A Deep Neural Network Architecture to Model Reference Evapotranspiration Using a Single Input Meteorological Parameter

Sowmya Mangalath Ravindran¹ · Santosh Kumar Moorakkal Bhaskaran² · Sooraj Krishnan Nair Ambat³

Received: 29 April 2021 / Accepted: 22 September 2021 / Published online: 2 October 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract

Hydro-agrological research considers the reference evapotranspiration (ET_o), driven by meteorological variables, crucial for achieving precise irrigation in precision agriculture. ET_o modelling based on a single meteorological parameter would be beneficial in places where the collection of climatic parameters is challenging. The aim of this research is to develop a deep neural network (DNN) architecture that predicts daily ET_o with a single input parameter selected based on the feature importance (FI) score generated by the machine learning techniques, random forest (RF), and extreme gradient boosting (XGBoost). This study also investigated the potential of SHapley Additive exPlanations to interpret and validate the outcomes of the feature selection methods by assessing the contributions of each feature to the ET_o prediction. These methods recommended solar radiation as a significant parameter in the datasets of three California Irrigation Management System (CIMIS) weather stations located in distinct ET_o zones. Three ET_o models (DNN-Ret, XGB-Ret, and RF-Ret) were built using solar radiation as the sole input, and CIMIS ET_o as the output. The performance evaluation of the developed models proved that DNN-Ret outperformed XGB-Ret and RF-Ret regardless of the dataset, with coefficients of determination (R^2) ranging from 0.914 to 0.954 in the local scenario, with an average decrease of 8–9.5% in mean absolute error and root mean squared error, and an improvement of 2.6–2.9% in Nash–Sutcliffe efficiency and 1.7–2% increase in R^2 . The overall result analysis highlighted the efficiency of DNN-Ret in the single input parameter based ET_o modelling in diverse climatic zones.

Keywords Reference evapotranspiration · Deep neural network · Feature importance · Extreme gradient boosting · Random forest · SHAP

✉ Sowmya Mangalath Ravindran
sowmyamr@cusat.ac.in

¹ Department of Computer Applications, Cochin University of Science and Technology, Kochi, Kerala 682022, India

² Division of Information Technology, School of Engineering, Cochin University of Science and Technology, Kochi, Kerala 682022, India

³ Naval Physical and Oceanographic Laboratory, Defence R&D Organization, Kochi, Kerala 682021, India

1 Introduction

In this modern era of advanced technologies, the significance of proper water management is indispensable, with a skyrocketing demand for fulfilling the domestic, industrial, and agricultural needs of the booming population. As about 70% of all water drained from rivers and aquifers worldwide is diverted to crop production and other agricultural activities, agriculture captures the major share of water resources in most countries (Shang et al. 2019). Though various strategies have been proposed to ensure an adequate water management method in irrigation (Roy 2021), the role of agricultural irrigation scheduling is pivotal in a water-scarce country like India. The accurate estimation of the quantity of water for crop cultivation, and the forecasting of various agronomic and climatic conditions for its growth, aids the farmers as it enhances the effectiveness of irrigation scheduling and water resource management (Feng et al. 2017b; Mohammadi and Mehdizadeh 2020). The concept of crop water requirement, that is, the amount of water needed to meet the water loss through evapotranspiration (ET), serves as a potential tool for a reliable estimation of water consumption (Allen et al. 1998).

Reference evapotranspiration (ET_o), a variant of ET, is the percentage of ET of the reference crop, a grass with distinctive characteristics (Allen et al. 1998). The precise estimate of ET_o for a specific location and crop coefficient (K_c) for various growing seasons of a particular crop are multiplied together to obtain the actual ET (ET_c) of that crop. Various mathematical equations have been proposed by hydrologists (Xiang et al. 2020) to quantify ET_o and subsequently estimate ET_c. The calculation is based solely on meteorological variables. Of these empirical models, the Food and Agricultural Organization of the United States (FAO) has developed a combination type of empirical equation, Penman–Monteith (FPM-56) (Allen et al. 1998), which has gained total acceptance among the hydrological research community and is strongly recommended as a standard method to estimate ET_o. However, the need to accumulate a large number of meteorological variables for applying the FPM-56 equation paved the way for the implementation of other empirical methods, which use fewer input meteorological parameters (Xiang et al. 2020). Nevertheless, the inconsistent behavior and complex and data-intensive calculation of conventional empirical methods led to the application of artificial intelligence (AI) in ET_o modelling. Using the “black-box” aspect of AI, various soft computing approaches have been used as an ET_o prediction mechanism and have successfully addressed the non-linear interaction between inputs and ET_o (Chia et al. 2020b).

Minimal input parameter-based ET_o modelling research has acquired wide recognition in hydro-meteorological communities, especially in developing nations where weather stations and data collection methods are few (Adamala et al. 2015; Debnath et al. 2015; Roy 2021). Among these, certain recent studies have implemented the AI techniques such as artificial neural networks (ANN) and their derivatives (Ferreira et al. 2019; Reis et al. 2019; Ferreira and da Cunha 2020a; Sowmya et al. 2020; Bellido-Jiménez et al. 2021; Kaya et al. 2021), adaptive neuro fuzzy inference systems (ANFIS) (Petković et al. 2020; Üneş et al. 2020), support vector machine (SVM) and their hybrid variants (Chia et al. 2020a; Tikhmarine et al. 2020; Ahmadi et al. 2021), tree based soft computing methods (Fan et al. 2018; Wu et al. 2020), gene expression programming (GEP) (Kazemi et al. 2021; Muhammad et al. 2021), hybrid ML models (Mohammadi and Mehdizadeh 2020; Zhu et al. 2020; Kisi et al. 2021; Gong et al. 2021), and ensemble of ML models (Wu et al. 2021; Martín et al. 2021) for the ET_o modelling. Most of these studies have reported success by experimenting with different input parameters for modelling, such as only

temperature (Bellido-Jiménez et al. 2021), a combination of temperature and solar radiation (Chia et al. 2020a), temperature and relative humidity (Ferreira and da Cunha 2020a), temperature and wind speed (Nagappan et al. 2020) or a variety of other combinations to minimize meteorological data usage. Furthermore, when temperature and radiation data were included to the input set, these investigations demonstrated excellence and superiority compared to the equivalent traditional empirical techniques (Chia et al. 2020a; Üneş et al. 2020). In temperature-based studies, Rahimikhoob (2010), Wang et al. (2011), Feng et al. (2017b) and Sanikhani et al. (2019) used maximum and minimum air temperature and the extra-terrestrial radiation data for their modelling process, whereas Reis et al. (2019) and Adamala (2018) utilized only maximum and minimum temperature and proved their superiority to the Hargreaves-Samani empirical method. Recently, Bellido-Jiménez et al. (2021) employed several machine learning techniques to estimate ETo using intra-daily temperature-based variables. This study reported reasonably good performance compared to previous temperature-based studies. However, AI-based ETo modelling research based on a single input parameter is quite rare in the literature. Hence, one of the objectives of the present work is to create an ETo prediction model that uses a single and relevant input parameter while maintaining estimation accuracy comparable to FPM-56, consequently enriching the limited input parameter-based ETo research.

Feature engineering and feature selection are extensively used in machine learning algorithms to boost the prediction accuracy by eliminating irrelevant features that impair the model's generalization capability. ANFIS is often used in several domains to identify the optimal parameters for prediction tasks. In agriculture sector, Kuzman et al. (2021) analyzed the impact of various environmental parameters on the fertilizers for an optimal crop yield. Multiple studies have been done in the energy domain to determine the ideal parameters that enhance the performance of Kusum biodiesel (Petković et al. 2021), maximize the fatty acid methyl ester yield and exergy efficiency (Milić et al. 2021), maximize the forecasting efficiency of the water-jet assisted underwater laser cutting parameters (Nikolić et al. 2016), and so on. In the educational domain, Gavrilović et al. (2018) detected the significant factors that influence the mathematics lectures. Petković et al. (2017a) applied ANFIS to select the most influential inputs for precipitation concentration index estimation.

ET domain also applied the ANFIS to select the dominant parameters in the ETo assessment (Petković et al. 2015, 2020), where sunshine hours and global solar radiation were found to be the most relevant parameters. Most dominant features have been identified by Xing et al. (2016) using path analysis theory, and sunshine hours have been chosen as the best attribute for creating input combinations for ETo prediction. To ascertain optimal input combinations for support vector regression (SVR), Mohammadi and Mehdizadeh (2020) experimented relief, random forest (RF), principal component analysis (PCA), and Pearson's correlation as pre-processing data techniques. Based on the findings obtained from ANFIS, multilinear regression (MLR) and SVM, Üneş et al. (2020) selected temperature and solar radiation as the inputs for ETo modelling. Certain related researches have attempted genetic algorithm (Jovic et al. 2018), PCA (Nagappan et al. 2020), maximum information coefficient techniques (Chen et al. 2020a), subset regression analysis (Afzaal et al. 2020), and gamma test (Patil and Deka 2017) to analyze the effect of input parameters on ETo. These investigations were conducted in various climatic zones and most of them have chosen solar radiation as one of the most influential parameters on ETo. However, none of them has reported an accurate ETo model utilizing only the solar radiation data.

ETo is a meteorological parameter, and its specific and complicated relationship with other meteorological factors depends greatly on the diverse climatic conditions of different global areas. Thus, the identification of the highly significant meteorological parameter

that contributes the most to ETo estimation for each meteorological area under study will lead to the development of a cost-effective ETo prediction model, particularly in limited input parameter ETo modelling experiments. Ranking and filtering the features as per the calculated feature importance (FI) score is a type of pre-modelling task, and RF and extreme gradient boosting (XGBoost) belong to those classes of embedded techniques that integrate feature selection into their modelling process, and are referred to as RF-FI and XGB-FI, respectively, throughout the paper. Petković et al. (2017b) have modelled the multi target regression problem using RF coupled with Genie 3 ranking method. By identifying key features from the XGBoost FI scores and recursively eliminating the low ranked features, Shi et al. (2019) obtained an accurate predictive model from their study. Zhang et al. (2018) have adopted an XGBoost framework based on the top ranked features that were chosen using RF-FI methodology. In their ETo modelling study, Wang et al. (2019) assessed the importance of meteorological variables on ETo using the FI ranking capability of both RF and GEP. Although these approaches have been successfully implemented in several other domains, they have not yet been extensively applied in ETo datasets for optimal feature selection.

For validating the proposed FI techniques, this study utilized the SHAP (SHapley Additive exPlanations), a cutting-edge methodology for explaining the output of machine learning models by measuring the contribution of each input feature to the predicted output value (Lundberg and Lee 2017). The SHAP values have been widely used to measure the importance of features in various domains, such as gold price prediction (Jabeur et al. 2021), online review (Meng et al. 2021), COVID-19 diagnosis (Zoabi et al. 2021), and traumatic brain injury prognostication (Farzaneh et al. 2021). Effrosynidis and Arampatzis (2021) examined the consistency and effectiveness of various feature selection techniques in environmental datasets and demonstrated SHAP's excellence. In ET domain, Başağaoğlu et al. (2021) applied the SHAP value analysis to identify the contributions of hydro-climatic variables on different ET parameters. This study identified that shortwave solar radiation, air temperature, and relative humidity are the optimal features for the daily ETo prediction. The encouraging results of these studies motivated the authors to apply FI analysis approaches in the current study to select the most relevant meteorological parameters for ETo modelling. This will eventually help to construct unique ETo estimate models according to diverse geographic locations and climatic circumstances, as well as to address the seasonal fluctuations of ETo.

One of the key challenges of the machine learning technique is to enhance prediction accuracy and produce more reliable results-producing models. This can be achieved through providing more data to train the model, introducing large architectures, and providing more computing resources. The neural network research on ETo modelling was commenced by Kumar et al. (2002) using a shallow feed-forward multi-layer perceptron (FFMLP), and this has currently reached in the implementation of deep neural network architectures (de Oliveira e Lucas et al. 2020; Roy 2021). Most ET-based hydrological studies rely primarily on FFMLP for modelling purposes with back-propagation as a learning algorithm (Antonopoulos and Antonopoulos 2017; Abrishami et al. 2019). Later FFMLP was refined to the radial basis function network using Gaussian as a transfer function (Trajkovic 2005), generalized regression neural network (Feng et al. 2017b), extreme learning machine (ELM), and its hybridised variants (Reis et al. 2019; Zhu et al. 2020). Recently, deep learning models have gained a great deal of attention among other AI models attributed to their features such as increased accuracy, robustness, efficiency, decreased computational costs, and overall modelling effectiveness (Han et al. 2019). In the domain of hydrology, especially in ETo modelling, the exploration of deep learning architectures

has been sporadic. Saggi and Jain (2019) introduced the deep neural network (DNN) concept in ETo modelling by depicting the relationship between the entire set of seven input parameters and ETo. This research showed the potential of the deep learning architecture to model ETo by comparing its performance with the gradient boosting machine, the generalized linear model, and the RF. Özgür and Yamaç (2020) tested the power of DNN in ETo estimation by using SeLU activation function. A reduced feature one-dimensional convolution neural network (CNN) model for ETo prediction was developed by Nagappan et al. (2020) and succeeded with promising results. Sowmya et al. (2020) developed various DNN variants for ETo prediction using different feature combinations, and analyzed their predictive performances. Both local and regional based ETo forecasting studies (Ferreira and da Cunha 2020a, b), and ETo time-series forecasting study (de Oliveira e Lucas et al. 2020) demonstrated the strength of CNN and its variants. Research conducted on the temporal convolutional neural network (TCN) (Chen et al. 2020a, b) and recurrent neural network (RNN) architectures (Afzaal et al. 2020) highlighted the advantages of deep learning models over traditional machine learning and empirical methods. Because the present study uses only one input parameter for modelling, the machine learning methods for modelling require sophisticated hyper-parameter optimization. The feature learning capabilities of deep learning algorithms and the encouraging findings of previously mentioned deep learning-based ETo modelling studies have motivated the authors to select the deep learning architectures for this complex modelling task.

The tree-based ETo estimation models, such as RF (Feng et al. 2017a; Wang et al. 2019; Karimi et al. 2020), gradient boosting decision tree (Ponraj and Vigneswaran 2020), XGBoost (Fan et al. 2018; Wu and Fan 2019), and light gradient boosting machine (Fan et al. 2019) were observed to provide more stable results, faster predictions, management of large datasets, and also the capability to prevent over-fitting, in comparison to other soft computing techniques. Recently, Wu et al. (2021) unveiled stacking and blending ensemble ETo models and demonstrated their supremacy over basic machine learning and empirical models in prediction precision, stability, portability, and computing cost under complete and minimal input scenarios. While the strength of the ensemble and boosting techniques have given rise to these models and highlighted their state-of-the-art impacts in various studies, their applications in the ETo modelling are very abundant in the literature. Therefore, the present study selected the ensemble techniques, XGBoost and RF, to create baseline models for evaluating the proposed DNN architecture for ETo modelling.

The remarkable features of deep learning architectures, the relevance of feature selection approaches in machine learning algorithms, the significance of the most prominent single input parameter in ETo modelling, and the state-of-the-art outcomes of bagging and boosting techniques have motivated the authors to execute this research. The novelty of the present study is the proposal of a methodology to develop a DNN architecture for ETo modelling, which utilizes only one meteorological input parameter. Another contribution is the application of FI techniques such as RF-FI, XGB-FI and SHAP to identify the most influential feature for the ETo calculation in various climatic zones. The authors implemented the proposed methodology by taking the daily meteorological data from the dataset of three California Irrigation Management System (CIMIS) weather stations located in distinct ETo zones in California as a case, resulting in the development of the solar radiation based ETo prediction model, which is another contribution made by the current study. To execute the methodology, the authors set the objectives of the case study as: (1) Applying the feature selection techniques RF-FI, XGB-FI and SHAP to the datasets to identify a single input parameter that significantly contributes to the ETo prediction; (2) Building a DNN architecture, tree-based ensemble RF and XGBoost models to predict daily ETo by

using the selected parameter as input and FPM-56 ETo as output; and (3) Evaluating the performance of the proposed DNN model by considering the RF and XGBoost models as a baseline.

2 Materials and Methods

2.1 Study Site Description

CIMIS is a free data resource administered by the Department of Water Resources that functions by gathering input from over 145 weather stations of California (Kisi 2011). The proposed case study, taking into consideration the varied landmass and atmosphere of California, selected the following three weather stations of CIMIS: Oakville, Pomona and Alturas. These comprise distinct ETo zones. Oakville lies in the ETo zone of the inland San Francisco Bay area and has 1254.76 mm of ETo as the monthly average. A cool and temperate climate also characterizes it because of the coastal influence with an average annual temperature of about 8.4 °C and an average yearly rainfall of 794 mm. The mid-central valley ETo zone, Pomona, has peak summer sunshine and wind, and experiences a warm climate with a monthly average ETo of 1447.8 mm. The annual mean temperature and precipitation of Pomona are 17.9 °C and 421 mm, respectively. The north-eastern plain, Alturas, at 1331 m above sea level, has a mild and temperate climate with an estimated annual temperature of 8.2 °C. There is comparatively more precipitation in winter than in summer, with an average value of 322 mm per year and a mean monthly ETo average of 1102.36 mm. More geographical details of the study sites are described in Table 1.

2.2 Dataset Used

The data collection at CIMIS weather stations is carried out on a per-minute basis. The measured and estimated parameters are then recorded on an hourly and daily basis in the CIMIS database. The datasets used in this study include the daily data from the three automated weather stations of CIMIS for 32 years, from March 1989 to October 2020. CIMIS weather stations have used various measurement instruments for data collection, such as solar radiation pyranometers, air temperature thermistors, relative humidity sensors, and wind speed anemometers (Kisi 2011), and these gauged data are expected to meet input data criteria at different points of the current analysis. In addition to the measured parameters, the ETo value, which is the output parameter of this study, is the estimated parameter and is computed for each CIMIS weather station using the CIMIS Penman equation (Kisi 2011), and it is given in Eq. (1). The FPM-56 equation (Allen et al. 1998), as presented in Eq. (2) is revised by integrating a wind function in addition to other climatic parameters to form the CIMIS Penman equation:

Table 1 Study site description

Station	Latitude	Longitude	Altitude
Oakville	38.26° N	122.25° W	57.912 m
Pomona	34.4° N	117.49° W	222.504 m
Alturas	41.26° N	120.29° W	1342.644 m

$$ETo = \left(\frac{\Delta}{\Delta + \gamma} \right) R_n + \left(1 - \frac{\Delta}{\Delta + \gamma} \right) (e_s - e_a) f_u \quad (1)$$

$$ETo = \frac{0.408\Delta(R_n - G) + \gamma \left(\frac{900}{T_m + 273} \right) u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (2)$$

where Δ is the slope of vapour pressure curve (kPa/°C), γ is the psychometric constant (kPa/°C), R_n is the net solar radiation (MJ/m²/day), e_s is mean saturation vapour pressure (kPa), e_a is actual vapour pressure (kPa), f_u is the wind function (m/s), G is the soil heat flux (MJ/m²/day), T_m is the daily mean temperature (°C), and u_2 is the wind speed at 2 m height (m/s).

The hourly measured weather data such as maximum, minimum, and average air temperature ($T_{\max}$, $T_{\min}$, T_{avg}), solar radiation (R_s), maximum and minimum relative humidity ($RH_{\max}$, $RH_{\min}$), and wind speed (u_2) at two meters, and other values estimated from these measured parameters are given as input parameters for the CIMIS Penman equation to estimate hourly ETo and summed it up over 24 h to estimate daily ETo. The descriptive statistics of the datasets are tabulated in Table 2.

2.3 Dataset Pre-processing

2.3.1 Removal of Missing Values and Outliers

Missing data causes bias in the forecasting model, and these models are often very vulnerable to outliers too. Hence, it should be managed to ensure the proper data quality and achieve better prediction accuracy (Lin and Tsai 2020). In the aforementioned datasets, a count of 456, 716 and 320 records lack feature values from Oakville, Pomona and Alturas stations, respectively. Certain feature values indicate anomalies in all datasets, mainly in ETo, $T_{\max}$, $T_{\min}$, T_{avg} , R_s and u_2 data. As a pre-processing step, the datasets were prepared by removing the missing values and inconsistent outliers. The correlation coefficients depicting the difference in the pairwise correlation between the features and ETo before and after the pre-processing phase are listed in Table 3. This comparison table shows a slight increase in correlation between the ETo and $T_{\max}$, T_{avg} , R_s and $RH_{\min}$ in Oakville dataset. The Pomona dataset also achieved a high correlation between ETo and all other features except R_s . However, there was no discrepancy in correlation in Alturas dataset because it contained negligible missing and outlier values.

2.3.2 Normalization

The learning process of the neural network simulation is worse influenced by the disparity in the dimensions of the input features of the dataset. Normalization maps the input and output variables to a similar magnitude, and modifies the input data within the domain of the transfer function of the neural network, resulting in much faster learning and convergence (Antonopoulos and Antonopoulos 2017; Abrishami et al. 2019). In this analysis, the min–max scaling procedure is used to normalize the attributes between 0 and 1 using Eq. (3) before loading them into the model training:

Table 2 Descriptive statistics of the features of the datasets

Dataset	Feature	Max ^a	Min ^b	Mean	SD ^c	Correlation with ETo
Oakville	T _{max} ^d	43.9	1.1	23.0	7.0	0.83
	T _{min} ^e	24.4	−8.3	7.0	4.1	0.51
	T _{avg} ^f	31.7	−2.0	15.0	5.0	0.79
	R _S ^g	381	5.0	200	95	0.97
	u ₂ ^h	4.0	0	1.25	0.88	0.17
	RH _{max} ⁱ	100.0	20.4	93.3	9.1	−0.25
	RH _{min} ^j	100.0	0	45.2	19.5	−0.6
	ETo ^k	9.1	0	3.4	2.0	1
Pomona	T _{max}	42.8	8.3	24.9	6.2	0.79
	T _{min}	25.0	−4.4	10.4	4.7	0.56
	T _{avg}	32.0	2.5	17.6	5.0	0.75
	R _S	363	12	207	83	0.96
	u ₂	4.0	0	1	0.69	0.09
	RH _{max}	100.0	19.0	89.3	11.3	−0.01
	RH _{min}	100.0	0	42.0	17.3	−0.33
	ETo	7.6	0	3.4	1.8	1
Alturas	T _{max}	40.0	−13.3	17.1	9.8	0.87
	T _{min}	17.8	−32.2	−0.64	6.2	0.68
	T _{avg}	27.3	−22.0	8.2	7.5	0.85
	R _S	410	9	195	103	0.94
	u ₂	9.0	0	1.7	1.4	−0.1
	RH _{max}	100.0	40.0	90.5	9.2	−0.28
	RH _{min}	100.0	0	38.9	20.2	−0.68
	ETo	8.4	0	3.2	2.2	1

^aMaximum^bMinimum^cStandard deviation^dMaximum temperature (°C)^eMinimum temperature (°C)^fAverage temperature (°C)^gSolar radiation (Watt/m²)^hWind speed (m/s)ⁱMaximum relative humidity (%)^jMinimum relative humidity (%)^kReference evapotranspiration (mm/day)

$$X_{norm} = \frac{(X_n - X_{min})}{(X_{max} - X_{min})} \quad (3)$$

where X_{norm} is the normalized value, X_n is the actual value, X_{max} is the maximum value, and X_{min} is the minimum value of the attributes.

Table 3 Correlation coefficients between the predictor features and ETo before and after data pre-processing

Dataset	Data pre-processing	Correlation coefficients between ETo and						
		$T_{\max}$	$T_{\min}$	T_{avg}	R_s	u_2	$RH_{\max}$	$RH_{\min}$
Oakville	✗	0.82	0.51	0.79	0.96	0.17	-0.25	-0.61
	✓	0.83	0.51	0.80	0.97	0.17	-0.25	-0.6
Pomona	✗	0.78	0.55	0.74	0.96	0.08	-0.018	-0.34
	✓	0.79	0.56	0.75	0.96	0.085	-0.012	-0.33
Alturas	✗	0.87	0.68	0.85	-0.09	0.94	-0.28	-0.68
	✓	0.87	0.68	0.85	0.94	-0.1	-0.28	-0.68

2.4 Feature Importance Measuring Techniques for Feature Selection

As the AI models usually portray the relationship between input and output variables, the role of dimensionality reduction as a pre-processing technique, particularly in limited data cases, is high and results in a low-cost solution for AI modelling tasks (Chen et al. 2020a). FI is one of the input selection methods, which includes several techniques to assess the value of each attribute in model building, and consequently makes the right predictions. Based on an estimated score, the relevance of the feature in the predictive modelling process is ranked, and an input selection can be rendered by preserving the high score features and removing the low score features from the dataset (Chen et al. 2019). This score aims to provide an insight into the data and model under consideration, reduce the dimensionality, and decrease the memory and time required for the modelling task (Petković et al. 2017b). Tree-based ensemble learning algorithms such as RF and XGBoost have a built-in ability to rank features according to their relevance because its construction itself is focused on the reduction in the criterion used to pick split points such as variance in regression and Gini in classification (Joharestani et al. 2019; Shi et al. 2019).

2.4.1 Random Forest

RF is a forest of many classification and regression trees (CARTs) modelled by Breiman (2001) to simulate the input–output relationship in both classification and regression problems intelligently with strong tolerance to outliers and anomalies. The robust interpretability of the RF bagging technique is because of its ability to infer the significance of each variable in the tree construction decision (Zhang et al. 2018). The RF measures the noise or impurity of the features; that is, it checks how much each feature degrades the performance of the model, or it tests how much each attribute reduces impurity. The more the feature lowers the impurity, the more important it is. As a result, the final importance of the feature is obtained by taking the average decrease in impurity from each attribute across trees (Petković et al. 2017b).

Component trees in the RF train independently on a different dataset that is sampled with replacement from the original data. The out-of-bag error (OOB) is calculated in all data points by using the value of the attribute from another OOB point that was not used during training (Wang et al. 2019). Based on the fact that the more frequent splitting

attribute lowers the training error, feature ranking is conducted according to the frequency of splitting of branches in the trees. Hence, the features selected at the top of the trees are usually more important than the features selected at the end nodes of the trees (Zhang et al. 2018; Wang et al. 2019).

2.4.2 XGBoost

Boosting is an ensemble strategy that integrates a group of weak learners to maximize forecast accuracy. XGBoost is a member of the boosting algorithm family that uses as its core the gradient-boosting paradigm, which creates a tree ensemble by collecting information from previously generated trees (Chen and Guestrin 2016). Generally, the weak models are CARTs, which are grown one by one by reducing the error in predictive modelling tasks and disbursing weights according to the model result. The weights refer to the relative contribution of a feature to the model and signify the importance of an attribute in the modelling process. This employs an additive training procedure to build a boosted tree learner from the weak learners by minimizing the variance in the model output and trying to fit into the tree ensemble yielding an optimal loss function (Wu and Fan 2019). Hence, once the boosted trees are built, the importance score metric can be directly retrieved for each attribute (Shi et al. 2019). The potential to tackle overfitting problems and parallelize computation makes XGBoost a state-of-the-art machine learning technique (Fan et al. 2018).

2.4.3 SHapley Additive exPlanations (SHAP)

SHAP is a game-theoretic technique introduced by Lundberg and Lee (2017) to explain the prediction of any machine learning model, either globally or locally. This is a unified measure of feature importance based on Shapley values, computed by quantifying each feature's contribution to the prediction and aggregating it across the whole population. The collective Shapley values offer SHAP's global interpretability. The local interpretability of SHAP is driven by the Shapley values of each instance of the dataset, which assists in detecting and assessing the impacts of features on the output. Consequently, these values assist the SHAP in distinguishing between features that push the model's prediction higher and those that drive the prediction lower.

Lundberg et al. (2018) proposed TreeSHAP, a fast and model-specific variant of SHAP designed specifically for tree-based ensemble machine learning models such as decision trees, RF, and gradient boosted trees like XGBoost. TreeSHAP incorporates interaction effects of features into the Shapley values using the conditional expectation function and explains the model through sophisticated visualization of individual feature characterizations, which excels conventional feature attribution approaches. The visualization tools such as feature importance plots, SHAP dependence plots, and force plots to depict individual SHAP value provide global and local interpretability to tree ensemble models.

2.5 Deep Neural Network

DNN is an improved variant of ANN; however, unlike ANN, the nodes in DNN are strongly interdependent and imply long-term dependence through weight sharing. As a recursive feed-forward neural network, it uses more than one hidden layer in addition to one input and one output layer to achieve reduced overfitting and enhanced generalizability (Vieira et al. 2020). The number of neurons in the input and output layer of DNN aligns

with the number of predictor and response variables. The numerous hidden layers responsible for the entire computation enable the DNN to receive and process huge volume of input data. There should be an appropriate number of neurons in the hidden layers, enabling the DNN to learn more sophisticated data characteristics (Lecun et al. 2015).

The DNN learning mechanism involves the iterative execution of feed-forward and error back-propagation cycles until the optimal degree of precision has been attained (Saggi and Jain 2019; Sowmya et al. 2020). DNN's generalization efficiency relies on the activation function used. The activation function triggered in each neuron includes hyperbolic tan function, rectified linear unit function (ReLU), logistic sigmoid function, and so on (Zhu et al. 2018; Datta 2020). During the training process, the use of higher number of hidden layers may raise gradient vanishing, which can be eliminated by using ReLU as the activation function (Datta 2020). It can be represented as Eq. (4):

$$ReLU(k) = \begin{cases} k, & \text{if } k > 0 \\ 0, & \text{if } k \leq 0 \end{cases} \quad (4)$$

The training period of the neural network requires iterative weight adjustment, which necessitates certain random initialization of weights at the beginning, and this mode of initialization is another cause of the gradient boosting and vanishing phenomena. Glorot and Bengio (2010) and He et al. (2015) have suggested several ways to generate values from some statistical distributions to initialize weights. All of these initialization methods aim to maintain the gradient variance consistent across the layers and adjust it to the activation function employed. As the selection of weight values for one layer depends on the previous layer's neuron characteristics, it allows achieving a smoother and more efficient gradient descent without any saturation. Among the different kernel initializers, *he* initialization derived from Normal and Uniform distribution (*he_normal* and *he_uniform*) (He et al. 2015) is currently leading in deep learning research. Moreover, Datta (2020) demonstrated the combined effect of using ReLU with *he* initialization and surveyed several studies that implemented it.

The back-propagation error mechanism reinforces the model through a derivative method of error calculation and weight and bias updation. The weights of each neuron are adjusted with a loss function, and the loss function has to be minimized during the training process to increase model performance (Sowmya et al. 2020; Vieira et al. 2020). For regression, the widely used loss function is mean squared error (MSE) and it is computed using Eq. (5):

$$MSE = \frac{1}{N} \sum_{i=1}^N (X_{pred} - X_{org})^2 \quad (5)$$

where N is the total data record count, X_{pred} and X_{org} are the predicted and actual output values, respectively.

The Stochastic Gradient Descent Optimizer, Adam, is mainly used for weight optimization in all deep learning architectures (Zhu et al. 2018). Thus, the intensity of the signal or output of each neuron moving to the next layer depends on the weight, bias, and activation function and is presented in Eq. (6):

$$Y = \varnothing \left(\sum (w_i * x_i + b) \right) \quad (6)$$

where $\varnothing$ is the activation function, w_i is the neuron weight, x_i is the neuron input, and b is the neuron bias.

Fig. 1 Operational workflow diagram of the proposed methodology, which includes six phases from data collection to ETo prediction model evaluation and deployment

Table 4 FI scores of the predictor features of the datasets (The boldface entries indicate the highest RF-FI and XGB-FI scores obtained by R_S compared to other features in Oakville, Pomona, and Alturas datasets)

Features	Oakville		Pomona		Alturas	
	RF-FI	XGB-FI	RF-FI	XGB-FI	RF-FI	XGB-FI
$T_{\max}$	0.01628	0.01489	0.03503	0.04369	0.01583	0.01151
$T_{\min}$	0.00405	0.00556	0.00323	0.00442	0.00475	0.00641
T_{avg}	0.01106	0.01732	0.00807	0.01304	0.05919	0.10382
R_S	0.94037	0.92592	0.93037	0.91111	0.88669	0.83213
u_2	0.00381	0.01348	0.00328	0.01153	0.00925	0.02226
$RH_{\max}$	0.00909	0.00709	0.00658	0.00554	0.00896	0.00778
$RH_{\min}$	0.01534	0.01574	0.01342	0.01065	0.01534	0.01609

In addition to the factors mentioned above, such as hidden layer counts, neuron counts in hidden layers, activation functions, optimization, certain other factors are also required to describe the DNN architecture, such as momentum, dropout, learning rate, regularisation, and epoch size (Sowmya et al. 2020; Vieira et al. 2020).

The workflow of the proposed methodology of the present study is depicted in Fig. 1.

2.6 Implementation Specifics of the Proposed Methods

2.6.1 Input Data Selection

The present study employed three FI measuring techniques, i.e., RF-FI, XGB-FI and SHAP, to identify the highest FI score feature from the seven predictor parameters of the Oakville, Pomona and Alturas datasets. Through computing a score by quantifying the noise that adversely affects its predictive modelling performance, RF-FI and XGB-FI methods ranked the features. The scores are tabulated in Table 4 for each dataset. Table 5 lists the ranks of the predictor features as per the average FI score of RF-FI and XGB-FI applied to each dataset.

Table 4 shows a similarity between the RF-FI and XGB-FI scores with only a negligible variation in all datasets. According to the average score analysis described in Table 5,

Table 5 FI Rank list of the predictor features (The boldface entries indicate the average RF-FI and XGB-FI scores obtained by the top-ranked feature R_S in Oakville, Pomona, and Alturas datasets)

Features	Oakville		Pomona		Alturas	
	Score	Rank	Score	Rank	Score	Rank
$T_{\max}$	0.016	2	0.039	2	0.014	5
$T_{\min}$	0.005	7	0.004	7	0.006	7
T_{avg}	0.014	4	0.011	4	0.082	2
R_S	0.933	1	0.920	1	0.859	1
u_2	0.009	5	0.007	5	0.016	3
$RH_{\max}$	0.008	6	0.006	6	0.008	6
$RH_{\min}$	0.015	3	0.012	3	0.015	4

R_S ranked first in the Oakville, Pomona and Alturas datasets, with scores of 0.93, 0.92 and 0.86, respectively. The dataset-wise average score analysis reported an improvement of 98.8%, 98.6%, and 97.2% in FI score of R_S relative to all other features in Oakville, Pomona and Alturas, respectively. The average FI scores of the second-ranked feature in the Oakville ($T_{\max}$), Pomona ($T_{\max}$) and Alturas (T_{avg}) datasets were 0.016, 0.039 and 0.082, respectively, showing a decline of 98.3%, 95.7%, and 90.5% in comparison to the FI score of R_S . These factors imply the strength of the relationship between R_S and ETo.

The summary plot generated from TreeSHAP, as shown in Fig. 2, was utilized to validate the FI rank list that resulted from RF-FI and XGB-FI. It displays the order of seven variables based on their importance in influencing the ETo in the three weather station datasets, Oakville, Pomona and Alturas. Features with large absolute Shapley values are significant, and the global importance measure of a dataset is calculated by aggregating these values per feature throughout the data and plotting them in decreasing order of importance. The plots in Fig. 2 reported that R_S is the most important feature in the ETo prediction model design, and therefore, supported the findings of both RF-FI and XGB-FI. Moreover, $T_{\min}$ and $RH_{\max}$ were found as the least significant features. There are minor discrepancies relating the positions of $T_{\max}$, T_{avg} and $RH_{\min}$ in the rank list compared to the plot's interpretation. However, the variations in the Shapley values of these variables are so negligible compared to R_S that the inclusion of these parameters would not enhance prediction accuracy and can be ignored. As a result, the interpretation of the summary plot reinforces the fact that the single predictor variable, R_S , is ideal for modelling the ETo in this case study. A more in-depth explanation of feature analysis with SHAP is given in Sect. 3.1.

2.6.2 Predictive Model Development

In this study, three models were designed for ETo prediction: bagging-based RF-Ret, boosting-based XGB-Ret and deep learning-based DNN-Ret using only one input parameter, R_S , which was found to have the most influencing effect on the output parameter, ETo. The training process involved in the model building consumed 80% example records of the weather station datasets. The remaining 20% records were reserved for testing purposes, and the splitting count of the data records for the training and testing processes is shown in Fig. 3. For implementing these models, Google cloud's virtual machine platform with a graphical processing unit was explored, along with packages like Scikit-learn ver.0.22.2,

Fig. 2 SHAP summary plots showing the relative importance of each feature in the ETo prediction at: **a** Oakville; **b** Pomona; and **c** Alturas datasets. FI is computed by averaging the absolute Shapley values per feature across the dataset

Fig. 3 Training/testing records split

XGBoost ver.0.90, Tensorflow ver.2.3.0, Keras ver. 2.4.3 and Python programming language.

Hyper parameters of the RF-Ret and XGB-Ret models were chosen using a grid search technique. The $n_estimator$ parameter indicates the tree count in the RF-Ret forest (Ferreira and da Cunha 2020a), and in the case of XGB-Ret, it is the count of boosted trees to fit (Ferreira and da Cunha 2020b). This parameter has the potential to manage the computational complexity and generalizability of these models. Both models used the search space (30, 50, 70, 100, 200, 300) for the $n_estimator$. The max_depth is the tree depth control parameter of the model, and thus, handles the overfitting problem effectively throughout the modelling process. The max_depth values experimented in this study for RF-Ret were 3, 5, 7 and 9, while those for XGB-Ret were 3, 5, 7, 9, 15 and 20. Another relevant parameter for designing the XGB-Ret is the $learning_rate$, which governs the weighting of new trees applied to the model and its search space included (0.05, 0.1, 0.15, 0.2). The optimal hyper parameters of these models resulted from the grid search were tabulated in Table 6.

The DNN-Ret architecture for ETo modelling at different stations and its hyper parameters were identified by trial-and-error. As this study focuses on a single input parameter-based regression model, the input and output layers of the DNN-Ret requires only one neuron. The hidden layer count and the number of neurons in each layer are characterized by the consistency and number of training data records taken into account and vary with regard to the dataset. Here, the DNN-Ret models have examined with one to four hidden layers with 10, 15, 20, 25, 30, 40 and 60 neurons per hidden layer. Four hidden layers of (40, 60, 60, 40) neurons combination performed better in Oakville and Alturas data scenarios, while (10, 15, 15, 10) worked better in Pomona dataset. The DNN-Ret model architecture (40-60-60-40) deployed in the current study is shown in Fig. 4.

The initial weights of DNN-Ret were chosen from a uniform distribution ranging from $-\sqrt{6/f1}$ to $+\sqrt{6/f1}$, where $f1$ is the fan_{in} (the number of input units in the weight tensor) as per the $he_uniform$ kernel initialiser (He et al. 2015). $ReLU$ activation in hidden layers and $linear$ activation in the output layer were also used during feed-forward operation of model training. The weights and biases were modified and refined during the back propagation process with *Adam* optimizer and *MSE* loss function, resulting in a globally minimum cost value. With a batch size of 16 and a default learning rate, the entire model training was carried out for a maximum of 200 epochs using the early stopping criterion.

2.7 Criteria for Model Performance Assessment

As this case study belongs to the regression category, statistical performance metrics such as mean absolute error (MAE), root mean square error (RMSE), Nash–Sutcliffe efficiency coefficient (NSE), and coefficient of determination (R^2) between actual and predicted ETo were

Table 6 Optimal parameters from grid search for RF-Ret and XGB-Ret models

Dataset	RF-Ret		XGB-Ret		
	max_depth	n_estimator	max_depth	n_estimator	learning_rate
Oakville	5	300	3	100	0.1
Pomona	5	70	3	100	0.1
Alturas	5	100	3	100	0.1

Fig. 4 DNN-Ret structure (40-60-60-40) having input layer with one neuron, four hidden layers with 40, 60, 60, and 40 neurons, respectively, and output layer with one neuron

used to analyse the performance of the proposed DNN model and other baseline machine learning models of this study. RMSE, the goodness of the fit metric, is the standard deviation of the discrepancy between the predicted and the actual values (Yaseen et al. 2018; Wu et al. 2021). As analogous to RMSE, the goodness of fit measure that does not govern the direction of the value is considered as MAE (Yaseen et al. 2018; Wu et al. 2021). In relation to the calculated data variance, NSE shows how much the map of observed and simulated data matches the 1:1 line by evaluating the relative magnitude of the residual variance (Saggi and Jain 2019; Wu et al. 2021). R^2 is an assessment parameter which shows the degree of the linear relationship between the expected and actual values of the output variable (Yaseen et al. 2018; Wu et al. 2021). All these criteria are presented using Eqs. (7) to (10):

$$MAE = \frac{1}{n} \sum_{i=1}^n |ET_i^p - ET_i^a| \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (ET_i^p - ET_i^a)^2} \quad (8)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (ET_i^p - ET_i^a)^2}{\sum_{i=1}^n (\overline{ET_i^a} - ET_i^a)^2} \quad (9)$$

$$R^2 = \left(\frac{\sum_{i=1}^n (ET_i^p - \overline{ET_i^p})(ET_i^a - \overline{ET_i^a})}{\sqrt{\left(\sum_{i=1}^n (ET_i^p - \overline{ET_i^p})^2\right)\left(\sum_{i=1}^n (ET_i^a - \overline{ET_i^a})^2\right)}} \right)^2 \quad (10)$$

where n is the number of data records, ET_i^p is the predicted ETo values, ET_i^a is the actual FPM-56 ETo values, $\overline{ET_i^p}$ is the mean of predicted ETo values, $\overline{ET_i^a}$ is the mean of actual ETo values.

3 Results and Discussion

3.1 Feature Analysis

Figure 5 presents the distribution of the impacts of each feature on ETo prediction by plotting the Shapley value of that particular feature for every sample. The points on this beeswarm plot represent Shapley values of the features related to the daily records of meteorological data, providing insight into the importance and association of each of the seven features on the ETo prediction. The red and blue hues in the figure represent the higher and lower values of ETo prediction, respectively. All three plots in Fig. 5 clearly indicate that R_s is the most significant feature in all datasets. Higher R_s values, shown in red color, lead to a higher ETo prediction value, indicating a positive relationship between R_s and ETo. However, the larger $RH_{\min}$ and $RH_{\max}$ values result in a low ETo prediction value, which is depicted in blue in the plot, and of these, $RH_{\min}$ ranks top in FI measure. Although these plots demonstrate a positive association between ETo and the temperature variables, the correlation may cause some redundancy in these features and push these parameters down in the FI rank list.

SHAP force plots show the marginal influence of features on the predicted model input by highlighting the features that contribute most to push the output from the base value to the expected output value. This is illustrated using two colors: red for higher value prediction and blue for lower value prediction. For example, in Fig. 6, the lower value of ETo prediction, 1.42 mm, is attributed to R_s with 42 Watt/m², $RH_{\min}$ with 87.1%, u_2 with 4 m/s and T_{avg} with 16.65 °C. However, the values of 350 Watt/m² for R_s , 20.2% for $RH_{\min}$, and 100% for $RH_{\max}$ raise the ETo forecast to 6.41 mm. This confirms the idea that greater R_s and u_2 and lower $RH_{\max}$ and $RH_{\min}$ drive the ETo to a higher value, with R_s contributing more.

To further investigate the association between features and ETo prediction output, dependence plots of SHAP can be used. This helps in examining the significant impact of individual predictor variables and their interactions. The dependence plot in Fig. 7a demonstrates the impact of R_s on ETo as the value of T_{avg} rises from 8 to 22 °C. The ETo prediction value increases with increasing R_s value, particularly from 200 Watt/m², which is indicated as the red zone. The reverse effect can also be seen in Fig. 7b, where $RH_{\min}$ values more than 75% and R_s values less than 200 Watt/m² generate a lower value of ETo prediction, which is indicated as the blue zone. In addition, when compared to other features in the dataset, this finding shows a significant positive association between R_s and ETo prediction.

3.2 Model Performance Comparison

The ETo prediction models RF-Ret, XGB-Ret and DNN-Ret developed in this study were evaluated using the MAE, RMSE, NSE and R^2 metrics. Results are listed in Tables 7 and 8. In model predictions, MAE and RMSE denote the magnitude of error, and NSE and R^2 represent the efficiency of the model prediction. Hence, the lower value of MAE and

Fig. 5 SHAP beeswarm plot displaying the combined view of FI and feature impacts on the ETo prediction in the datasets: **a** Oakville, **b** Pomona and **c** Alturas. Each point on the x-axis corresponds to a Shapley value for a feature of daily meteorological data record instance. Seven features are positioned along y-axis by their mean absolute Shapley values. The color spectrum from blue to red represents the value of the feature ranging from low to high

RMSE towards 0 and the higher values of NSE and R^2 towards 1 determine the model's efficiency.

This section analyses the results of the proposed study in the local and general context. In the local scenario, the model performance assessment is carried out study station-wise, whereas performance assessments are averaged across the study stations and analysed in the general scenario.

3.2.1 Evaluation of the Developed Models in the Local Scenario

The performance metrics of the ETo prediction models developed for the Oakville, Pomona, and Alturas stations during the test period are listed in Table 7. The table entries show that the DNN-Ret significantly outperformed both the RF-Ret and XGB-Ret models at all weather stations in this study.

Fig. 6 SHAP force plots for the interpretation of individual ETo predictions in the Oakville dataset. The base value is the mean prediction ETo value calculated from the training data. The amount of positive and negative contribution of features to ETo prediction is shown by red and blue arrows: **a** depicts low ETo prediction of 1.42 mm with low R_S value; **b** depicts high ETo prediction of 6.41 mm with high R_S value

Fig. 7 SHAP dependence plots depicting the association and interaction between predictor variables, with feature values on x-axis and Shapley values on y-axis: **a** R_S dependence plot with T_{avg} interaction; **b** RH_{min} dependence plot with R_S interaction

At Oakville station, the MAE values of the developed models during testing decreased from 0.356 to 0.335 mm/day through 0.348 mm/day, gaining a percentage decrease of 5.9 for the DNN-Ret with respect to RF-Ret model and 3.7 with respect to XGB-Ret model.

Table 7 The performance metrics of ETo prediction models during the testing period (The boldface entries indicate the best performance metrics reported by DNN-Ret compared to XGB-Ret and RF-Ret during the testing period)

Dataset	Model	MAE(mm/day)	RMSE(mm/day)	NSE	R ²
Oakville	DNN-Ret	0.335	0.472	0.95	0.954
	XGB-Ret	0.348	0.496	0.937	0.946
	RF-Ret	0.356	0.505	0.934	0.944
Pomona	DNN-Ret	0.325	0.464	0.934	0.938
	XGB-Ret	0.343	0.484	0.918	0.926
	RF-Ret	0.349	0.491	0.917	0.924
Alturas	DNN-Ret	0.478	0.646	0.911	0.914
	XGB-Ret	0.545	0.741	0.87	0.886
	RF-Ret	0.551	0.75	0.868	0.883

Table 8 The mean performance metrics of ETo prediction models during the training and testing periods (The boldface entries indicate the best mean performance metrics reported by DNN-Ret at CIMIS weather stations compared to XGB-Ret and RF-Ret during the training and testing periods)

Model	Period	MAE(mm/day)	RMSE(mm/day)	NSE	R ²
DNN-Ret	Training	0.370	0.508	0.937	0.941
	Testing	0.379	0.527	0.932	0.935
XGB-Ret	Training	0.399	0.540	0.920	0.926
	Testing	0.412	0.574	0.908	0.919
RF-Ret	Training	0.396	0.536	0.922	0.927
	Testing	0.419	0.582	0.906	0.917

The RMSE value of 0.472 mm/day for DNN-Ret indicates a decline of 6.5% and 4.8% relative to RF-Ret and XGB-Ret. There is a marginal gap in NSE (0.934 and 0.937) and R² (0.944 and 0.946) between RF-Ret and XGB-Ret. However, 0.95 of NSE and 0.954 of R² shows a substantial increase in DNN-Ret forecast accuracy compared to the other two models.

The test records of Pomona weather station dataset proved the DNN-Ret model as the ideal ETo estimation model because it showed the lowest MAE (0.325 mm/day) and RMSE (0.464 mm/day) as well as the highest NSE (0.934) and R² (0.938) relative to other models. In MAE and RMSE, the DNN-Ret obtained a 6.9% and 5.5% decrease compared to RF-Ret, and a 5.2% and 4.1% decrease compared to XGB-Ret. Although the DNN-Ret showed a drop in the MAE and RMSE error estimates relative to the Oakville dataset, there was no significant improvement in the NSE and R² estimates. However, in the Pomona ETo variants, DNN-Ret showed a considerable improvement in NSE and R² (NSE=0.934, R²=0.938) relative to RF-Ret (NSE=0.917, R²=0.924) and XGB-Ret (NSE=0.918, R²=0.926).

The ETo models developed using the Alturas dataset exhibited lower efficiency compared to the models developed in other stations, irrespective of the architecture employed. On an average, the models reported a MAE of 0.525 mm/day, an RMSE of 0.712 mm/day, an NSE of 0.883, and an R² of 0.894 in Alturas. This may be caused by the characteristics

of the dataset, as this station is at a high altitude and experiences a mild and temperate climate. However, DNN-Ret maintained its consistency in performance excellence when compared to RF-Ret (13.2% and 13.9% decrease in MAE and RMSE) and XGB-Ret (12.2% and 12.8% decrease in MAE and RMSE). Both RF-Ret and XGB-Ret underperformed, with NSE values of 0.868 and 0.87 and R^2 values of 0.883 and 0.886, respectively, compared to DNN-Ret, which scored $NSE=0.911$ and $R^2=0.914$.

A comparison of the relationship between the predicted and actual daily ETo values for the developed models at Oakville, Pomona and Alturas during the testing period is illustrated in Figs. 8, 9 and 10 respectively. The line charts, which depict the parallel distribution between the actual and predicted values, clearly indicate that the predicted ETo values of DNN-Ret closely follow the corresponding FPM-56 ETo values in all stations (Figs. 8a, 9a and 10a). The plotted data points in the scatter plots of DNN-Ret (Figs. 8d, 9d and 10d) show more closeness towards the 1:1 regression line relative to XGB-Ret and RF-Ret in Pomona and Alturas weather stations. At Oakville station, XGB-Ret exhibited a similar scattered ETo estimate to DNN-Ret with an R^2 of 0.95.

Fig. 8 Line charts and scatter plots showing the comparison of predicted and actual ETo values and their relationship for **a, d** DNN-Ret, **b, e** XGB-Ret, and **c, f** RF-Ret at Oakville weather station during the testing period

Fig. 9 Line charts and scatter plots showing the comparison of predicted and actual ETo values and their relationship for **a, d** DNN-Ret, **b, e** XGB-Ret, and **c, f** RF-Ret at Pomona weather station during the testing period

Figure 11 presents the performance comparison metrics for the developed R_S based models at Oakville, Pomona and Alturas weather stations. The plotted data points for each station exhibit the performance excellence of DNN-Ret by showing a lower MAE and RMSE and a higher NSE and R^2 compared to other models.

3.2.2 Evaluation of the Developed Models in the General Scenario

Table 8 lists the mean performance metrics of the models during the training and testing phases of the model construction process using the three CIMIS data sets. Figure 12 illustrates the average model performance evaluation of DNN-Ret, XGB-Ret, and RF-Ret during both training and testing phases.

Regardless of the dataset, DNN-Ret exhibited superior results with MAE=0.370 mm/day, RMSE=0.508 mm/day, NSE=0.937, and $R^2=0.941$ during training, and MAE=0.379 mm/day, RMSE=0.527 mm/day, NSE=0.932, and $R^2=0.935$ during testing. The negligible variation in the training and testing metric data of the DNN-Ret

Fig. 10 Line charts and scatter plots showing the comparison of predicted and actual ETo values and their relationship for **a, d** DNN-Ret, **b, e** XGB-Ret, and **c, f** RF-Ret at Alturas weather station during the testing period

signifies its capability to cope with the overfitting issues. The flexibility shown by the model for simulating the dataset records with different climatic attributes is evident from the average percentage decrease in both MAE and RMSE and a percentage rise in both NSE and R^2 quantities. With the same 9.5% decrease in both MAE and RMSE and 2.9% and 2% increase in NSE and R^2 with respect to RF-Ret and an 8% decrease in MAE and 8.2% reduction in RMSE as well as a 2.6% increase in NSE and 1.7% increase in R^2 relative to XGB-Ret, DNN-Ret prove its consistency in ETo modelling performance at all stations.

The remaining RF-Ret and XGB-Ret models reported a relatively equivalent performance during the testing phase of the modelling with a magnitude of variation of 0.007 mm/day, 0.008 mm/day, 0.002, and 0.002 for MAE, RMSE, NSE and R^2 respectively. Yet, the RF-Ret showed a considerable difference in the evaluation parameter values during the training and testing phases. This variation signals some glimpses of overfitting in the RF-Ret model. Furthermore, a reduction in the MAE and RMSE and an

Fig. 11 Evaluation of the developed ETo prediction models in the local scenario: Plots comparing the performance evaluation metrics **a** MAE, **b** RMSE, **c** NSE and **d** R^2 for DNN-Ret, XGB-Ret, and RF-Ret with regard to Oakville, Pomona, and Alturas weather stations during the testing period

Fig. 12 Evaluation of the developed ETo prediction models in the general scenario: Plots comparing the performance evaluation metrics **a** MAE, **b** RMSE, **c** NSE and **d** R^2 for DNN-Ret, XGB-Ret, and RF-Ret during the training and testing periods, by averaging it across Oakville, Pomona, and Alturas weather stations

improvement in the NSE and R^2 values of the XGB-Ret assisted it to hit the second position among the developed ETo models.

4 Discussion

DNN-Ret, the DNN architecture implemented in the CIMIS datasets of the present study, outperformed XGB-Ret and RF-Ret by exhibiting the best performance measures such as MAE of 0.325 mm/day and RMSE of 0.464 mm/day in the Pomona weather station and NSE of 0.95 and R^2 of 0.954 in the Oakville weather station. In the general scenario, DNN-Ret won by achieving an average performance measure of MAE=0.379 mm/day, RMSE=0.527 mm/day, NSE=0.932, and R^2 =0.935. This study complements the previous research of Sowmya et al. (2020), which implemented the DNN architecture in CIMIS datasets for ETo prediction and concluded with a recommendation to use $T_{\max}$ and R_s for modelling. They attempted utilizing simply the R_s parameter as input in the ETo modelling and reported an RMSE of 0.61 mm/day and an R^2 of 0.91 in the Oakville dataset. The present study improves the results by providing a decrease of 23% in RMSE and an increase of 5% in R^2 through the sophisticated hyper parameter optimization of the DNN model architecture. Saggi and Jain (2019) modelled a DNN based ETo estimation process using the entire climatic parameters required by the FPM-56 equation. They reported RMSE ranging from 0.1921 to 0.2691 mm/day and R^2 ranging from 0.95 to 0.99. Özgür and Yamaç (2020) subsequently improved this model by using the SeLU activation function in their DNN architecture, yielding the best performance metrics of RMSE=0.2073 mm/day and R^2 =0.9934 at the Aksaray weather station, Turkey. Recently, Nagappan et al. (2020) developed a one-dimensional CNN ETo estimation model with only three input parameters extracted from PCA with R^2 =0.979 and RMSE=0.21 mm/day as the performance metrics. These values fall within the performance ranges of the previously mentioned full input parameter investigations (R^2 =0.95–0.99 and RMSE=0.19–0.27 mm/day). The observed results point to the fact that the efficiency of the ETo prediction model depends more on the model architecture and the optimization of hyper parameters done during the design phase rather than the number of input variables used. Therefore, this study focused on both of these criteria and demonstrated superior performance accuracy with only one input parameter when compared to prior minimal parameter-based deep learning studies that employed more than one parameter for modelling. For instance, Ferreira and da Cunha (2020a) reported that the CNN models developed with hourly temperature and humidity have reduced the RMSE values from 0.71 to 0.51 and raised the R^2 from 0.79 to 0.88 in the local context, and the lowest RMSE of 0.45 and the highest R^2 value of 0.87 were observed in the regional scenario. Chen et al. (2020b) reported that the temperature, radiation, and humidity-based TCN and long short-term memory models outscored the other machine learning and empirical models developed in the local scenario by 0.755–0.49 in RMSE and 0.83–0.92 in R^2 . In the pooled scenario, a maximum of 0.5 in RMSE and 0.91 in R^2 was also observed. The RNN architecture designed by Afzaal et al. (2020) also exhibited a minimum RMSE of 0.38 mm/day and a maximum R^2 of 0.92. Hence, the overall performance analysis confirmed the excellence of the proposed single input DNN model that exhibited RMSE values ranging from 0.472 to 0.646 mm/day and R^2 values ranging from 0.914 to 0.954 in the local context, and the best RMSE of 0.527 and R^2 of 0.935 in the general context.

The temperature parameter was the focus of the majority of the minimal parameter-based ETo modelling studies reported in the literature. Compared to other climate variables, the temperature measuring process tends to be very convenient and cost-effective, and thus, ETo soft computing based prediction studies using temperature are leading, mostly in humid areas. Reis et al. (2019) modelled ETo prediction using ANN and MLR, employing only maximum and minimum temperatures, and resulting in $NSE=0.687$ and 0.55 in local and general applications, respectively. In semiarid, arid, subhumid, and humid climatic zones, Adamala (2018) built generalized wavelet neural networks with RMSE values of 0.658 , 0.92 , 0.66 , and 0.62 mm/day and NSE values of 0.83 , 0.81 , 0.87 , and 0.76 , respectively. Although these statistical measures indicated lower performance when compared to the performance metric of current study, certain temperature-based studies revealed a slight improvement in MAE and R^2 . For instance, Sanikhani et al. (2019) assessed the performance of six AI models for ETo estimation using two weather station datasets in Turkey and observed that the RBNN and ANFIS-SC models performed well at the Isparta station, with $MAE=0.332$ mm/day, $RMSE=0.43$ mm/day, and $R^2=0.929$. Feng et al. (2017b) reported the efficiency of ELM in both local and pooled dataset implementation with $MAE=0.267$ and 0.263 mm/day, $NSE=0.891$ and 0.895 , respectively. Compared to the RMSE values of the present study, the highest performing MLP models published by Rahimikhoob (2010) ($RMSE=0.41$ mm/day and $R^2=0.95$), and Wang et al. (2011) ($RMSE=0.3$ mm/day and $R^2=0.92$) have shown a modest performance enhancement. This could be because these studies, in addition to temperature, included the extra-terrestrial radiation parameter in ETo modelling.

Solar radiation plays a vital role in the quantification of ETo, in addition to temperature. Numerous ETo modelling studies have documented the efficacy of solar radiation to improve the predictive performance of soft computing models (Fan et al. 2018; Petković et al. 2020; Üneş et al. 2020; Kazemi et al. 2021). Radiation-based empirical models commonly use temperature and solar radiation as the input parameters for estimation, which have always resulted in a lower outcome compared to the prediction performance of soft computing models employing the same input parameters for modelling (Antonopoulos and Antonopoulos 2017; Chia et al. 2020a; Zhu et al. 2020). However, no studies have been reported in the literature intending to estimate ETo using only solar radiation data using soft computing methods, although various studies attempted it for their model's performance comparison. Üneş et al. (2020) attempted to model ETo using different combinations of input parameters, including individual parameters alone, using MLR and SVM. They reported the same correlation coefficient of 0.749 , MAE of 0.579 and 0.620 mm/day, and MSE of 0.634 and 0.533 mm/day for the solar radiation input model. Kaya et al. (2021) tested several input combinations for the ETo modelling. The single input parameter modelling results demonstrated that the solar radiation input model using SVR outperformed the MLP and MLR models with $MAE=0.575$ mm/day, $RMSE=0.814$ mm/day, and $R^2=0.888$. The performance metrics of the single input parameter-based radial basis M5 model tree models developed by Kisi et al. (2021) using three Turkish weather station datasets revealed that the solar radiation input model developed at the Isparta weather station outperformed the temperature model with $MAE=0.524$ mm/day, $RMSE=0.678$ mm/day, $NSE=0.835$, and $R^2=0.927$. With the gradient boosting with categorical feature support technique, the ETo estimation models developed by Wu et al. (2020) achieved an accuracy with an average RMSE of 0.567 mm/day, MAE of 0.420 mm/day, and Adj- R^2 of 0.8514 . These observations demonstrated that DNN-Ret surpassed all the above-mentioned solar radiation input models. In comparison to the results presented by Kaya et al. (2021) and Kisi et al. (2021), DNN-Ret showed a significant reduction in MAE from 0.575 and

0.524 to 0.325 mm/day, a reduction in RMSE from 0.814 and 0.678 to 0.464 mm/day, a rise from 0.835 to 0.95 in NSE, and a rise from 0.888 and 0.927 to 0.954 in R^2 .

One of the objectives of this research is to bring minimalism in terms of the number of input parameters, and therefore, it focused on selecting the most dominant single climatic parameter on ETo for modelling. The FI-based input selection phase of this study revealed that the R_S is the most influential parameter on ETo in all stations considered. The SHAP validated and interpreted the tree based XGB-FI and RF-FI approaches and supported the finding. The summary plots and dependency plots generated during the feature analysis phase of SHAP identified temperature ($T_{\max}$ and T_{avg}) and $RH_{\min}$ were exhibiting some interaction and associativity with R_S in the ETo prediction, although this varied depending on the dataset. However, regardless of the datasets, the R_S parameter maintained consistency in FI and was chosen as the sole parameter as input to the proposed DNN architecture. However, R_S is not a widely used measure in underdeveloped countries, and hence, R_S -based ETo modelling studies are scarce in those regions. This may be because R_S measurement equipment, such as pyranometers, is very costly, and certain research has also found that the influence of R_S on ETo predictions differs for various climatic areas (Fan et al. 2019). Difficulty in capturing R_S can be overcome by using the Angstrom formula as it defines a connection between extra-terrestrial radiation and relative duration of sunshine without R_S data (Patil and Deka 2017). Allen et al. (1998) suggested other procedures to estimate the missing R_S data from $T_{\max}$ and $T_{\min}$. In some climatic zones, RH and temperature have a greater impact on ETo than R_S . The scope of this research is not restricted to R_S as the sole single input, but to any parameter that has the most significant influence on ETo. As a result, this study provides a generic approach for picking a single input parameter and developing a DNN architecture for ETo estimation based on that parameter. The inclusion of RH and temperature will certainly improve the performance of ETo models; nevertheless, that variation in the prediction accuracy might well be regained through an efficient deep learning architecture. Moreover, seasonal variations in ETo have been identified in certain studies (Gong et al. 2006), which can be addressed by applying machine learning-based clustering techniques to classify the climatic characteristics of the data and model them accordingly (Chen et al. 2020b).

5 Conclusions

This paper presented a new methodology for ETo modelling employing only one meteorological variable as input and exploring the state-of-the-art deep learning approach, DNN, for model architecture construction. This study proposed two FI measuring techniques that are integrated into the tree-based ensemble machine learning techniques, XGBoost (XGB-FI) and RF (RF-FI), to select the most influential input parameter for ETo modelling. This research also investigated the potential of a game theory approach, SHAP, to interpret the input selection methods and analyze the features of the dataset. SHAP ensured the reliability of the selected meteorological parameter in ETo modelling by measuring the contributions of each feature to the ETo prediction. Because the number of input parameters positively correlates with the efficiency of the machine learning model, single input parameter modelling may result in a decrease in prediction accuracy. However, this can be mitigated through sophisticated optimization of DNN hyper parameters, and thus, establish an efficient mapping between input and output.

In this research, the proposed methodology was implemented by modelling the CIMIS Penman ETo paradigm with DNN and using a single climatic parameter as input, without compromising accuracy. Initially, the study identified solar radiation as the most influential parameter for ETo prediction using XGB-FI and RF-FI, which was verified using the SHAP technique. SHAP's visualization interpretations confirmed the findings of XGB-FI and RF-FI and recommended solar radiation as the top-ranked feature in all CIMIS datasets under consideration for the case study. The model building process incorporated three distinct datasets from various ET zones during model training, aiming for a specific model suitable for each evapotranspiration zone. For each of the three CIMIS weather stations, three models were developed using the DNN, XGBoost, and RF frameworks, with CIMIS-Penman ETo as a target feature and solar radiation as an input feature, namely DNN-Ret, XGB-Ret, and RF-Ret, respectively. The performance of the models was verified in both local and general scenarios, and from the overall result analysis, DNN-Ret had a comparable performance enhancement relative to the other two models. The best accuracy was observed for DNN-Ret at Oakville weather station with MAE=0.335 mm/day, RMSE=0.472 mm/day, NSE=0.95, $R^2=0.954$. The mean metric values of DNN-Ret exhibited a decrease of 9.5% in MAE and RMSE, 2.9% increase NSE and 2% increase in R^2 compared to RF-Ret. A similar performance enhancement of DNN-Ret was also visible with respect to XGB-Ret, with approximately 8% decrease in MAE and RMSE, 2.6% increase in NSE, and 1.7% increase in R^2 . The XGB-Ret and RF-Ret were similar in performance, and of these, XGB-Ret performed more consistently with both training and testing datasets. Therefore, from the overall efficiency assessment, DNN-Ret, XGB-Ret, and RF-Ret were ranked first, second, and third, respectively.

The hydro and agro-communities are pursuing ET research outcomes based on minimal climate data, especially in developing countries where the weather station network is not well established and the collection of meteorological parameters such as wind speed and relative humidity is not feasible. This research is immensely significant in such situations, as it utilizes only one input parameter, exploits the feature learning capability of DNN in input and output mapping, and achieves robust performance in ETo estimation. However, this study could not be applied in a global scenario because it is focused on developing a unique ETo model that depicted the climatic characteristics of a specific ET zone. ETo is merely a meteorological parameter that is highly influenced by environmental and seasonal variations. Therefore, the order of significance of features influencing ETo varies depending on the characteristics of various climatic zones. Hence, a generalized and consistent single input parameter-based ETo prediction model suitable for all types of climates is needed in the ET domain to provide a cost-effective and accurate solution for ETo prediction. The future research can include the development of deep learning models hybridized with unsupervised machine learning techniques like clustering to aggregate data from diverse climatic zones into a pooled dataset and subsequently utilize it to build a global ETo model.

Author contributions RMS: Conceptualization, Methodology, Software, Validation, Data curation, Writing—Original draft. BMSK: Conceptualization, Investigation, Resources, Project administration, Writing—Review and editing. KNAS: Formal analysis, Software, Validation, Visualization, Resources, Writing—Review and editing.

Funding This research did not receive any specific grant from funding agencies.

Data Availability Data set is freely available from the CIMIS website and other materials that support the findings of this study are available on request from the corresponding author.

Code Availability Available with the corresponding author.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- Abrishami N, Sepaskhah AR, Shahrokhnia MH (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>
- Adamala S (2018) Temperature based generalized wavelet-neural network models to estimate evapotranspiration in India. *Inf Process Agric* 5:149–155. <https://doi.org/10.1016/j.inpa.2017.09.004>
- Adamala S, Raghuwanshi NS, Mishra A (2015) Generalized quadratic synaptic neural networks for ETo Modeling. *Environ Process* 2:309–329. <https://doi.org/10.1007/s40710-015-0066-6>
- Afzaal H, Farooque AA, Abbas F, Acharya B, Esau T (2020) Computation of evapotranspiration with artificial intelligence for precision water resource management. *Appl Sci*. <https://doi.org/10.3390/app10051621>
- Ahmadi F, Mehdizadeh S, Mohammadi B, Pham QB, Doan TN, Vo ND (2021) Application of an artificial intelligence technique enhanced with intelligent water drops for monthly reference evapotranspiration estimation. *Agric Water Manag* 244:106622. <https://doi.org/10.1016/j.agwat.2020.106622>
- Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration—guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. FAO, United Nations, 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>
- Başağaoğlu H, Chakraborty D, Winterle J (2021) Reliable evapotranspiration predictions with a probabilistic machine learning framework. *Water*. <https://doi.org/10.3390/w13040557>
- Bellido-Jiménez JA, Estévez J, García-Marín AP (2021) New machine learning approaches to improve reference evapotranspiration estimates using intra-daily temperature-based variables in a semi-arid region of Spain. *Agric Water Manag*. <https://doi.org/10.1016/j.agwat.2020.106558>
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32. <https://doi.org/10.1023/A:1010933404324>
- Chen T, Guestrin C (2016) Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. pp 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chen M, Liu Q, Chen S, Liu Y, Zhang CH, Liu R (2019) XGBoost-based algorithm interpretation and application on post-fault transient stability status prediction of power system. *IEEE Access* 7:13149–13158. <https://doi.org/10.1109/ACCESS.2019.2893448>
- Chen Z, Sun S, Wang Y, Wang Q, Zhang X (2020a) Temporal convolution-network-based models for modeling maize evapotranspiration under mulched drip irrigation. *Comput Electron Agric* 169:105206. <https://doi.org/10.1016/j.compag.2019.105206>
- Chen Z, Zhu Z, Jiang H, Sun S (2020b) Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. *J Hydrol* 591:125286. <https://doi.org/10.1016/j.jhydrol.2020.125286>
- Chia MY, Huang YF, Koo CH (2020a) Support vector machine enhanced empirical reference evapotranspiration estimation with limited meteorological parameters. *Comput Electron Agric* 175:105577. <https://doi.org/10.1016/j.compag.2020.105577>
- Chia MY, Huang YF, Koo CH, Fung KF (2020b) Recent advances in evapotranspiration estimation using artificial intelligence approaches with a focus on hybridization techniques—a review. *Agronomy*. <https://doi.org/10.3390/agronomy10010101>
- Datta L (2020) A survey on activation functions and their relation with xavier and he normal initialization. *arXiv Preprint*. <https://arxiv.org/abs/2004.06632>

- de Oliveira e Lucas P, Alves MA, de Lima e Silva PC, Guimarães FG (2020) Reference evapotranspiration time series forecasting with ensemble of convolutional neural networks. *Comput Electron Agric* 177:105700. <https://doi.org/10.1016/j.compag.2020.105700>
- Debnath S, Adamala S, Raghuvanshi NS (2015) Sensitivity analysis of FAO-56 Penman-Monteith method for different agro-ecological regions of India. *Environ Process* 2:689–704. <https://doi.org/10.1007/s40710-015-0107-1>
- Effrosynidis D, Arampatzis A (2021) An evaluation of feature selection methods for environmental data. *Ecol Inform* 61:101224. <https://doi.org/10.1016/J.ECOINF.2021.101224>
- 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>
- Fan J, Ma X, Wu L, Zhang F, Yu X, Zeng W (2019) Light Gradient Boosting Machine: an efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. *Agric Water Manag* 225:105758. <https://doi.org/10.1016/j.agwat.2019.105758>
- Farzaneh N, Williamson CA, Gryak J, Najarian K (2021) A hierarchical expert-guided machine learning framework for clinical decision support systems: an application to traumatic brain injury prognostication. *npj Digit Med* 4:78. <https://doi.org/10.1038/s41746-021-00445-0>
- 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>
- Ferreira LB, da Cunha FF (2020a) New approach to estimate daily reference evapotranspiration based on hourly temperature and relative humidity using machine learning and deep learning. *Agric Water Manag* 234:106113. <https://doi.org/10.1016/j.agwat.2020.106113>
- Ferreira LB, da Cunha FF (2020b) Multi-step ahead forecasting of daily reference evapotranspiration using deep learning. *Comput Electron Agric* 178:105728. <https://doi.org/10.1016/J.COMPAG.2020.105728>
- Ferreira LB, da Cunha FF, de Oliveira RA, Fernandes Filho EI (2019) Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM—a new approach. *J Hydrol* 572:556–570. <https://doi.org/10.1016/j.jhydrol.2019.03.028>
- Gavrilović S, Denić N, Petković D, Živić NV, Vujičić S (2018) Statistical evaluation of mathematics lecture performances by soft computing approach. *Comput Appl Eng Educ* 26:902–905. <https://doi.org/10.1002/cae.21931>
- Glorot X, Bengio Y (2010) Understanding the difficulty of training deep feedforward neural networks. In: *Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR workshop and conference proceedings*. pp 249–256
- Gong L, Xu C, Chen D, Halldin S, Chen YD (2006) Sensitivity of the Penman-Monteith reference evapotranspiration to key climatic variables in the Changjiang (Yangtze River) basin. *J Hydrol* 329:620–629. <https://doi.org/10.1016/J.JHYDROL.2006.03.027>
- Gong D, Hao W, Gao L, Feng Y, Cui N (2021) Extreme learning machine for reference crop evapotranspiration estimation: model optimization and spatiotemporal assessment across different climates in China. *Comput Electron Agric* 187:106294. <https://doi.org/10.1016/j.compag.2021.106294>
- Han Z, Zhao J, Leung H, Ma KF, Wang W (2019) A review of deep learning models for time series prediction. *IEEE Sens J* 21:7833–7848. <https://doi.org/10.1109/jsen.2019.2923982>
- He K, Zhang X, Ren S, Sun J (2015) Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: *Proceedings of the IEEE international conference on computer vision 2015*. pp 1026–1034. <https://doi.org/10.1109/ICCV.2015.123>
- Jabeur SB, Mefteh-Wali S, Viviani J-L (2021) Forecasting gold price with the XGBoost algorithm and SHAP interaction values. *Ann Oper Res*. <https://doi.org/10.1007/s10479-021-04187-w>
- Joharestani MZ, Cao C, Ni X, Bashir B, Talebiesfandarani S (2019) PM2.5 prediction based on random forest, XGBoost, and deep learning using multisource remote sensing data. *Atmosphere (Basel)* 10:373. <https://doi.org/10.3390/atmos10070373>
- Jovic S, Nedeljkovic B, Golubovic Z, Kostic N (2018) Evolutionary algorithm for reference evapotranspiration analysis. *Comput Electron Agric* 150:1–4. <https://doi.org/10.1016/j.compag.2018.04.003>
- Karimi S, Shiri J, Marti P (2020) Supplanting missing climatic inputs in classical and random forest models for estimating reference evapotranspiration in humid coastal areas of Iran. *Comput Electron Agric* 176:105633. <https://doi.org/10.1016/j.compag.2020.105633>

- Kaya YZ, Zelenakova M, Üneş F, Demirci M, Hlavata H, Mesáros P (2021) Estimation of daily evapotranspiration in Košice City (Slovakia) using several soft computing techniques. *Theor Appl Climatol* 144:287–298. <https://doi.org/10.1007/s00704-021-03525-z>
- Kazemi MH, Majnooni-Heris A, Kisi O, Shiri J (2021) Generalized gene expression programming models for estimating reference evapotranspiration through cross-station assessment and exogenous data supply. *Environ Sci Pollut Res* 28:6520–6532. <https://doi.org/10.1007/s11356-020-10916-8>
- Kisi O (2011) Modeling reference evapotranspiration using evolutionary neural networks. *J Irrig Drain Eng* 137:636–643. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000333](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000333)
- Kisi O, Keshtegar B, Zounemat-Kermani M, Heddam S, Trung NT (2021) Modeling reference evapotranspiration using a novel regression-based method: radial basis M5 model tree. *Theor Appl Climatol*. <https://doi.org/10.1007/s00704-021-03645-6>
- Kumar M, Raghuwanshi NS, Singh R, Wallender WW, Pruitt WO (2002) Estimating evapotranspiration using artificial neural network. *J Irrig Drain Eng* 128:224–233. [https://doi.org/10.1061/\(ASCE\)0733-9437\(2002\)128:4\(224\)](https://doi.org/10.1061/(ASCE)0733-9437(2002)128:4(224))
- Kuzman B, Petković B, Denić N, Petković D, Ćirković B, Stojanović J, Milić M (2021) Estimation of optimal fertilizers for optimal crop yield by adaptive neuro fuzzy logic. *Rhizosphere* 18:100358. <https://doi.org/10.1016/J.RHISPH.2021.100358>
- Lecun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444. <https://doi.org/10.1038/nature14539>
- Lin WC, Tsai CF (2020) Missing value imputation: a review and analysis of the literature (2006–2017). *Artif Intell Rev* 53:1487–1509. <https://doi.org/10.1007/s10462-019-09709-4>
- Lundberg SM, Lee S-I (2017) A unified approach to interpreting model predictions. In: *Proceedings of the 31st international conference on neural information processing systems*. pp 4768–4777
- Lundberg SM, Erion GG, Lee S-I (2018) Consistent individualized feature attribution for tree ensembles. *arXiv Preprint*. <https://arxiv.org/abs/1802.03888>
- Martín J, Sáez JA, Corchado E (2021) On the suitability of stacking-based ensembles in smart agriculture for evapotranspiration prediction. *Appl Soft Comput* 108:107509. <https://doi.org/10.1016/j.asoc.2021.107509>
- Meng Y, Yang N, Qian Z, Zhang G (2021) What makes an online review more helpful: an interpretation framework using XGBoost and SHAP values. *J Theor Appl Electron Commer Res* 16:466–490. <https://doi.org/10.3390/jtaer16030029>
- Milić M, Petković B, Selmi A, Petković D, Jermisittiparsert K, Radivojević A, Milovancevic M, Khan A, Vidosavljević ST, Denić N, Kuzman B (2021) Computational evaluation of microalgae biomass conversion to biodiesel. *Biomass Convers Biorefinery*. <https://doi.org/10.1007/s13399-021-01314-2>
- Mohammadi B, Mehdizadeh S (2020) Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. *Agric Water Manag* 237:106145. <https://doi.org/10.1016/j.agwat.2020.106145>
- Muhammad MKI, Shahid S, Ismail T, Harun S, Kisi O, Yaseen ZM (2021) The development of evolutionary computing model for simulating reference evapotranspiration over Peninsular Malaysia. *Theor Appl Climatol* 144:1419–1434. <https://doi.org/10.1007/s00704-021-03606-z>
- Nagappan M, Gopalakrishnan V, Alagappan M (2020) Prediction of reference evapotranspiration for irrigation scheduling using machine learning. *Hydrol Sci J* 65:2669–2677. <https://doi.org/10.1080/02626667.2020.1830996>
- Nikolić V, Petković D, Lazov L, Milovančević M (2016) Selection of the most influential factors on the water-jet assisted underwater laser process by adaptive neuro-fuzzy technique. *Infrared Phys Technol* 77:45–50. <https://doi.org/10.1016/J.INFRARED.2016.05.021>
- Özgür A, Yamaç SS (2020) Modelling of daily reference evapotranspiration using deep neural network in different climates. *arXiv Preprint*. <https://arxiv.org/abs/2006.01760>
- Patil AP, Deka PC (2017) Performance evaluation of hybrid Wavelet-ANN and Wavelet-ANFIS models for estimating evapotranspiration in arid regions of India. *Neural Comput Appl* 28:275–285. <https://doi.org/10.1007/s00521-015-2055-0>
- Petković D, Gocic M, Trajkovic S, Shamshirband S, Motamedi S, Hashim R, Bonakdari H (2015) Determination of the most influential weather parameters on reference evapotranspiration by adaptive neuro-fuzzy methodology. *Comput Electron Agric* 114:277–284. <https://doi.org/10.1016/j.compag.2015.04.012>
- Petković D, Gocic M, Trajkovic S, Milovančević M, Šević D (2017a) Precipitation concentration index management by adaptive neuro-fuzzy methodology. *Clim Change* 141:655–669. <https://doi.org/10.1007/s10584-017-1907-2>

- Petković M, Džeroski S, Kocev D (2017b) Feature ranking for multi-target regression with tree ensemble methods. In: International conference on discovery science. Springer, pp 171–185. https://doi.org/10.1007/978-3-319-67786-6_13
- Petković B, Petković D, Kuzman B, Milovančević M, Wakil K, Ho LS, Jermisittiparsert K (2020) Neuro-fuzzy estimation of reference crop evapotranspiration by neuro fuzzy logic based on weather conditions. *Comput Electron Agric* 173:105358. <https://doi.org/10.1016/j.compag.2020.105358>
- Petković D, Barjaktarovic M, Milošević S, Denić N, Spasić B, Stojanović J, Milovancevic M (2021) Neuro fuzzy estimation of the most influential parameters for Kusum biodiesel performance. *Energy* 229:120621. <https://doi.org/10.1016/J.ENERGY.2021.120621>
- Ponraj AS, Vigneswaran T (2020) Daily evapotranspiration prediction using gradient boost regression model for irrigation planning. *J Supercomput* 76:5732–5744. <https://doi.org/10.1007/s11227-019-02965-9>
- Rahimikhoob A (2010) Estimation of evapotranspiration based on only air temperature data using artificial neural networks for a subtropical climate in Iran. *Theor Appl Climatol* 101:83–91. <https://doi.org/10.1007/s00704-009-0204-z>
- Reis MM, da Silva AJ, Zullo Junior J, Santos LD, Azevedo AM, Lopes ÉM (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>
- Roy DK (2021) Long short-term memory networks to predict one-step ahead reference evapotranspiration in a subtropical climatic zone. *Environ Process* 8:911–941. <https://doi.org/10.1007/s40710-021-00512-4>
- Saggi MK, 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>
- Sanikhani H, Kisi O, Maroufpoor E, Yaseen ZM (2019) Temperature-based modeling of reference evapotranspiration using several artificial intelligence models: application of different modeling scenarios. *Theor Appl Climatol* 135:449–462. <https://doi.org/10.1007/s00704-018-2390-z>
- Shang C, Chen W-H, Stroock AD, You F (2019) Robust model predictive control of irrigation systems with active uncertainty learning and data analytics. *IEEE Trans Control Syst Technol* 28:1493–1504. <https://doi.org/10.1109/TCST.2019.2916753>
- Shi X, Wong YD, Li MZF, Palanisamy C, Chai C (2019) A feature learning approach based on XGBoost for driving assessment and risk prediction. *Accid Anal Prev* 129:170–179. <https://doi.org/10.1016/j.aap.2019.05.005>
- Sowmya MR, Kumar MBS, Ambat SK (2020) Comparison of deep neural networks for reference evapotranspiration prediction using minimal meteorological data. In: 2020 advanced computing and communication technologies for high performance applications (ACCTHPA). pp 27–33. <https://doi.org/10.1109/ACCTHPA49271.2020.9123201>
- Tikhamarine Y, Malik A, Souag-Gamane D, Kisi O (2020) Artificial intelligence models versus empirical equations for modeling monthly reference evapotranspiration. *Environ Sci Pollut Res* 27:30001–30019. <https://doi.org/10.1007/s11356-020-08792-3>
- Trajkovic S (2005) Temperature-based approaches for estimating reference evapotranspiration. *J Irrig Drain Eng* 131:316–323. [https://doi.org/10.1061/\(ASCE\)0733-9437\(2005\)131:4\(316\)](https://doi.org/10.1061/(ASCE)0733-9437(2005)131:4(316))
- Üneş F, Kaya YZ, Mamak M (2020) Daily reference evapotranspiration prediction based on climatic conditions applying different data mining techniques and empirical equations. *Theor Appl Climatol* 141:763–773. <https://doi.org/10.1007/s00704-020-03225-0>
- Vieira S, Lopez Pinaya WH, Garcia-Dias R, Mechelli A (2020) Deep neural networks. *Mach Learn Methods Appl Brain Disord*. <https://doi.org/10.1016/B978-0-12-815739-8.00009-2>
- Wang YM, Traore S, Kerh T, Leu JM (2011) Modelling reference evapotranspiration using feed forward backpropagation algorithm in arid regions of Africa. *Irrig Drain* 60:404–417. <https://doi.org/10.1002/ird.589>
- Wang S, Lian J, Peng Y, Hu B, Chen H (2019) Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi, China. *Agric Water Manag* 221:220–230. <https://doi.org/10.1016/j.agwat.2019.03.027>
- Wu L, Fan J (2019) Comparison of neuron-based, kernel-based, tree-based and curve-based machine learning models for predicting daily reference evapotranspiration. *PLoS ONE* 14:e0217520. <https://doi.org/10.1371/journal.pone.0217520>
- Wu T, Zhang W, Jiao X, Guo W, Hamoud YA (2020) Comparison of five boosting-based models for estimating daily reference evapotranspiration with limited meteorological variables. *PLoS ONE* 15:e0235324. <https://doi.org/10.1371/journal.pone.0235324>
- Wu T, Zhang W, Jiao X, Guo W, Hamoud YA (2021) Evaluation of stacking and blending ensemble learning methods for estimating daily reference evapotranspiration. *Comput Electron Agric* 184:106039. <https://doi.org/10.1016/j.compag.2021.106039>

- Xiang K, Li Y, Horton R, Feng H (2020) Similarity and difference of potential evapotranspiration and reference crop evapotranspiration—a review. *Agric Water Manag* 232:106043. <https://doi.org/10.1016/J.AGWAT.2020.106043>
- Xing X, Liu Y, Zhao W, Kang D, Yu M, Ma X (2016) Determination of dominant weather parameters on reference evapotranspiration by path analysis theory. *Comput Electron Agric* 120:10–16. <https://doi.org/10.1016/j.compag.2015.11.001>
- Yaseen ZM, Allawi MF, Yousif AA, Jaafar O, Hamzah FM, El-Shafie A (2018) Non-tuned machine learning approach for hydrological time series forecasting. *Neural Comput Appl* 30:1479–1491. <https://doi.org/10.1007/s00521-016-2763-0>
- Zhang D, Qian L, Mao B, Huang C, Huang B, Si Y (2018) A data-driven design for fault detection of wind turbines using random forests and XGboost. *IEEE Access* 6:21020–21031. <https://doi.org/10.1109/ACCESS.2018.2818678>
- Zhu X, Li J, Zhu M, Jiang Z, Li Y (2018) An evaporation duct height prediction method based on deep learning. *IEEE Geosci Remote Sens Lett* 15:1307–1311. <https://doi.org/10.1109/LGRS.2018.2842235>
- Zhu B, Feng Y, Gong D, Jiang S, Zhao L, Cui N (2020) Hybrid particle swarm optimization with extreme learning machine for daily reference evapotranspiration prediction from limited climatic data. *Comput Electron Agric* 173:105430. <https://doi.org/10.1016/j.compag.2020.105430>
- Zoabi Y, Deri-Rozov S, Shomron N (2021) Machine learning-based prediction of COVID-19 diagnosis based on symptoms. *npj Digit Med* 4:3. <https://doi.org/10.1038/s41746-020-00372-6>

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