, M., Akib, S., Petković, D., Piri, J., Trajkovic, S., 2016. Estimation of Reference Evapotranspiration Using Neural Networks and Cuckoo Search Algorithm. J. Irrig. Drain. Eng. 142, 04015044. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000949

Solomatine DP. Applications of data-driven modeling and machine learning in control of water resurces. In: Mohammadian M, Sarker RA, Yao X, editors. Computational intelligence in Control: Idea Group Publishing; 2002. p. 197–217.

Solomatine DP, Dulal KN. Model trees as an alternative to neural networks in rainfall—runoff modelling. Hydrological Sciences Journal. 2003;48:399-411.

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106, 7183–7192. https://doi.org/10.1029/2000JD900719

Vapnik, V. 2013. The Nature of Statistical Learning Theory; Springer Science & Business Media: Berlin, Germany.

Yaseen ZM, Jaafar O, Deo RC, Kisi O, Adamowski J, Quilty J, et al. Stream-flow forecasting using extreme learning machines: A case study in a semi-arid region in Iraq. Journal of Hydrology. 2016;542:603-14.

Young C-C, Liu W-C, Wu M-C. A physically based and machine learning hybrid approach for accurate rainfall-runoff modeling during extreme typhoon events. Applied Soft Computing. 2017;53:205-16.

Xu CY, Singh VP. 2001. Evaluation and generalization of temperature-based methods for calculating evaporation. Hydrological Processes 15(2):305-319. https://doi.org/10.1002/hyp.119

Wang, P., Liu, C., Li, Y., 2019. Estimation method for ET0 with PSO-LSSVM based on the HHT in cold and arid data-sparse area. Cluster Comput. 22, 8207–8216. https://doi.org/10.1007/s10586-018-1726-x

Wen, X., Si, J., He, Z., Wu, J., Shao, H., Yu, H., 2015. Support-Vector-Machine-Based Models for Modeling Daily Reference Evapotranspiration with Limited Climatic Data in Extreme Arid Regions. Water Resour. Manag. 29, 3195–3209. https://doi.org/10.1007/s11269-015-0990-2

Wu, L., Peng, Y., Fan, J., Wang, Y., 2019. Machine learning models for the estimation of monthly mean daily reference evapotranspiration based on cross-station and synthetic data. Hydrol. Res. 50, 1730–1750. https://doi.org/10.2166/nh.2019.060

Zhang, Y., Wei, Z., Zhang, L., Du, J., 2019. Applicability evaluation of different algorithms for daily reference evapotranspiration model in KBE system. Int. J. Comput. Sci. Eng. 18, 361. https://doi.org/10.1504/IJCSE.2019.099074