



Estimation of daily reference evapotranspiration with limited climatic data using machine learning approaches across different climate zones in New Mexico

Esmail Mokari^{1,2} · David DuBois² · Zohrab Samani¹ · Hamid Mohebzadeh³ · Koffi Djaman²

Received: 24 May 2021 / Accepted: 2 November 2021 / Published online: 9 November 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2021

Abstract

Reference evapotranspiration (ET_0) is a major factor for water resource management. Although the FAO Penman–Monteith model is the highly recommended for estimating ET_0 , its requirement of a complete climatic variables has made the application of this model complicated. The objective of this study was to investigate the potential of four machine learning (ML) models, namely extreme learning machine (ELM), genetic programming (GP), random forest (RF), and support vector regression (SVR), for estimating daily ET_0 with limited climatic data using a tenfold cross-validation method across different climate zones in New Mexico. Four input scenarios, namely S1 (T_{\max} (maximum air temperature), T_{\min} (minimum air temperature), RH_{ave} (average relative humidity), U_2 (wind speed at 2 m height), R_S (total solar radiation)), S2 (T_{\max} , T_{\min} , U_2 , R_S), S3 (T_{\max} , T_{\min} , R_S), and S4 (T_{ave} , R_S), were considered using climatic data during the 2009–2019 period from six selected weather stations across different climate zones. The results showed that the estimated daily ET_0 differed significantly following ML model types and input scenarios across different climate zones. The ML models under S1 scenario showed the best estimation accuracy during the testing stage in climate zones 1 and 5 (RMSE and MAE $< 0.5 \text{ mm day}^{-1}$). The ML models under S3 and S4 scenarios were found to be more preferred at climate zones 1, 5, and 8 (RMSE and MAE $< 1 \text{ mm day}^{-1}$). The estimation accuracy of ML models was decreased with lack of RH_{ave} and U_2 data in input scenarios although the ML models based on S4 scenario (only T_{ave} and R_S) showed acceptable ET_0 estimations particularly in the climate zone 5 ($0.5 \text{ mm day}^{-1} < \text{RMSE} < 0.6 \text{ mm day}^{-1}$). The SVR and ELM were the best ML models for all input scenarios in the studied climate zones where these models showed the best stabilities in the testing stages.

1 Introduction

Estimating crop water use (actual evapotranspiration, ET_a) is highly crucial as it is used for irrigation scheduling and planning (Kisi 2016). ET_a can be measured directly in the field using an eddy covariance flux tower (Samani et al. 2011) or weighing lysimeters (Ding et al. 2010). However, these methods are largely limited because of the high cost of instrumentation. Alternatively, ET_a can be calculated

using crop coefficient (K_c) and reference evapotranspiration (ET_0). ET_0 can be estimated using the FAO Penman–Monteith model (Allen et al. 1998) which is the most accurate and recommended method, but requiring several input climate variables has made this method pretty complex for its utilization (Mokari et al. 2021). Thus, developing models with fewer input climatic data is highly required particularly where the climatic data are not completed. Several investigators have developed simplified empirical ET_0 models using the limited input climatic data over the last decades (Hargreaves and Samani 1985; Romanenko 1961; Tabari et al. 2013). However, these models have showed less suitability for daily ET_0 estimation (Torres et al. 2011). The estimation of ET_0 can be a complex and non-linear process where it is quite complicated for empirical models to take into account these complex processes (Fan et al. 2018).

Alternatively, machine learning (ML) models have shown their capability to be used as powerful tools to estimate ET_0 since they do not require any specific knowledge of internal

✉ Esmail Mokari
emokari@nmsu.edu

¹ Dept. of Civil Engineering, New Mexico State Univ,
Las Cruces, NM 88003, USA

² Dept. of Plant and Environmental Sciences, New Mexico
State Univ, Las Cruces, NM 88003, USA

³ School of Engineering, University of Guelph, Guelph, ON,
Canada

variables (Wang et al. 2017). Several ML models including artificial neural networks (ANN), support vector machine (SVM), random forest (RF), extreme learning machine (ELM), and genetic programming (GP) have been investigated by various researchers to estimate ET_0 (Fan et al. 2018; Feng et al. 2017a, 2017b; Gocic et al. 2016; Traore et al. 2016; Wen et al. 2015; Yin et al. 2017). Among these models, the SVM and ELM models have showed the best estimation accuracies compared to the other ML models (Abdullah et al. 2015; Fan et al. 2018; Feng et al. 2017b; Patil and Deka 2016; Yin et al. 2017, 2016). Patil and Deka (2016) evaluated three ML models, namely ELM, SVM, and ANN. They reported similar performance when both ELM and SVM models were used to estimate weekly ET_0 where these two models outperformed the ANN model. Yin et al. (2017) forecasted ET_0 variability with aid of the ELM and SVM models in the north west of China. Both ML models showed great performances for estimating ET_0 where a slightly better performance was observed using the ELM model. Fan et al. (2018) compared the ELM, SVM, and four tree-based models to estimate daily ET_0 across China. They found that the ELM model slightly outperformed the SVM model where both models showed a better accuracy compared with the four tree-based models.

Although various ML models, the SVM and ELM models in particular, have been frequently used to estimate ET_0 , their potential to estimate ET_0 has not been comprehensively investigated in the regions with different climate zones. The related literature reviews show that several efforts made on estimating ET_0 using ML models have mainly focused on a specific climate zone including arid climate zone (Shiri 2018; Wen et al. 2015), semi-arid climate zone (Tabari et al. 2012), Mediterranean climate zone (Kisi 2016), a warm and humid climate zone (Feng et al. 2017b), and a maritime climate zone (Shiri et al. 2012). Wen et al. (2015) estimated daily ET_0 using the SVM model with four combinations of climatic data in the extremely arid regions of China. The findings showed the SVM model estimated daily ET_0 more accurately than the ANN and empirical models. Kisi (2016) investigated the performance of three different ML models, namely SVM, multivariate adaptive regression splines (MARS), and M5 Model Tree (M5Tree), to estimate ET_0 in the Mediterranean climate zone. They reported that the SVM model performed better than the MARS and M5Tree models.

To the best of our knowledge, there are limited studies on evaluating the potential of ML models for estimating ET_0 in the region with different climate zones. It is crucial to know how powerful the ML models are when the climate zone is changed. For examples, Shiri et al. (2014) comprehensively assessed the empirical, semi-empirical, ML models for estimating ET_0 across three different climate zones in Iran. The most accurate results were observed in the humid climate zone

while the poorest ET_0 estimations were found in the arid climate zone. New Mexico (NM) is comprised of eight climate zones where the climatic data (particularly air temperature) are highly variable between the different climate zones. Thus, the objectives of the present study were to (1) comprehensively evaluate the potential of ML models including SVM, ELM, GP, and RF for estimating daily ET_0 across different climate zones in NM during the 2009–2019 period where no studies are available on estimating daily ET_0 using ML models; and (2) assess the effects of different input combinations of climatic data on the estimation accuracy of daily ET_0 across different climate zones of NM.

2 Materials and methods

2.1 Study area

Based on the topographic features, the NM state is divided into eight climate zones (Karl and Koss 1984) where six of them including climate zones 1, 2, 3, 5, 7, and 8 were studied in this study (Fig. 1). The mean annual temperature varies from 20 °C in climate zone 8 to 4.4 °C in climate zone 2 in the north with high mountains and valleys (NM climate center: [Climate in New Mexico](#)). In summer, although daytime temperature can exceed 37 °C in climate zone 8, the average monthly maximum temperature during July as the warmest month is slightly over 32 °C. The average annual rainfall differs from less than 254 mm in the southern parts such climate zone 8 to more than 508 mm at higher elevations such as climate zone 2. The potential evaporation in the state is much higher than average annual rainfall and it can reach 1854 mm in the southeast parts such as climate zone 7 (NM climate center: [Climate in New Mexico](#)).

2.2 Data collection and input scenarios

Continuous time series of daily climatic data including maximum air temperature (T_{max}), minimum air temperature (T_{min}), maximum relative humidity (RH_{max}), minimum relative humidity (RH_{min}), wind speed at 2 m height (U_2), and total solar radiation (R_S) during the 2009–2019 period were collected from six weather stations across different climate zones of NM (Fig. 1). The collected data were analyzed to determine the missing and outlier data. Days with missing and outlier data were removed. The average relative humidity (RH_{ave}) was also calculated using RH_{max} and RH_{min} . The FAO Penman–Monteith model (Allen et al. 1998), the most accepted method for estimating ET_0 , was applied as follows:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma(900/(T_m + 273))u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

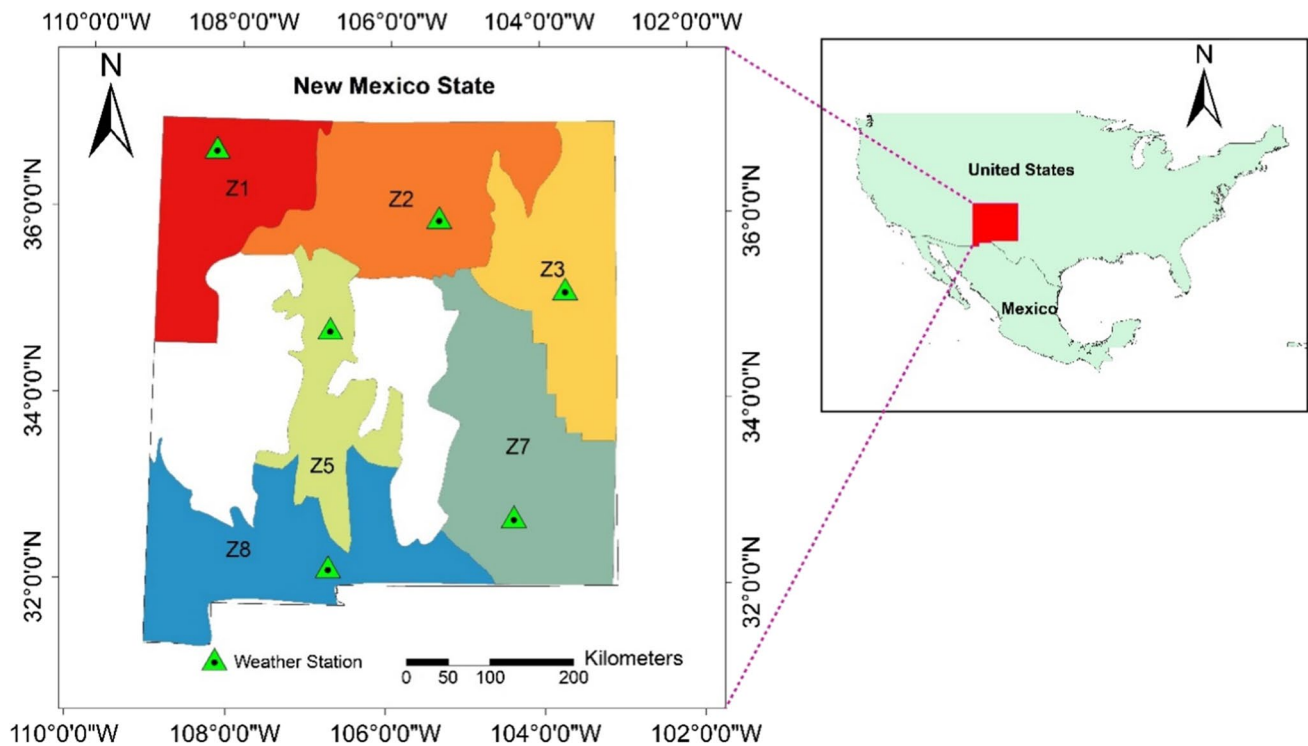


Fig. 1 The geographical locations of the six weather stations across different climate zones (Z1 to Z8) in NM state

where ET_0 is the reference evapotranspiration (mm/day), R_n is the net radiation at the crop surface ($\text{MJ m}^{-2} \text{day}^{-1}$), G is the soil heat flux density ($\text{MJ m}^{-2} \text{day}^{-1}$), T_m is the mean daily air temperature at 2 m height ($^{\circ}\text{C}$), u_2 is the wind speed at 2 m height (m/s), e_s is the saturation vapor pressure (kPa), e_a is the actual vapor pressure (kPa), $e_s - e_a$ is the saturation vapor pressure deficit (kPa), Δ is the slope of the vapor pressure curve ($\text{kPa } ^{\circ}\text{C}^{-1}$), and γ is the psychrometric constant ($\text{kPa } ^{\circ}\text{C}^{-1}$).

Four different input scenarios were determined to assess the effects of different input combinations of climatic data on the estimation accuracy of daily ET_0 across different climate zones. The scenarios were S1 (T_{\max} , T_{\min} , RH_{ave} , U_2 , R_S), S2 (T_{\max} , T_{\min} , U_2 , R_S), S3 (T_{\max} , T_{\min} , R_S), and S4 (T_{ave} , R_S).

2.3 Applied machine learning (ML) models

2.3.1 Extreme learning machine (ELM)

ELM, known as an advanced method of the single-hidden layer feed-forward neural networks (SLFNs), is a model with a single input layer, a hidden layer, and an output. This standard form of the model is classified as a type of ANN model. The computation process in the ELM is faster compared to the traditional ANN. In ELM, the hidden biases and input-hidden weights are generated randomly when the hidden nodes are selected.

Then, the hidden layer outputs are computed. Finally, the hidden-output weights are determined using the Moore–Penrose generalized inverse. More detailed information about this model can be found in the literature (Huang et al. 2006).

2.3.2 Genetic programming (GP)

GP is a data-driven technique developed by Koza (1992) which is used for finding a highly fit individual in the space of possible solutions. In this method, individuals are mathematical formulas created by combinations of functions such as $\sin(\alpha)$ and variables. The GP model applies evolutionary computation to find the best individual for the optimized fitness values. Generally, the GP model follows five steps to find the fittest individual: (1) an initial random population of individuals composed of functions and variables is created; (2) the fitness of each individual in the population is validated with a problem-specific fitness function and the most appropriate individuals are selected to survive in the new population as parents; (3) once parents are selected, they create better types known as offspring or new generations by producing algorithms known as genetic operators; (4) then, the individuals are assessed for fitness; and (5) the process from (2) to (4) is repeated over several generations until an individual satisfies a given success criterion.

2.3.3 Random forest (RF)

RF is an ensemble ML model which has been widely applied for several regression and classification problems (Breiman 2001). This model includes several random and simple decision trees. Fundamentally, the target of the RF model is to create a large random subset of decorrelated regression trees with bootstrap from samples and features. This model is divided into two main parts, i.e., randomness and ensemble learning. More details about this model can be found in Breiman (2001).

2.3.4 Support vector regression (SVR)

SVR is a frequently used ML model for classification and regression purposes (Cortes and Vapnik 1995). The structural risk reduction (SRR) concept is employed in this model as an alternative of the empirical risk reduction concept which is frequently used by ANN models. Based on the SRR concept, the upper bound to the generalization error is minimized instead of the training error which is resulted in an optimum network structure (Lin et al. 2006). The SVR is originated in a fundamental hypothesis known as the nonlinear mapping of the principal data into a higher dimensional feature space. The performance of linear regression in the feature space is employed by the kernels. The radial basis function (RBF) is found to be the best kernel among several kernels used in the SVR (Barzegar et al. 2017). Therefore, the RBF kernel was applied in the present study.

2.4 Cross-validation and model parameterization

In this study, a tenfold cross-validation method was applied for the training period to determine the optimum parameters for the applied ML models. Then, the optimum values of each ML model were used to estimate ET_0 for the testing period. Normalization, which is a part of data preparation for ML models, was also applied to match the consistency of the ML models. Table 1 shows the optimized values of the four different ML models with different input scenarios at six different climate zones.

In ELM, the weights and biases of the hidden layer are generated using the random computation. The random initialization of the weights in ELM can result in different outputs of the networks for identical numbers of neurons. To find the best weights, 1000 ELMs with the selected number of hidden neurons were trained in the training period and the best weights that minimize the objective function were maintained. Then, the selected structure with optimized weights was used to estimate ET_0 in the validation phase. A three-layer ELM model with a sigmoid activation function was developed. Therefore, the optimum number of hidden neurons was found using the tenfold cross-validation approach.

In GP, the two main parameters population size and generation size need to be optimized to produce a great performance. The optimum parameters were determined with aid of the harmony search algorithm (Zong Woo et al. 2001) using the tenfold cross-validation method.

With respect to the RF model, the main parameter which is known as ntree (the number of decision trees) was determined using the tenfold cross-validation approach based on the reported numerical ranges by Belgiu and Drăguț (2016).

In SVR, the RBF kernel function was applied including three main parameters of structural parameter (γ), penalty coefficient (C), and tolerance threshold (ϵ). The different values for each parameter including γ (20 values between 0.0001 and 10,000), C (20 values from 0.0001 to 10,000), and ϵ (10 values between 0.001 and 1) were assessed and the optimized values were determined (Zhang et al. 2018).

2.5 Model performance evaluation

Quantitative measures (Despotovic et al. 2015) including coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) were used to assess the performance of the different ML models for estimating daily ET_0 as follows:

$$R^2 = \frac{\sum_{i=1}^n (P_i - PM_i)^2}{\sum_{i=1}^n (PM_i - PM_{ave})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (P_i - PM_i)^2}{N}} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_i - PM_i| \quad (4)$$

where P_i is the i th value of estimated ET_0 for ML models, PM_i is the i th value of ET_0 for the FAO Penman–Monteith model, PM_{ave} is the average of ET_0 values for the FAO Penman–Monteith model, and N is the number of paired values.

Higher values of R^2 (closer to 1) show more efficient models while lower values of RMSE and MAE indicate a better model performance.

3 Results and discussions

The statistical results of the four ML models employing different input scenarios to estimate daily ET_0 across different climate zones in NM are shown in Tables 2, 3, 4, 5, 6, and 7. It is clear that the estimated daily ET_0 values varied significantly following ML model types and input scenarios. For example, with respect to climate zone 5

Table 1 Optimized parameters of the four different ML models with different input scenarios at six different climate zones

ML model		ELM	GP		RF	SVR		
Optimized parameters		<i>n</i>	<i>P</i>	<i>G</i>	<i>nt</i>	γ	<i>C</i>	ϵ
Zone 1	S1	94	4500	21	100	0.056	334.22	0.001
	S2	70	4500	31	100	0.056	334.22	0.001
	S3	21	3500	11	61	0.167	334.22	0.001
	S4	9	3500	21	98	0.222	334.22	0.001
Zone 2	S1	72	2500	31	67	0.167	334.22	0.112
	S2	83	4500	11	84	0.5	1	0.001
	S3	17	4500	11	99	0.389	1	0.001
	S4	13	2500	11	51	0.001	2000.33	0.112
Zone 3	S1	83	4500	41	90	0.056	334.22	0.001
	S2	67	4500	41	98	0.056	667.44	0.001
	S3	10	4000	41	99	0.167	1000.66	0.001
	S4	43	2500	41	76	0.333	1333.88	0.001
Zone 5	S1	93	3000	31	100	0.056	334.22	0.001
	S2	43	4000	21	95	0.111	1	0.001
	S3	42	2500	21	94	0.056	334.22	0.001
	S4	9	4000	11	44	0.278	1	0.001
Zone 7	S1	69	4000	41	97	0.056	3000	0.112
	S2	53	4500	21	100	0.056	1000.66	0.112
	S3	33	4500	11	100	0.111	667.44	0.001
	S4	5	2500	11	64	0.222	334.22	0.223
Zone 8	S1	82	2500	11	90	0.056	334.22	0.001
	S2	73	4500	21	92	0.056	334.22	0.001
	S3	50	3500	21	98	0.056	334.22	0.001
	S4	12	1000	11	91	0.222	667.44	0.001

n nodes, *P* population, *G* generation, *nt* number of decision trees, γ structural parameter, *C* penalty coefficients, ϵ tolerance threshold

ELM extreme learning machine, *GP* genetic programming, *RF* random forest, *SVR* support vector regression

S1: T_{max} (maximum air temperature), T_{min} (minimum air temperature), RH_{ave} (average relative humidity), U_2 (wind speed at 2 m height), R_S (total solar radiation); S2: T_{max} , T_{min} , U_2 , R_S ; S3: T_{max} , T_{min} , R_S ; S4: T_{ave} , R_S

(Table 5), the RF model was found to be the best ML model for all input scenarios during the training stage where the highest R^2 and lowest RMSE and MAE were observed. The SVR and ELM models indicated great performances to estimate daily ET_0 for all input scenarios during the testing stage. However, the GP and RF models were found to be comparable models for estimating daily ET_0 to the SVR and ELM models. Apparently, all ML models under S1 and S2 scenarios showed the best estimation accuracy compared with other input scenarios (Table 5). The average RMSE of S1 and S2 scenarios during the training and testing stages was 0.18 mm day^{-1} and 0.22 mm day^{-1} , respectively. However, ML models under S4 scenario (only T_{ave} and R_S as input climatic data) were also able to estimate daily ET_0 pretty accurately with having an average RMSE of 0.48 mm day^{-1}

and 0.56 mm day^{-1} during the training and testing stages, respectively (Table 5).

All applied ML models showed different accuracies under various input scenarios across different climate zones (Tables 2, 3, 4, 5, 6, and 7). Similar to climate zone 5, the ML models under S1 and S2 scenarios had the best estimation accuracy during the testing stage across other climate zones. Both the SVR and ELM models provided the best estimation of daily ET_0 for all input scenarios across all studied climate zones during the testing stage followed by RF and GP models with acceptable accuracy (Tables 2, 3, 4, 5, 6, and 7). Therefore, the accuracy ranking is $SVR = ELM > RF > GP$ according to the statistical indicators provided in the Tables 2, 3, 4, 5, 6, and 7. With respect to the lack of complete dataset, daily ET_0 estimated by ML models under S1 and S2 scenarios were observed

Table 2 Statistical values of the four different ML models with different input scenarios during the training and testing at climate zone 1

Input scenario/ Model	Training			Testing		
	MAE (mm/day)	RMSE (mm/day)	R ²	MAE (mm/day)	RMSE (mm/day)	R ²
S1						
ELM	0.111	0.150	0.995	0.120	0.166	0.994
GP	0.313	0.422	0.967	0.316	0.424	0.967
RF	0.071	0.1	0.998	0.182	0.257	0.988
SVR	0.104	0.146	0.996	0.113	0.154	0.995
S2						
ELM	0.166	0.228	0.990	0.182	0.250	0.988
GP	0.291	0.413	0.970	0.289	0.409	0.970
RF	0.083	0.120	0.997	0.221	0.301	0.983
SVR	0.160	0.226	0.990	0.174	0.245	0.988
S3						
ELM	0.475	0.659	0.920	0.485	0.686	0.912
GP	0.533	0.767	0.895	0.537	0.759	0.895
RF	0.187	0.271	0.986	0.506	0.724	0.902
SVR	0.452	0.670	0.920	0.469	0.701	0.912
S4						
ELM	0.486	0.692	0.911	0.494	0.703	0.907
GP	0.548	0.803	0.894	0.545	0.795	0.894
RF	0.20	0.289	0.984	0.531	0.762	0.891
SVR	0.478	0.712	0.909	0.480	0.714	0.908

ELM extreme learning machine, GP genetic programming, RF random forest, SVR support vector regression, MAE mean absolute error, RMSE root mean square error, R² coefficient of determination

S1: T_{max} (maximum air temperature), T_{min} (minimum air temperature), RH_{ave} (average relative humidity), U₂ (wind speed at 2 m height), R_S (total solar radiation); S2: T_{max}, T_{min}, U₂, R_S; S3: T_{max}, T_{min}, R_S; S4: T_{ave}, R_S

The bold values are the best statistical measures among all ML models

Table 3 Statistical values of the four different ML models with different input scenarios during the training and testing at climate zone 2

Input scenario/ Model	Training			Testing		
	MAE (mm/day)	RMSE (mm/day)	R ²	MAE (mm/day)	RMSE (mm/day)	R ²
S1						
ELM	0.284	0.418	0.920	0.300	0.442	0.903
GP	0.437	0.636	0.825	0.416	0.600	0.826
RF	0.110	0.169	0.987	0.292	0.442	0.902
SVR	0.215	0.004	0.930	0.254	0.442	0.903
S2						
ELM	0.321	0.462	0.902	0.326	0.478	0.887
GP	0.438	0.627	0.827	0.416	0.585	0.831
RF	0.128	0.189	0.984	0.329	0.492	0.880
SVR	0.292	0.476	0.901	0.291	0.473	0.890
S3						
ELM	0.539	0.707	0.772	0.518	0.677	0.774
GP	0.607	0.807	0.708	0.581	0.759	0.717
RF	0.214	0.286	0.965	0.537	0.715	0.750
SVR	0.517	0.734	0.766	0.486	0.690	0.768
S4						
ELM	0.570	0.753	0.742	0.542	0.715	0.749
GP	0.616	0.819	0.703	0.589	0.770	0.709
RF	0.237	0.319	0.956	0.588	0.789	0.70
SVR	0.573	0.771	0.734	0.538	0.719	0.743

ELM extreme learning machine, GP genetic programming, RF random forest, SVR support vector regression, MAE mean absolute error, RMSE root mean square error, R² coefficient of determination

S1: T_{max} (maximum air temperature), T_{min} (minimum air temperature), RH_{ave} (average relative humidity), U₂ (wind speed at 2 m height), R_S (total solar radiation); S2: T_{max}, T_{min}, U₂, R_S; S3: T_{max}, T_{min}, R_S; S4: T_{ave}, R_S

The bold values are the best statistical measures among all ML models

Table 4 Statistical values of the four different ML models with different input scenarios during the training and testing at climate zone 3

Input scenario/ Model	Training			Testing		
	MAE (mm/day)	RMSE (mm/day)	R ²	MAE (mm/day)	RMSE (mm/day)	R ²
S1						
ELM	0.117	0.159	0.996	0.129	0.178	0.995
GP	0.445	0.627	0.949	0.452	0.635	0.945
RF	0.075	0.107	0.998	0.193	0.268	0.99
SVR	0.106	0.150	0.997	0.116	0.163	0.996
S2						
ELM	0.225	0.306	0.987	0.240	0.342	0.983
GP	0.351	0.489	0.970	0.363	0.503	0.968
RF	0.104	0.146	0.997	0.288	0.394	0.978
SVR	0.228	0.320	0.986	0.241	0.340	0.984
S3						
ELM	0.676	0.893	0.895	0.680	0.903	0.889
GP	0.771	1.044	0.868	0.763	1.037	0.869
RF	0.268	0.362	0.983	0.717	0.960	0.874
SVR	0.656	0.909	0.894	0.674	0.919	0.890
S4						
ELM	0.713	0.935	0.885	0.720	0.937	0.879
GP	0.781	1.041	0.861	0.861	1.030	0.859
RF	0.291	0.391	0.980	0.794	1.040	0.852
SVR	0.702	0.953	0.883	0.716	0.959	0.878

ELM extreme learning machine, GP genetic programming, RF random forest, SVR support vector regression, MAE mean absolute error, RMSE root mean square error, R² coefficient of determination

S1: T_{max} (maximum air temperature), T_{min} (minimum air temperature), RH_{ave} (average relative humidity), U₂ (wind speed at 2 m height), R_S (total solar radiation); S2: T_{max}, T_{min}, U₂, R_S; S3: T_{max}, T_{min}, R_S; S4: T_{ave}, R_S
The bold values are the best statistical measures among all ML models

Table 5 Statistical values of the four different ML models with different input scenarios during the training and testing at climate zone 5

Input scenario/ Model	Training			Testing		
	MAE (mm/day)	RMSE (mm/day)	R ²	MAE (mm/day)	RMSE (mm/day)	R ²
S1						
ELM	0.104	0.141	0.991	0.118	0.154	0.989
GP	0.272	0.358	0.944	0.261	0.347	0.946
RF	0.067	0.095	0.996	0.164	0.229	0.977
SVR	0.103	0.140	0.991	0.111	0.146	0.99
S2						
ELM	0.110	0.151	0.989	0.121	0.164	0.988
GP	0.255	0.344	0.953	0.257	0.344	0.954
RF	0.066	0.092	0.996	0.168	0.234	0.976
SVR	0.121	0.165	0.988	0.125	0.167	0.987
S3						
ELM	0.396	0.519	0.883	0.399	0.535	0.873
GP	0.433	0.589	0.856	0.413	0.568	0.862
RF	0.161	0.214	0.980	0.420	0.556	0.863
SVR	0.394	0.543	0.876	0.380	0.531	0.878
S4						
ELM	0.419	0.544	0.871	0.411	0.541	0.870
GP	0.444	0.581	0.857	0.431	0.570	0.857
RF	0.174	0.232	0.977	0.460	0.603	0.840
SVR	0.418	0.561	0.866	0.404	0.553	0.866

ELM extreme learning machine, GP genetic programming, RF random forest, SVR support vector regression, MAE mean absolute error, RMSE root mean square error, R² coefficient of determination

S1: T_{max} (maximum air temperature), T_{min} (minimum air temperature), RH_{ave} (average relative humidity), U₂ (wind speed at 2 m height), R_S (total solar radiation); S2: T_{max}, T_{min}, U₂, R_S; S3: T_{max}, T_{min}, R_S; S4: T_{ave}, R_S
The bold values are the best statistical measures among all ML models

Table 6 Statistical values of the four different ML models with different input scenarios during the training and testing at climate zone 7

Input scenario/Model	Training			Testing		
	MAE (mm/day)	RMSE (mm/day)	R^2	MAE (mm/day)	RMSE (mm/day)	R^2
S1						
ELM	0.347	0.474	0.959	0.384	0.519	0.951
GP	0.604	0.794	0.887	0.588	0.779	0.892
RF	0.145	0.205	0.992	0.412	0.576	0.940
SVR	0.295	0.486	0.959	0.339	0.544	0.948
S2						
ELM	0.477	0.631	0.927	0.496	0.669	0.918
GP	0.651	0.852	0.871	0.662	0.869	0.868
RF	0.195	0.265	0.987	0.510	0.696	0.912
SVR	0.453	0.639	0.926	0.471	0.666	0.920
S3						
ELM	0.961	1.25	0.712	1.033	1.352	0.670
GP	1.098	1.424	0.632	1.106	1.439	0.625
RF	0.376	0.503	0.958	1.102	1.443	0.628
SVR	0.962	1.286	0.701	1.01	1.356	0.668
S4						
ELM	1.067	1.384	0.651	1.137	1.471	0.609
GP	1.155	1.498	0.592	1.186	1.534	0.574
RF	0.434	0.572	0.946	1.213	1.572	0.565
SVR	1.079	1.416	0.636	1.115	1.454	0.617

ELM extreme learning machine, *GP* genetic programming, *RF* random forest, *SVR* support vector regression, *MAE*: mean absolute error, *RMSE* root mean square error, R^2 coefficient of determination

S1: T_{\max} (maximum air temperature), T_{\min} (minimum air temperature), RH_{ave} (average relative humidity), U_2 (wind speed at 2 m height), R_S (total solar radiation); S2: T_{\max} , T_{\min} , U_2 , R_S ; S3: T_{\max} , T_{\min} , R_S ; S4: T_{ave} , R_S

The bold values are the best statistical measures among all ML models

more accurately compared with the calculated daily ET_0 values by the FAO Penman–Monteith model across all studied climate zones. However, the ML models under S3 and S4 scenarios were found to be more preferred at climate zones 1, 5, and 8 (Tables 2, 3, 4, 5, 6, and 7).

Input scenarios had a major key in the estimation accuracy of ML models. The ML models under S1 scenario produced a better accuracy compared with other scenarios although the difference between S1 and S2 scenarios was negligible for some climate zones. The findings showed that the estimation accuracy of ML models was decreased with lack of RH_{ave} and U_2 data in input scenarios where this reduction was the worst in climate zone 7 (RMSE and MAE > 1.5 mm day⁻¹). Therefore, RH_{ave} and U_2 data played a key role in the estimation accuracy of daily ET_0 using ML models across different climate zones in NM. However, the results of ML models based on S4 scenario (only T_{ave} and R_S) showed acceptable ET_0 estimations particularly in climate zone 5 where RMSE varied between 0.5 and 0.6 mm day⁻¹. Findings are in agreement with previous studies which showed more input climatic data improved the model estimation accuracy but

the contribution of climatic data for estimating ET_0 varied across different climate zones (Antonopoulos and Antonopoulos 2017; Fan et al. 2018).

Figure 2 shows the scatter plots of the calculated ET_0 by the FAO-PM model and estimated values by the four ML models for the best scenario under different climate zones in the testing stage. The SVR model provided more scattered estimations for climate zones 1, 2, 3, 5, and 8, whereas the ELM model produced more scattered estimations for climate zone 7 (Fig. 2). Generally, the estimated ET_0 by SVR and ELM models were observed to be closer to the calculated ET_0 by the FAO-PM model (Fig. 2). This trend showed that SVR and ELM models produced accurate estimations of daily ET_0 . The findings are in agreement with previous studies. Fan et al. (2018) reported the ELM and SVM models as the best combination of estimation accuracy and stability for estimating ET_0 in different climate zones of China. Wen et al. (2015) showed the potential of the SVR model than the ANN model for the accurate estimation of daily ET_0 in the extreme arid regions of China. Feng et al. (2017b) reported that the ELM model could be successfully used for estimating ET_0 in southwest of China.

Table 7 Statistical values of the four different ML models with different input scenarios during the training and testing at climate zone 8

Input scenario/Model	Training			Testing		
	MAE (mm/day)	RMSE (mm/day)	R ²	MAE (mm/day)	RMSE (mm/day)	R ²
S1						
ELM	0.089	0.119	0.996	0.095	0.136	0.995
GP	0.288	0.381	0.964	0.280	0.369	0.966
RF	0.058	0.084	0.998	0.152	0.214	0.988
SVR	0.085	0.113	0.996	0.090	0.124	0.996
S2						
ELM	0.133	0.187	0.991	0.146	0.214	0.988
GP	0.327	0.503	0.942	0.320	0.505	0.941
RF	0.069	0.099	0.997	0.189	0.272	0.981
SVR	0.133	0.190	0.990	0.142	0.209	0.988
S3						
ELM	0.481	0.640	0.895	0.485	0.645	0.893
GP	0.625	0.823	0.835	0.617	0.809	0.841
RF	0.191	0.260	0.983	0.502	0.669	0.885
SVR	0.483	0.661	0.890	0.489	0.658	0.891
S4						
ELM	0.496	0.662	0.888	0.504	0.663	0.887
GP	0.571	0.750	0.858	0.566	0.749	0.858
RF	0.199	0.272	0.981	0.535	0.716	0.869
SVR	0.491	0.669	0.887	0.498	0.671	0.887

ELM extreme learning machine, GP genetic programming, RF random forest, SVR support vector regression, MAE mean absolute error, RMSE: root mean square error, R² coefficient of determination

S1: T_{max} (maximum air temperature), T_{min} (minimum air temperature), RH_{ave} (average relative humidity), U₂ (wind speed at 2 m height), R_S (total solar radiation); S2: T_{max}, T_{min}, U₂, R_S; S3: T_{max}, T_{min}, R_S; S4: T_{ave}, R_S

The bold values are the best statistical measures among all ML models

Figure 3 shows the training and testing RMSE of the best ML model in each climate zone for various input scenarios. SVR and ELM were the best ML models in all climate zones for all input scenarios (Fig. 3). For S1 and S2 scenarios, the corresponding models provided the best estimation accuracy (RMSE < 0.5 mm day⁻¹) in both the training and testing stages for all climate zones except climate zone 7 (Fig. 3). However, the corresponding ML models showed acceptable estimation accuracy (RMSE < 1 mm day⁻¹) when S3 and S4 scenarios were employed (Fig. 3). The percentage increase in testing RMSE over training RMSE for the best ML model in each climate zone under various input scenarios is also shown in Fig. 3. The SVR and ELM models provided the highest stability in the testing stage where either decreases or the smallest increases in RMSE were observed (Fig. 3). The stability of ML models has been a key factor for estimating ET₀ because it can affect the estimation accuracy significantly. Fan et al. (2018) reported a large percentage increase in testing RMSE when the RF and M5Tree models were used to estimate daily ET₀ across China. However, they found the SVM and ELM models as the

most stable models with the RMSE of less than 10.1% in the testing stage.

4 Conclusion

The present study assessed the potential of ML models including extreme learning machine (ELM), genetic programming (GP), random forest (RF), and support vector regression (SVR) for estimating daily ET₀ using various input scenarios across different climate zones in NM during the 2009–2019 period. Findings showed that the estimation accuracy of daily ET₀ values was a function of ML model types and input scenarios across different climate zones. Both the SVR and ELM models provided the most accurate estimation of daily ET₀ during the testing stage followed by RF and GP models with acceptable accuracy in all studied climate zones. Daily ET₀ estimated by ML models under S1 and S2 scenarios were found more accurate compared with the calculated daily ET₀ values by the FAO Penman–Monteith model across all studied climate zones. However, the ML models under S3 and S4 scenarios were more preferred at climate zones 1, 5, and 8.

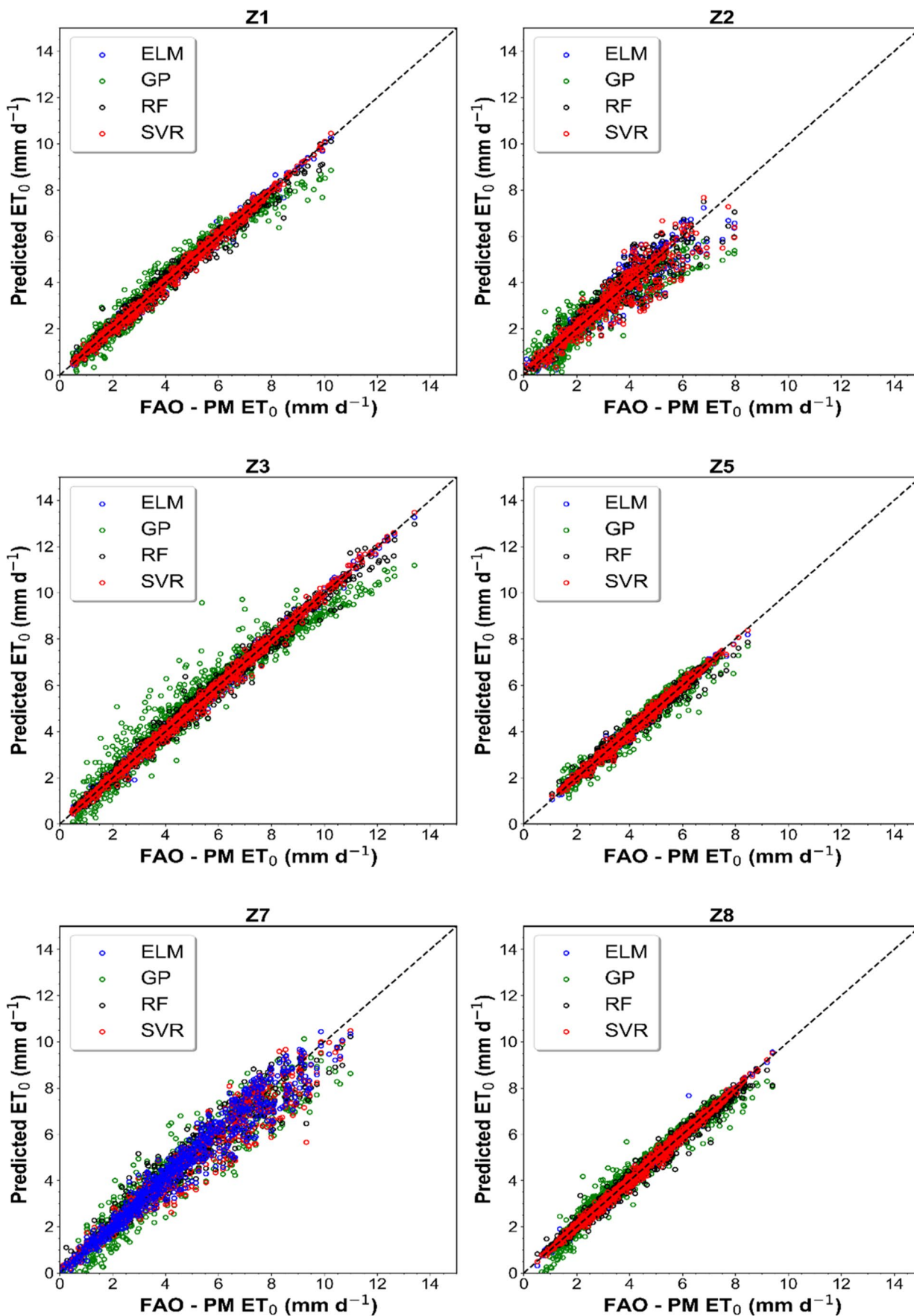


Fig. 2 Scatter plots of the calculated reference evapotranspiration (ET_0) by the FAO Penman–Monteith model (FAO-PM) and the estimated values by the four different ML models for best scenario across various climate zones in the testing stage. ELM extreme learning machine, GP genetic programming, RF random forest, SVR support vector regression. Z1 climate zone 1, Z2 climate zone 2, Z3 climate zone 3, Z5 climate zone 5, Z7 climate zone 7, Z8 climate zone 8

Input scenarios showed significant effects on the estimation accuracy of ML models. The ML models under S1 and S2 scenarios showed better accuracies than other scenarios. The estimation accuracy was decreased under missing RH_{ave} and U_2 data in input scenarios where this reduction was the worst in climate zone 7 (average RMSE of 1.5 mm day^{-1}). Therefore, RH_{ave} and U_2 data had a major role in the estimation accuracy of daily ET_0 across different climate zones in NM. With respect to the best input scenario for each climate zone, the SVR model showed more scattered estimations for climate zones 1, 2, 3, 5, and 8 whereas the ELM model produced more scattered estimations for climate zone 7. The SVR and ELM models offered the highest stability in the testing stage where either decreases or the smallest increases (less than 10%) in RMSE were found.

Findings provide guidelines for future investigators who need to study specific climate zones and identify appropriate ML models for the climate zone. The SVR model can be effectively applied to estimate ET_0 in regions where the mean annual temperature fluctuates between 4 and 20 °C. This model also has potential to estimate ET_0 in dried regions where the average monthly maximum temperature exceeds 32 °C. Estimation of ET_0 in windy climate zones can bring additional challenges. The results of this study suggest that the ELM model can be used for those regions. In addition, the results of this present study can be applied to forecast agriculture/rangeland productivity which is crucial for agricultural planning. As an example, estimated ET_0 using ML models in this study can be used to estimate rangeland aboveground biomass across New Mexico which is vitally important for grazing management. Rangeland’s production is directly affected by ET_0 . Thus, a model can relate estimated ET_0 using ML models to aboveground biomass. ML models are more convenient and comparably faster to be implemented than other models particularly when climate data are limited which was the case in this study. Generally, estimated ET_0 using ML models can be used as an input

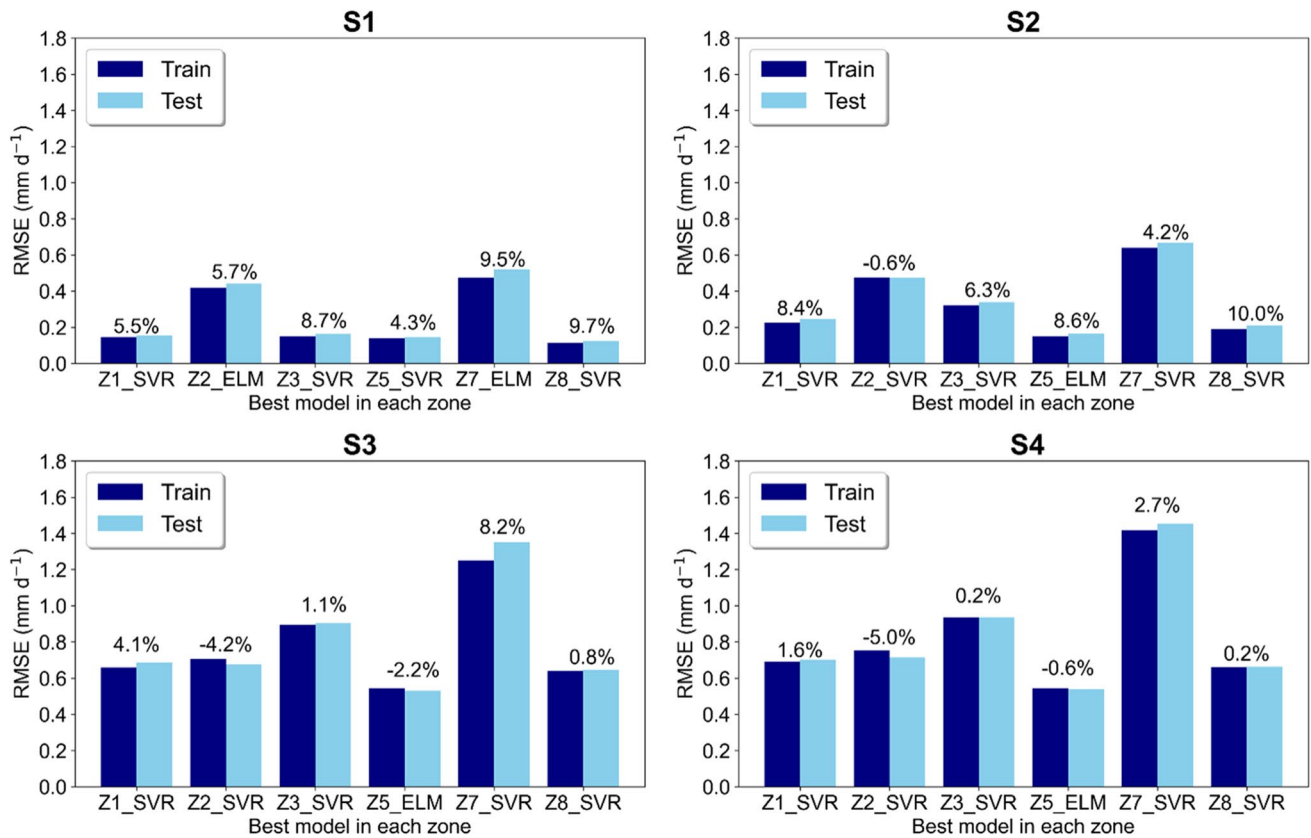


Fig. 3 Percentage increase in testing root mean square error (RMSE) over training RMSE for different input scenarios (S1, S2, S3, and S4) with the best ML model for each climate zone. ELM extreme learning

machine, GP genetic programming, RF random forest, SVR support vector regression. Z1 climate zone 1, Z2 climate zone 2, Z3 climate zone 3, Z5 climate zone 5, Z7 climate zone 7, Z8 climate zone 8

layer for variety of decision-making models where precision agriculture is practiced.

Acknowledgements The authors thank the National Institute of Food and Agriculture, US Department of Agriculture (award number: 2017-68007-26318) for supporting this work. The authors also thank New Mexico State University Climate Center for providing long-term climate data across New Mexico state.

Author contribution Conceptualization: E.M. and D.D.; Methodology: E.M.; H.M., and Z.S.; Technical investigation: D.D. and Z.S.; Writing (original draft preparation): E.M.; Writing (review and editing): K.D.; Supervision: Z.S. All authors have read and agreed to the published version of the manuscript.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

References

- Abdullah SS, Malek MA, Abdullah NS, Kisi O, Yap KS (2015) Extreme learning machines: a new approach for prediction of reference evapotranspiration. *J Hydrol* 527:184–195. <https://doi.org/10.1016/j.jhydrol.2015.04.073>
- Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration-guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56 Fao. Rome 300:D05109
- Antonopoulos VZ, Antonopoulos AV (2017) Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. *Comput Electron Agric* 132:86–96. <https://doi.org/10.1016/j.compag.2016.11.011>
- Barzegar R, Asghari Moghaddam A, Adamowski J, Fijani E (2017) Comparison of machine learning models for predicting fluoride contamination in groundwater. *Stoch Env Res Risk Assess* 31:2705–2718. <https://doi.org/10.1007/s00477-016-1338-z>
- Belgiu M, Drăguț L (2016) Random forest in remote sensing: a review of applications and future directions. *ISPRS J Photogramm Remote Sens* 114:24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32. <https://doi.org/10.1023/A:1010933404324>
- Climate in New Mexico, NM climate center, NMSU. <https://weather.nmsu.edu/climate/about/>
- Cortes C, Vapnik V (1995) Support-Vector Networks. *Mach Learn* 20:273–297. <https://doi.org/10.1007/bf00994018>
- Despotovic M, Nedic V, Despotovic D, Cvetanovic S (2015) Review and statistical analysis of different global solar radiation sunshine models. *Renew Sustain Energy Rev* 52:1869–1880. <https://doi.org/10.1016/j.rser.2015.08.035>
- Ding R, Kang S, Li F, Zhang Y, Tong L, Sun Q (2010) Evaluating eddy covariance method by large-scale weighing lysimeter in a maize field of northwest China. *Agric Water Manag* 98:87–95. <https://doi.org/10.1016/j.agwat.2010.08.001>
- Fan J et al (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, Gong D, Zhang Q, Zhao L (2017a) Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. *Agric Water Manag* 193:163–173. <https://doi.org/10.1016/j.agwat.2017.08.003>
- Feng Y, Peng Y, Cui N, Gong D, Zhang K (2017b) 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>
- Gocic M, Petković D, Shamshirband S, Kamsin A (2016) Comparative analysis of reference evapotranspiration equations modelling by extreme learning machine. *Comput Electron Agric* 127:56–63. <https://doi.org/10.1016/j.compag.2016.05.017>
- Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. *Appl Eng Agric* 1:96–99. <https://doi.org/10.13031/2013.26773>
- Huang G-B, Zhu Q-Y, Siew C-K (2006) Extreme Learning Machine: Theory and Applications. *Neurocomputing* 70:489–501. <https://doi.org/10.1016/j.neucom.2005.12.126>
- Karl TR, WJ Koss (1984) U.S. climate divisions, NOAA. <https://www.ncdc.noaa.gov/monitoring-references/maps/usclimate-divisions.php#history>
- Kisi O (2016) Modeling reference evapotranspiration using three different heuristic regression approaches. *Agric Water Manag* 169:162–172. <https://doi.org/10.1016/j.agwat.2016.02.026>
- Koza JR (1992) Genetic programming: on the programming of computers by means of natural selection vol 1. MIT press,
- Lin J-Y, Cheng C-T, Chau K-W (2006) Using support vector machines for long-term discharge prediction. *Hydrol Sci J* 51:599–612. <https://doi.org/10.1623/hysj.51.4.599>
- Mokari E, Samani Z, Heerema R, Ward F (2021) Evaluation of long-term climate change impact on the growing season and water use of mature pecan in Lower Rio Grande Valley. *Agric Water Manag* 252:106893. <https://doi.org/10.1016/j.agwat.2021.106893>
- Patil AP, Deka PC (2016) An extreme learning machine approach for modeling evapotranspiration using extrinsic inputs. *Comput Electron Agric* 121:385–392. <https://doi.org/10.1016/j.compag.2016.01.016>
- Romanenko VA (1961) Computation of the autumn soil moisture using a universal relationship for a large area Proc of Ukrainian Hydrometeorological Research Institute 3:12-25
- Samani Z, Bawazir S, Skaggs R, Longworth J, Piñon A, Tran V (2011) A simple irrigation scheduling approach for pecans. *Agric Water Manag* 98:661–664. <https://doi.org/10.1016/j.agwat.2010.11.002>
- Shiri J (2018) Improving the performance of the mass transfer-based reference evapotranspiration estimation approaches through a coupled wavelet-random forest methodology. *J Hydrol* 561:737–750. <https://doi.org/10.1016/j.jhydrol.2018.04.042>
- Shiri J, Kişi Ö, Landeras G, López JJ, Nazemi AH, Stuyt LCPM (2012) Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain). *J Hydrol* 414–415:302–316. <https://doi.org/10.1016/j.jhydrol.2011.11.004>
- Shiri J, Nazemi AH, Sadraddini AA, Landeras G, Kisi O, Fakhri Fard A, Marti P (2014) Comparison of heuristic and empirical

- approaches for estimating reference evapotranspiration from limited inputs in Iran. *Comput Electron Agric* 108:230–241. <https://doi.org/10.1016/j.compag.2014.08.007>
- Tabari H, Grismer ME, Trajkovic S (2013) Comparative analysis of 31 reference evapotranspiration methods under humid conditions. *Irrig Sci* 31:107–117. <https://doi.org/10.1007/s00271-011-0295-z>
- Tabari H, Kisi O, Ezani A, HosseinzadehTalaee P (2012) SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. *J Hydrol* 444–445:78–89. <https://doi.org/10.1016/j.jhydrol.2012.04.007>
- Torres AF, Walker WR, McKee M (2011) Forecasting daily potential evapotranspiration using machine learning and limited climatic data. *Agric Water Manag* 98:553–562. <https://doi.org/10.1016/j.agwat.2010.10.012>
- Traore S, Luo Y, Fipps G (2016) Deployment of artificial neural network for short-term forecasting of evapotranspiration using public weather forecast restricted messages. *Agric Water Manag* 163:363–379. <https://doi.org/10.1016/j.agwat.2015.10.009>
- Wang L, Kisi O, Zounemat-Kermani M, Li H (2017) Pan evaporation modeling using six different heuristic computing methods in different climates of China. *J Hydrol* 544:407–427. <https://doi.org/10.1016/j.jhydrol.2016.11.059>
- 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>
- Yin Z, Feng Q, Yang L, Deo RC, Wen X, Si J, Xiao S (2017) Future projection with an extreme-learning machine and support vector regression of reference evapotranspiration in a mountainous inland watershed in north-west China *Water* 9 <https://doi.org/10.3390/w9110880>
- Yin Z, Wen X, Feng Q, He Z, Zou S, Yang L (2016) Integrating genetic algorithm and support vector machine for modeling daily reference evapotranspiration in a semi-arid mountain area. *Hydrol Res* 48:1177–1191. <https://doi.org/10.2166/nh.2016.205>
- Zhang D et al (2018) Modeling and simulating of reservoir operation using the artificial neural network, support vector regression, deep learning algorithm. *J Hydrol* 565:720–736. <https://doi.org/10.1016/j.jhydrol.2018.08.050>
- Zong Woo G, Joong Hoon K, Loganathan GV (2001) A new heuristic optimization algorithm: harmony search. *SIMULATION* 76:60–68. <https://doi.org/10.1177/003754970107600201>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.