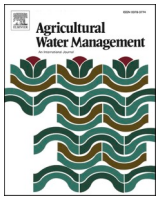
Agricultural Water Management 290 (2023) 108604

Contents lists available at ScienceDirect

Agricultural Water Management

journal homepage: www.elsevier.com/locate/agwat



Development of an enhanced bidirectional recurrent neural network combined with time-varying filter-based empirical mode decomposition to forecast weekly reference evapotranspiration

Masoud Karbasi [a,b,*], Mehdi Jamei [a,c,d], Mumtaz Ali [e], Anurag Malik [f], Xuefeng Chu [g], Aitazaz Ahsan Farooque [a,h], Zaher Mundher Yaseen [i,j]

[a] Canadian Centre for Climate Change and Adaptation, University of Prince Edward Island, St Peters, PE, Canada
[b] Water Engineering Department, Faculty of Agriculture, University of Zanjan, Zanjan, Iran
[c] New Era and Development in Civil Engineering Research Group, Scientific Research Center, Al-Ayen University, Thi-Qar, Nasiriyah 64001, Iraq
[d] Department of Civil Engineering, Faculty of Civil Engineering and Architecture, Shahid Chamran University of Ahvaz, Ahvaz, Iran
[e] UniSQ College, University of Southern Queensland, 4305 QLD, Australia
[f] Punjab Agricultural University, Regional Research Station, Bathinda, Punjab, India
[g] Department of Civil, Construction & Environmental Engineering, North Dakota State University, Fargo, ND, USA
[h] Faculty of Sustainable Design Engineering, University of Prince Edward Island, Charlottetown, PE, Canada
[i] Civil and Environmental Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia
[j] Interdisciplinary Research Centre for Membranes and Water Security, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia

ARTICLE INFO

Handling Editor- J.E. Fernández

Keywords:
Evapotranspiration
Agriculture engineering
Deep learning
Boruta feature selection
Recurrent neural network

ABSTRACT

Evapotranspiration is one of agricultural water management's most significant and impactful hydrologic processes. A new multi-decomposition deep learning-based technique is proposed in this study to forecast weekly reference evapotranspiration (ET_o) in western coastal regions of Australia (Redcliffe and Gold Coast). The time-varying filter-based empirical mode decomposition (TVF-EMD) technique was used to first break down the original meteorological variables/signals into intrinsic mode decomposition functions (IMFs), which included maximum and minimum temperature, relative humidity, wind speed, and solar radiation. Using a partial autocorrelation function (PACF), the significant lagged values were then calculated from the decomposed subsequences (i.e., IMFs). A novel Extra Tree- Boruta feature selection algorithm was used to extract important features from the decomposed IMFs. Four machine learning approaches, including bidirectional recurrent neural network (Bi-RNN), multi-layer perceptron neural network (MLP), random forest (RF), and extreme gradient boosting (XGBoost), were used to forecast weekly evapotranspiration using the TVF-EMD-based decomposed meteorological data. Different statistical metrics were applied to evaluate the model performances. The results showed that the decomposition of the input data by TVF-EMD significantly improved the accuracy compared with the non-decomposed inputs (single models without decomposition). The findings indicate that the TVF-BiRNN model, as presented, achieved the highest level of accuracy in simulating weekly ET0 at both the Redcliffe and Gold Coast stations (Redcliffe: R=0.9281, RMSE=3.8793 mm/week, MAPE = 9.2010%; Gold Coast: R=0.8717, RMSE=4.1169 mm/week, MAPE = 11.5408%). The novel hybrid modeling technique can potentially improve agricultural water management through its ability to generate more accurate ET_o estimates weekly. The proposed methodology exhibits potential applicability to various other environmental and hydrological modeling issues.

1. Introduction

Evaporation is defined physically as the transport of liquid water from surfaces like soil and plants into the atmosphere in a gaseous form (Monteith, 1965). Several hydro-climatological factors influence the evaporation simulation; for example, solar radiation and air temperature provide the energy required for the water transformation into a gaseous state. Additionally, the biological transpiration process involves

* Corresponding author at: Canadian Centre for Climate Change and Adaptation, University of Prince Edward Island, St Peters, PE, Canada.
E-mail address: m.karbasi@znu.ac.ir (M. Karbasi).

https://doi.org/10.1016/j.agwat.2023.108604
Received 15 August 2023; Received in revised form 11 November 2023; Accepted 20 November 2023
Available online 24 November 2023

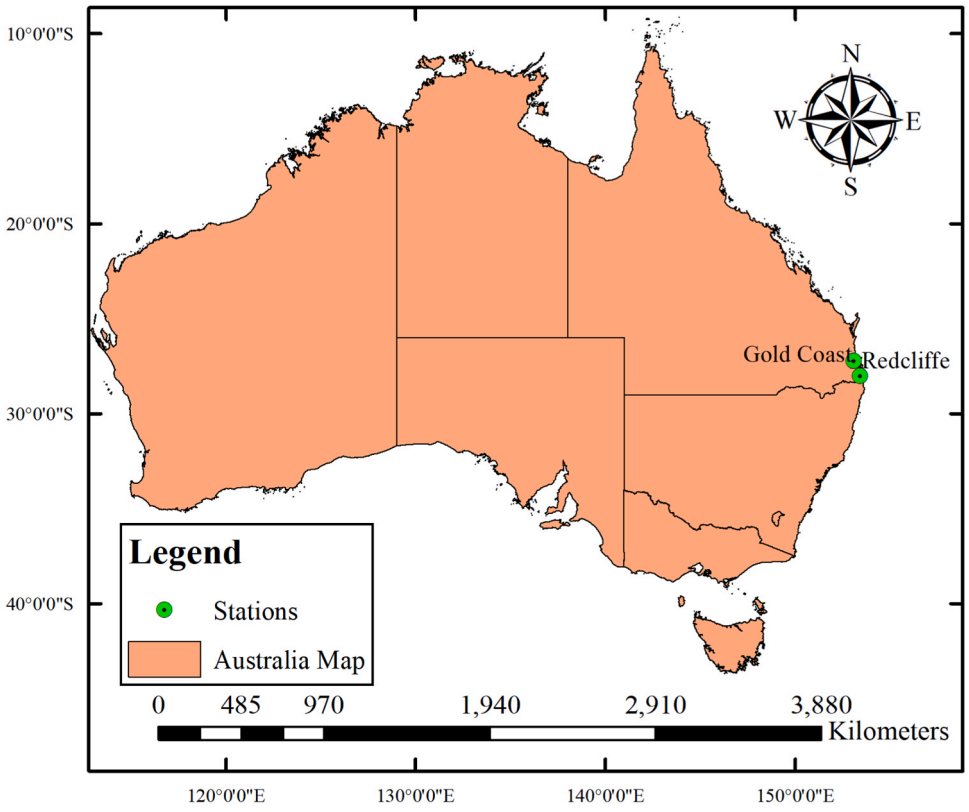


Fig. 1. Locations of Redcliffe and Gold Coast stations.

the movement of water from the plant's stomata into the atmosphere. The process of transpiration is often fueled by radiation, air humidity, air temperature, and wind (Brutsaert, 2013). Conversely, evapotranspiration (ET) is a complicated, non-linear process in which both evaporation and transpiration occur simultaneously. The amount of water lost from a cropped surface, measured in units of water depth over a predetermined time period, is known as the evapotranspiration rate (Keshtegar et al., 2022). In several domains, such as irrigation designing and planning, hydrology, crop water needs, drainage design, and water resource distribution (Torres et al., 2011), there is a need for accurate prediction of evapotranspiration (ET). Accurate prediction of evapotranspiration (ET) is crucial to agricultural water management. Reference evapotranspiration (ET_o) is often needed for this purpose. The FAO Penman-Monteith (PM) (Allen et al., 1998), Priestley-Taylor (PT) (Priestley and Taylor, 1972), and Hargreaves-Samani (HS) (Hargreaves and Samani, 1985) methods have been widely used to estimate ET_o (Lang et al., 2017). However, the main drawbacks of these methods are their variation of performance from one location to another and local calibration that limits the adequacy of the empirical method (Goyal et al., 2023). Machine learning (ML) models have been used as a dependable approach for resolving the problem of data unavailability in such areas (Yaseen, 2021). Potential ML models for simulating ET_o include evolutionary methods (Martí et al., 2015), fuzzy logic (Kisi, 2010), neural network (Nawandar et al., 2021), kernel method (Chen et al., 2022), decision tree (Ayaz et al., 2021), ensemble models (Mehr et al., 2019), random forest (Hameed et al., 2021), and hybrid ML models (Roy et al., 2021).

Deep learning (DL) models have been proven more feasible in hydrological and climatological applications in comparison with other ML models (Bhattarai et al., 2023; Fayer et al., 2023; Granata et al., 2022; Granata and Di Nunno, 2023; Kurth et al., 2018; Rasp et al., 2018; Sit et al., 2020). The excellent performance of DL models can be attributed to the superiority and reliability of their learning processes in simulating the intricate interactions between predictors and predictors (Danandeh Mehr et al., 2022; Fu et al., 2020). DL models can provide data output in a hierarchical arrangement, increasing the complexity of traditional neural network models (Bouguettaya et al., 2022). Convolutional neural networks, long short-term memory, deep neural networks, and recurrent neural networks are some of the different types of deep learning models (Sharifani and Amini, 2023). For ET_o process modeling, in particular, some research was adopted over the literature using the potential of DL models. An innovative strategy based on the hybridization of the multivariate variational mode decomposition with recurrent neural network for forecasting the daily scale of ET_o (Zheng et al., 2023). In another work, the integration of the classical Hargreaves-Samani method with long short-term memory was developed for daily records of ET_o (Yan et al., 2023). Other research on a similar concept was established for ET_o modeling using the capacity of DL models (Babaeian et al., 2022; Dong et al., 2021; Granata and Di Nunno, 2021; Wang et al., 2022; Yin et al., 2020). The existing literature reveals a few key issues that need to be resolved before the next generation of computer-aided modeling. The development of a DL model depends strongly on the time horizon (e.g., hourly, daily, weekly, monthly, or seasonally), which further affects its accuracy. The selection of appropriate input parameters to yield a better learning process using some advanced feature selection methods is yet to be tested. Hence, the identification of the appropriate optimization for the feature to be supplied for the prediction matrix is important. The current study aims to fill these gaps.

This study uses a special deep learning method, a bi-directional recurrent neural network (Bi-RNN), to predict weekly ET_o in combination with time-variable filter-based empirical mode decomposition (TVF-EMD). Redcliffe and Gold Coast stations in Australia were designated as test locations for the new modeling technique. The performance of the new model will be evaluated through the use of three ML algorithms, namely multi-layer perceptron neural network (MLP), random forest (RF), and extreme gradient boosting (XGBoost). Therefore, the primary goals of this study are: 1) use the Bi-RNN model to forecast the weekly ET_o, 2) generate a weekly ET_o forecast using a combination of

Table 1
Descriptive statistics of weekly meteorological variables for Redcliffe and Gold Coast stations.

Redcliffe						
Variable	Tmax (C°)	Tmin (C°)	RH (%)	Wind Speed (m/s)	Solar Radiation (MJ/sq m)	ET_o (mm/ week)
Mean	25.52	17.18	67.88	5.39	19.16	32.39
StD	3.09	4.22	6.71	1.36	5.47	10.36
Max	32.59	24.61	87.86	9.66	34.35	59.00
Min	18.07	7.14	40.14	2.16	8.08	14.10
Q1	22.96	13.46	63.86	4.23	14.55	23.30
Q2	25.84	17.46	67.71	5.46	18.53	32.00
Q3	28.07	21.04	72.14	6.40	23.51	40.40
Skewness	-0.14	-0.17	-0.22	0.17	0.30	0.22
Kurtosis	-0.95	-1.18	0.56	-0.48	-0.85	-0.94
Gold Coast						
Variable	Tmax (C°)	Tmin (C°)	RH (%)	Wind Speed (m/s)	Solar Radiation (MJ/sq m)	ET_o (mm/ week)
Mean	25.77	17.37	72.19	5.27	19.09	30.52
StD	2.92	3.75	8.08	1.25	5.53	8.70
Max	30.87	24.14	94.29	10.02	34.58	54.60
Min	18.11	8.41	47.93	1.52	8.14	8.90
Q1	23.19	14.17	67.21	4.30	14.33	23.10
Q2	26.04	17.60	72.43	5.12	18.44	29.80
Q3	28.26	20.77	77.71	6.05	23.65	37.10
Skewness	-0.18	-0.21	-0.13	0.65	0.29	0.17
Kurtosis	-1.13	-1.10	-0.09	0.56	-0.89	-0.73

StD: Standard Deviation; Q1: Quartile 1(25%), Q2: Quartile 2 (50%), Q3: Quartile 3 (75%); Skewness and Kurtosis values between − 1,+ 1, show variable distribution is near Normal

TVF-EMD and ML models, 3) use a unique Extra Tree-Boruta feature selection process to identify relevant intrinsic mode decomposition functions (IMFs) created from TVF-EMD in the forecasting procedure, 4) analyze the effectiveness of the models and 5) examine the differences between single and TVF-EMD-based models.

2. Materials and methods

2.1. Study area and data description

Fig. 1 shows the locations of Redcliffe and Gold Coast stations. Redcliffe is a town and suburb located in the Moreton Bay Region, Queensland, Australia. Its climate is warm and temperate in Redcliffe. It experiences heavy rains, even in the driest month. This climate falls under the Köppen and Geiger classification of Cfa (Humid subtropical climate). The average temperature in this area is 20.4 °C, and the annual precipitation is 1001 mm. Gold Coast is a city situated on the coast of Queensland, Australia, located to the north of the New South Wales border. Its climate is similar to that in Redcliffe. Its average temperature is 20.4 °C, and the annual precipitation is 1138 mm. The daily meteorological data, including minimum temperature (T_{min}), maximum temperature (T_{max}), wind speed, relative humidity (RH%), solar radiation, and reference evapotranspiration (ET_o), were obtained from the Australia Bureau of Meteorology (http://www.bom.gov.au/). These daily data were then transformed into a weekly scale. Table 1 shows the descriptive statistics of these basic weekly meteorological variables and ET_o for Redcliffe and Gold Coast stations. According to Table 1, the mean weekly ET_o values are 32.39 and 30.52 mm for Redcliffe and Gold Coast stations, respectively.

2.2. Data quality and integrity assessment

Typically, the evaluation of data quality and integrity is conducted through the use of the double-mass curve technique. This technique necessitates the availability of a record spanning over 30 years from numerous locations. However, in our specific scenario, we possess a dataset encompassing a duration of ten years (2012–2021) from two locations. Consequently, a straightforward examination can be performed by employing basic statistical analysis measures, such as minimum, maximum, mean, standard deviation, skewness, and kurtosis. Table 1 presents an overview of the statistical features of the weekly climatological data. We do the following analysis to ensure the data quality and validity:

1. Examining time series plots visually can help identify anomalies and incorrect spikes or declines in the data. Values that seemed suspect were compared to the original data sources.
2. Range checks are used to find values that go above predicted physical extremes and bounds. Erroneous data was confirmed or eliminated as needed.
3. Use composite estimates from nearby stations or interpolation to fill in the gaps left by missing data.
4. Consistency checks between related variables (e.g., temperature and humidity) to flag illogical relationships. Suspect data points were investigated.

2.3. PM equation

The FAO Penman-Monteith (PM) equation (Allen et al., 1998) has been widely used for the computation of reference evapotranspiration. The FAO PM equation was utilized in the present study to estimate weekly ET_o (Allen et al., 1998):

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} U_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34U_2)} \quad (1)$$

where ET_o = reference evapotranspiration, mm week^{-1}; R_n = net radiation, MJ m^{-2} week^{-1}; G = soil heat flux density, MJ m^{-2} week^{-1}; Δ = slope of saturation vapor pressure, kPa °C^{-1}, T = mean air temperature, °C; U_2 = wind speed at 2-m height above the ground surface, m s^{-1}; γ = psychrometric constant, kPa °C^{-1}; and $e_s - e_a$ = saturation and actual vapor pressures, kPa. Weekly ET_o values are obtained by summing seven days of ET_o, and time series are built using weekly ET_o.

2.4. Boruta-extra tree for feature selection

In modeling complex hydrological and environmental processes, nominating optimal inputs is crucial to better understand their impacts on the model output. A variety of methods are available for handling this issue. This study utilized the Boruta – Extra Tree (BET) algorithm for significant feature selection (Kursa et al., 2010) in weekly ET_o forecasting at the study sites. The working principles of the Boruta algorithm are optimized using the extra tree (Geurts et al., 2006) model for selecting the significant inputs and omitting the non-significant features. The extra tree is an ensemble machine learning model based on many united decision trees to perform regression or classification tasks (Geurts et al., 2006). The two central elements of the extra tree are the number of features randomly chosen for each node and the least sample size for partitioning a node, which enhances the model's forecast accuracy (Asadollah et al., 2020). The Boruta algorithm is applied to determine the Z-scores of each input predictor in reference to the shadow property. Finally, the Z-score metrics for each iteration are computed and repeatedly eliminate the features supposed to be very irrelevant based on the corresponding shadow attributes (when Z-scores < maximum Z-score among shadow attributes = unimportant inputs, and Z-scores > maximum Z-score among shadow attributes = important or effective inputs). The process ends when all features are confirmed or when the required iteration threshold is reached (Ahmed et al., 2021).

2.5. TVF-EMD

In 2017, Li et al. suggested a time-varying filter-based empirical

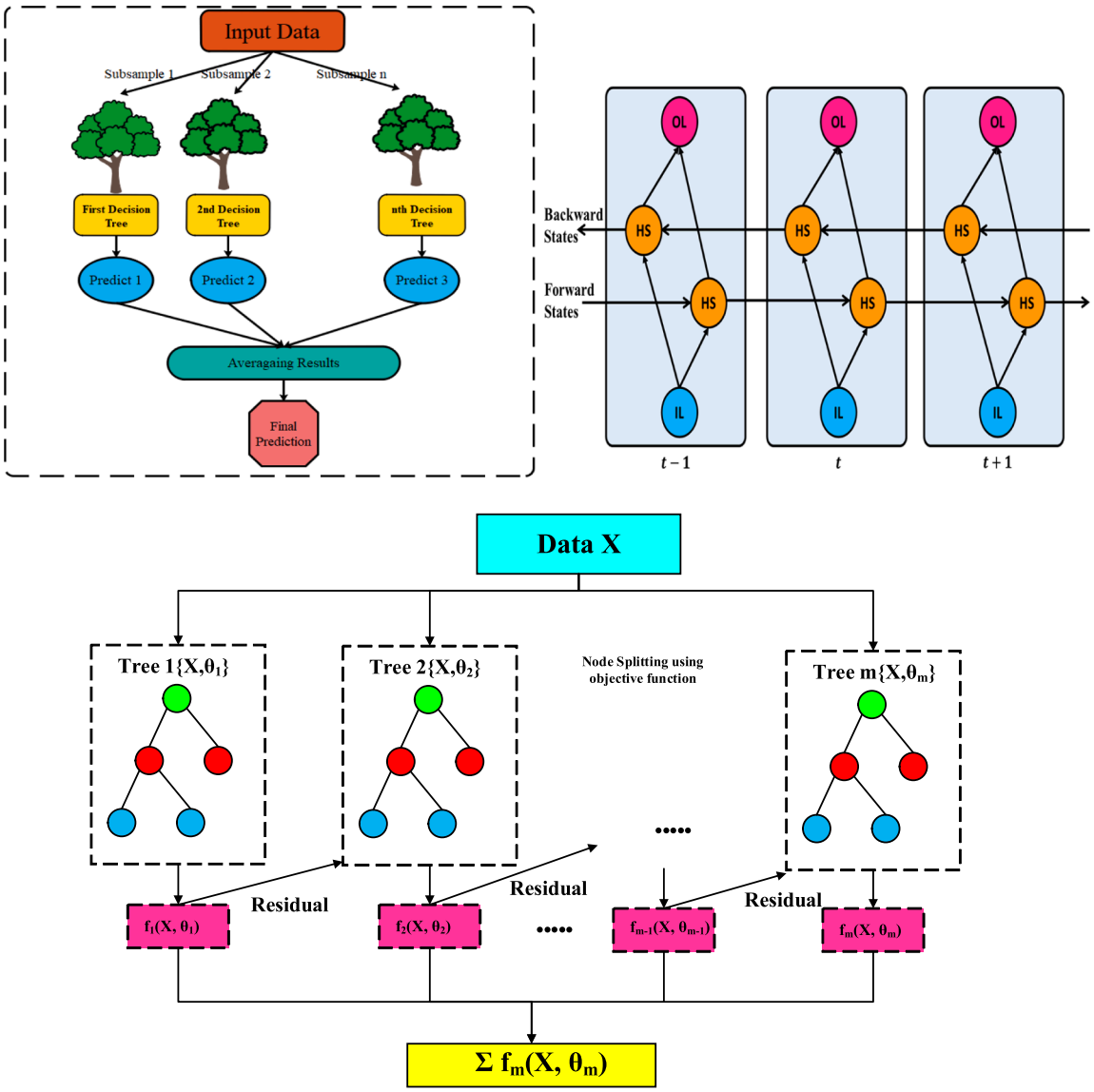


Fig. 2. Architecture of the Random Forest (upper left), BRNN model showing unfolded in time for three time steps (IL: input layer, HS: hidden state, and OL: output layer) (upper right), and XGBoost (Lower).

mode decomposition (TVF-EMD) model to control the consequences of the end effect and mode blending inside EMD and finish the shifting process using a time-varying filter. The performance of the TVF-EMD model is contingent on the correct selection of the bandwidth threshold (which has a direct influence on the separation and intermittence performance) and the B-spline order (which affects the filtering performance of the TVF) (Li et al., 2017). The TVF-EMD model fails to optimize the mode mixing problem. Similarly, accomplishing TVF needs to know the local cut-off frequency (LCOF) (Zhang et al., 2018). The TVF-EMD model has been observed to yield more practical outcomes when compared to other methods such as EWT (Empirical Wavelet Transforms), MEMD (Multivariate Empirical Mode Decomposition), EEMD (Ensemble Empirical Mode Decomposition), and VMD (Variational Mode Decomposition) (Wang et al., 2020; Zhang et al., 2018). The implementation of the TVF-EMD involves the following steps (Li et al., 2017; Song et al., 2021):

Step-1: Compute the LCOF by employing a B-spline approximation, which can be expressed as:

$$g_m^n(t) = \sum_{k=-\infty}^{\infty} c(k)\beta^n(t/m-k) \tag{2}$$

where, $\beta^n(t)$ represents the B-spline function and $c(k)$ is the B-spline coefficient. Eq. (2) indicates that the B-spline function is enlarged by the factor of m. The approximation is determined by n, m, and $c(k)$. Given B-spline order n and knots m, the B-spline approximation is governed by $c(k)$, which reduces the approximation error (ε_m^2) as follows:

$$\varepsilon_m^2 = \sum_{t=-\infty}^{+\infty} \left(x(t) - [c]_{\uparrow m} * b_m^n(t)\right)^2 \tag{3}$$

in which, $b_m^n(t)$ signifies $\beta^n(t/m)$, $[\bullet]_{\uparrow m}$ denotes up-sampling operation by m, and $*$ designates convolution operator. $c(k)$ is given by:

$$c(k) \quad = \left[p_m^n * x\right]_{\downarrow m}(k) \tag{4}$$

where, $[\bullet]_{\downarrow m}$ indicates down-sampling operation by m and p_m^n characterizes pre-filter. Thus, Eq. (2) can be rewritten as:

$$g_m^n(t) = \left[p_m^n * x\right]_{\downarrow m} * b_m^n(t) \tag{5}$$

Eq. (5) shows that the B-spline approximation is a type of low pass filtering. As a result, the LCOF is obtained from the input signal and used to construct the TVF. This process is carried out to obtain the LCOF,

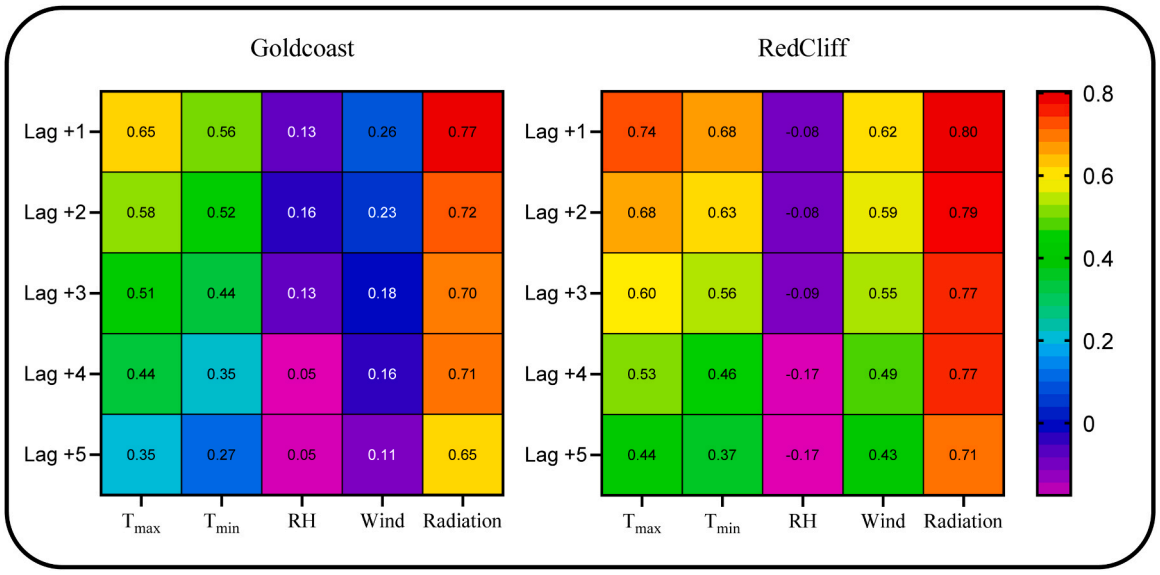


Fig. 3. Cross-correlation analysis for Gold Coast and Redcliffe stations.

Table 2
TVF EMD decomposition setting for input time series.

Station	Time Series	Number of Decomposed IMFs	B-spline Order	End_flag	Stopping criterion
Redcliffe	T_{max}	11	26	0	0.1
	T_{min}	14	26	0	0.1
	Wind Speed	13	26	0	0.1
	Solar Radiation	16	26	0	0.1
Gold Coast	T_{max}	11	26	0	0.1
	T_{min}	12	26	0	0.1
	Wind Speed	16	26	0	0.1
	Solar Radiation	12	26	0	0.1

$\varphi'_{bis}(t) = \varphi'_1(t) + \varphi'_2(t)/2$. Here, $\varphi'_1(t)$, and $\varphi'_2(t)$ are slow varying components. Realign $\varphi'_{bis}(t)$ to solve the issue of intermittence (i.e., noise) and obtain the final LCOF by interpolating among the peaks (or remainders).

Step-2: Achieve the local mean by filtering the input signal using a time-varying filter (i.e., B-spline approximation filter).

Step-3: Improve the stopping criterion by checking the residual signal as follows (Wang et al., 2020):

$$\theta(t) = \frac{B_{Loughlin}(t)}{\varphi_{avg}(t)} \tag{6}$$

in which, $B_{Loughlin}(t)$ and $\varphi_{avg}(t)$ are the Loughlin instantaneous bandwidth and the weighted average instantaneous frequency of separate components, respectively. For more information on the TVF-EMD model, readers can refer to Li et al. (2017).

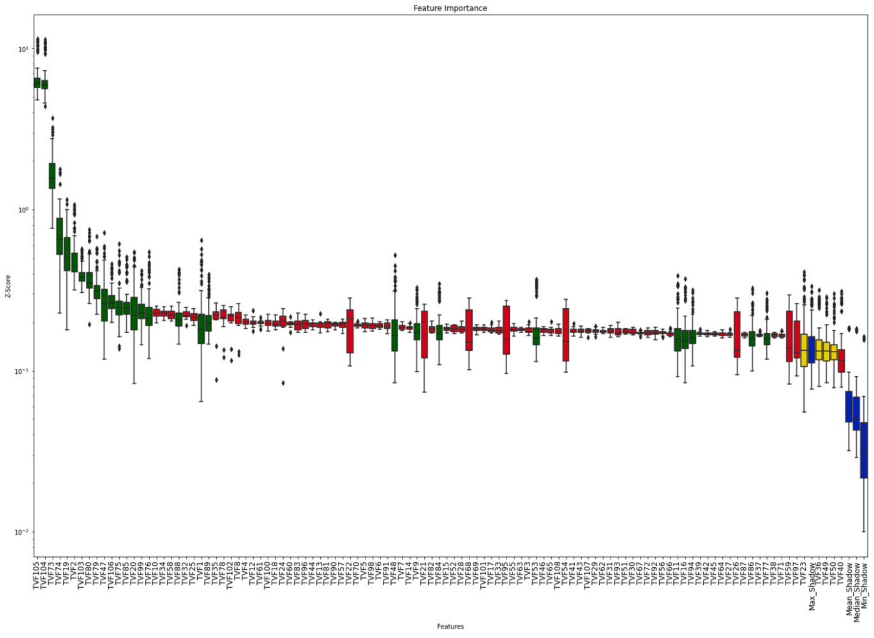


Fig. 4. Feature selecting process using the Boruta-Extra Tree method for Redcliffe station.

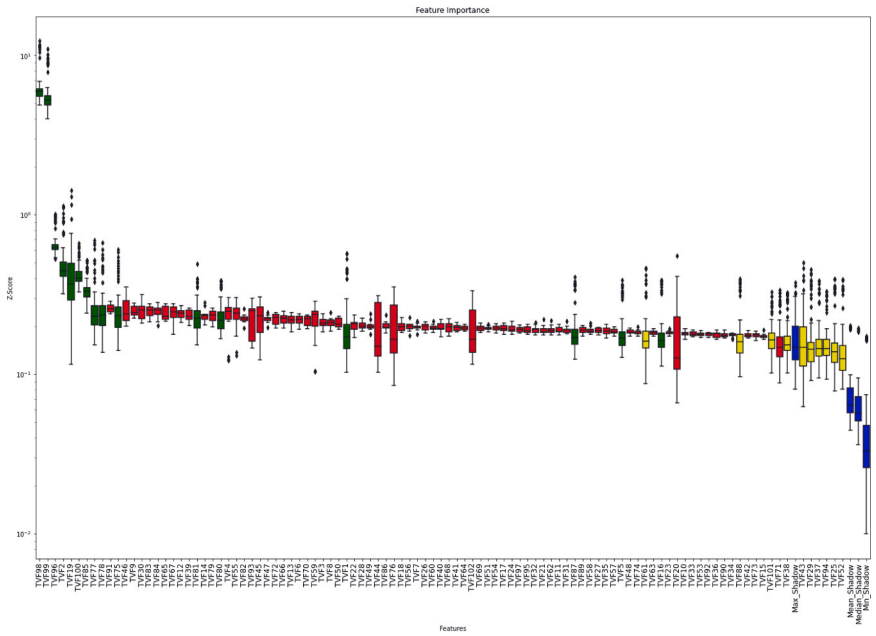


Fig. 5. Feature selecting process using the Boruta-Extra Tree method for Gold Coast station.

Table 3
Optimized model adjustment for the forecasting of ET_o.

Study site	**Models**	**Best parameters**
Redcliffe single models	Random Forest	N_Estimators: 20, Max-Depth:5
	XGBoost	Learn rate:0.07, Max-Depth:4, N_Estimators: 70
	Bi-RNN	Layers: 2, Number of Neurons: 180, Training Algorithm: Adam, Dropout: 0, Learning Rate: 0.000101, Epochs: 55, Batch Size: 64
	MLP	Training Algorithm: Levenberg Marquadt, Layers: 1, Number of Neuron: 2
Redcliffe TVF based models	Random Forest	N_Estimators: 300, Max-Depth:20
	XGBoost	Learn rate:0.17, Max-Depth:20, N_Estimators: 40
	Bi-RNN	Layers: 2, Number of Neurons: 200, Training Algorithm: Adam, Dropout: 0; Learning Rate: 0.0011, Epochs: 70, Batch Size: 64
	MLP	Training Algorithm: Levenberg Marquadt, Layers: 1, Number of Neuron: 5
Study site	**Models**	**Best parameters**
Gold Coast single models	Random Forest	N_Estimators: 20, Max-Depth: 3
	XGBoost	Learn rate:0.05, Max-Depth:4, N_Estimators: 100
	Bi-RNN	Layers: 2, Number of Neurons: 110, Training Algorithm: Adam, Dropout: 0, Training Algorithm: Adam; Learning Rate: 0.017, Epochs: 58, Batch Size: 92
	MLP	Training Algorithm: Levenberg Marquadt, Layers: 1, Number of Neuron: 3
Gold Coast TVF based models	Random Forest	N_Estimators: 200, Max-Depth:15
	XGBoost	Learn rate:0.07, Max-Depth:4, N_Estimators: 100
	Bi-RNN	Layers: 2, Number of Neurons: 130, Training Algorithm: Adam, Dropout: 0; Learning Rate: 0.0141, Epochs: 44, Batch Size: 64
	MLP	Training Algorithm: Levenberg Marquadt, Layers: 1, Number of Neuron: 4

2.6. Bidirectional recurrent neural networks

Deep-learning-based algorithms play an important role in handling multi-dimensional problems. Similarly, for analyzing the sequential data, the bidirectional recurrent neural networks (BRNN) can be a good choice (Schuster and Paliwal, 1997). A BRNN model can be trained using the entire available input sequences (information) in the past to predict the output. Moreover, BRNN can be effective in climatic and environmental forecasting (Jaihuni et al., 2022). The state of neurons in BRNN is divided into two directions: forward (positive direction) and backward (negative direction) states. The inputs of the backward state are not connected to the outputs from the forward state, and vice versa (Schuster and Paliwal, 1997). Fig. 2 shows the general structure of the BRNN model. In hidden neurons, BRNN has repetitive loops, which allow storing the previous input information for a moment so that future outputs can be predicted. The output of the hidden layer is retransmitted t times to the hidden layer. Once the number of epochs is finished, the output of the recursive neuron is sent to the next layer. In such a way, the output becomes more inclusive, and the preceding data are reserved for longer. Finally, the errors are reimbursed backward to update the weights (Apaydin et al., 2020).

2.7. Multi-layer perception neural network (MLP NN)

ANNs are effective machine learning tools to identify complicated non-linear correlations between inputs and outputs (Zhang et al., 1998). Multi-layer perceptron (MLP) is one of the widely used forms of neural networks. It can approximate any input/output map. MLPs are feed-forward multi-layer networks with one or more hidden layers that are trained by static back-propagation (Pashazadeh and Javan, 2020). The neural network utilized in this study has three layers: an input layer, a hidden layer, and an output layer. Each neuron in a particular layer is completely or partly interconnected with a large number of other neurons via weighted connections. The scalar weights dictate the strength of the connections between neurons that are linked. A weight of zero indicates that there is no link between two neurons, whereas a negative weight indicates a prohibitive association. MLP can be mathematically expressed as:

$$y = f\left(\sum_{i=1}^{n} w_i p_i + b\right) \tag{7}$$

where y is the target value, f denotes the activation function, p_i is the input vector, w_i is the weight vector, and b is the bias. A tangent sigmoid

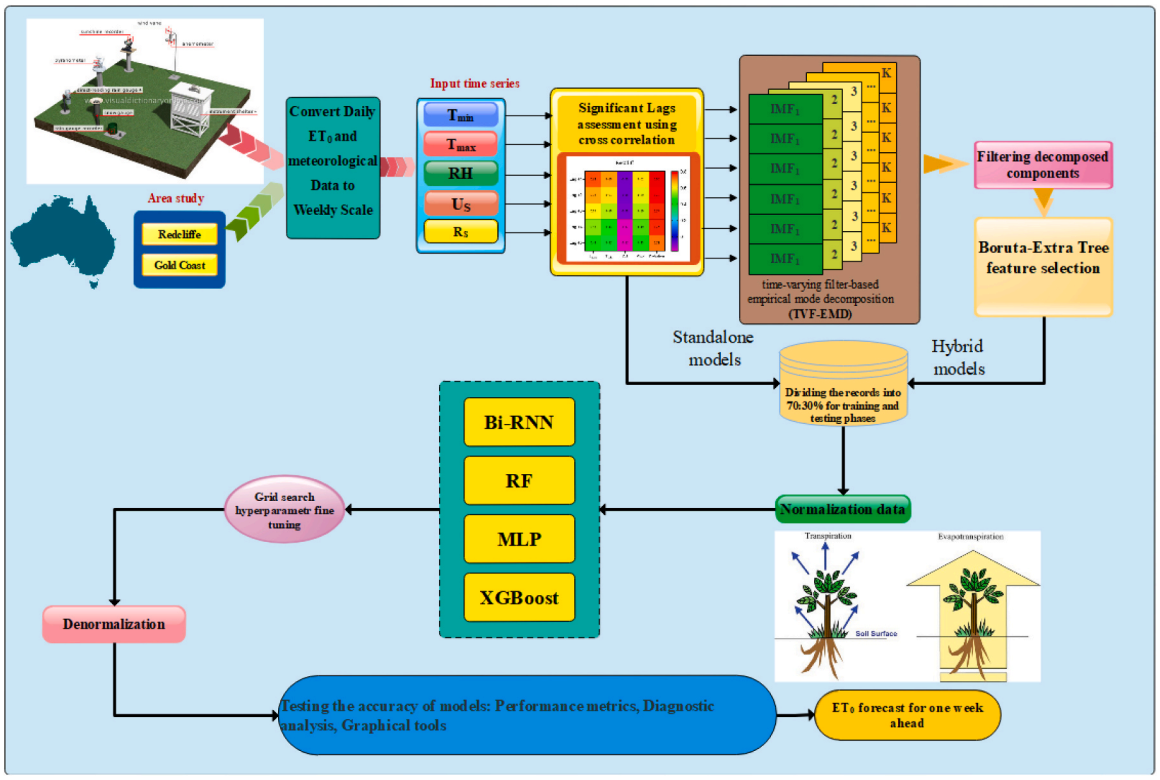


Fig. 6. Modeling flowchart.

Table 4
Modeling results for Redcliffe station (the best model is highlighted in bold blue).

Model	Data	R	RMSE	MAPE	NSE	I_A	$U_{95\%}$
TVF-BiRNN	**Train**	**0.9776**	**2.4700**	**6.3620**	**0.9435**	**0.9854**	**6.4702**
	Test	**0.9281**	**3.8793**	**9.2010**	**0.8588**	**0.9621**	**10.7304**
TVF-MLP	Train	0.9591	2.9520	7.4098	0.9191	0.9783	8.1719
	Test	0.9120	4.5957	12.1399	0.8018	0.9442	12.2654
TVF-RF	Train	0.9915	1.4048	3.4193	0.9817	0.9952	3.8961
	Test	0.8908	4.7927	11.7583	0.7845	0.9365	13.1700
TVF-XGBoost	Train	0.9967	0.8586	1.7748	0.9932	0.9983	2.3770
	Test	0.9055	4.5654	11.5553	0.8044	0.9437	12.4234
BiRNN	Train	0.9594	3.1649	7.9804	0.9072	0.9728	8.6916
	Test	0.8458	5.5086	12.8283	0.7153	0.9109	15.2910
MLP	Train	0.9057	4.4078	10.9044	0.8196	0.9490	12.2157
	Test	0.8342	5.7504	14.0130	0.6897	0.9053	15.8916
RF	Train	0.9439	3.4658	8.7338	0.8887	0.9684	9.6125
	Test	0.8300	5.8277	14.3701	0.6813	0.9029	16.0976
XGBoost	Train	0.9425	3.5158	8.7398	0.8855	0.9676	9.7307
	Test	0.8303	5.7806	13.7432	0.6864	0.9042	16.0208

activation function is employed in this study. For any variable s, it is defined as:

$$f(s) = \frac{2}{(1+e^{-2s})} - 1 \tag{8}$$

All datasets were normalized and split into testing and training data categories in this investigation. Previous research (Kişi, 2007; Rezaeianzadeh et al., 2014) has demonstrated that the Levenberg–Marquardt method yielded acceptable results for the majority of ANN applications. Hence, it was chosen for this investigation.

2.8. Random forest

An ensemble of regression (or classification) tree models is used in the RF algorithm (Breiman, 1999). Specifically, random subsamples of the original data are used to construct a succession of individual trees. Each subsample yields a decision tree, which is then used to predict the target variable. Ultimately, an ensemble average of all different trees is calculated (Merufinia et al., 2023). The procedure executes random binary trees that utilize a subset of the data over the bootstrapping method. A random selection of the training data is chosen and used to generate the model. The data that are not included are referred to as "out of bag" (OOB) (Naghibi et al., 2017). Each tree in the RF algorithm generates a prediction result. The final prediction value is determined by averaging the results from all individual trees. The prediction error is given by (Genuer et al., 2010):

$$MSE_{OOB} = \frac{\sum_{i=1}^{n_{tree}} \left(z_i - \hat{z}_i^{OOB}\right)^2}{n_{tree}} \tag{9}$$

where MSE_{OOB} is the mean squared error of the *OOB* data predictions,

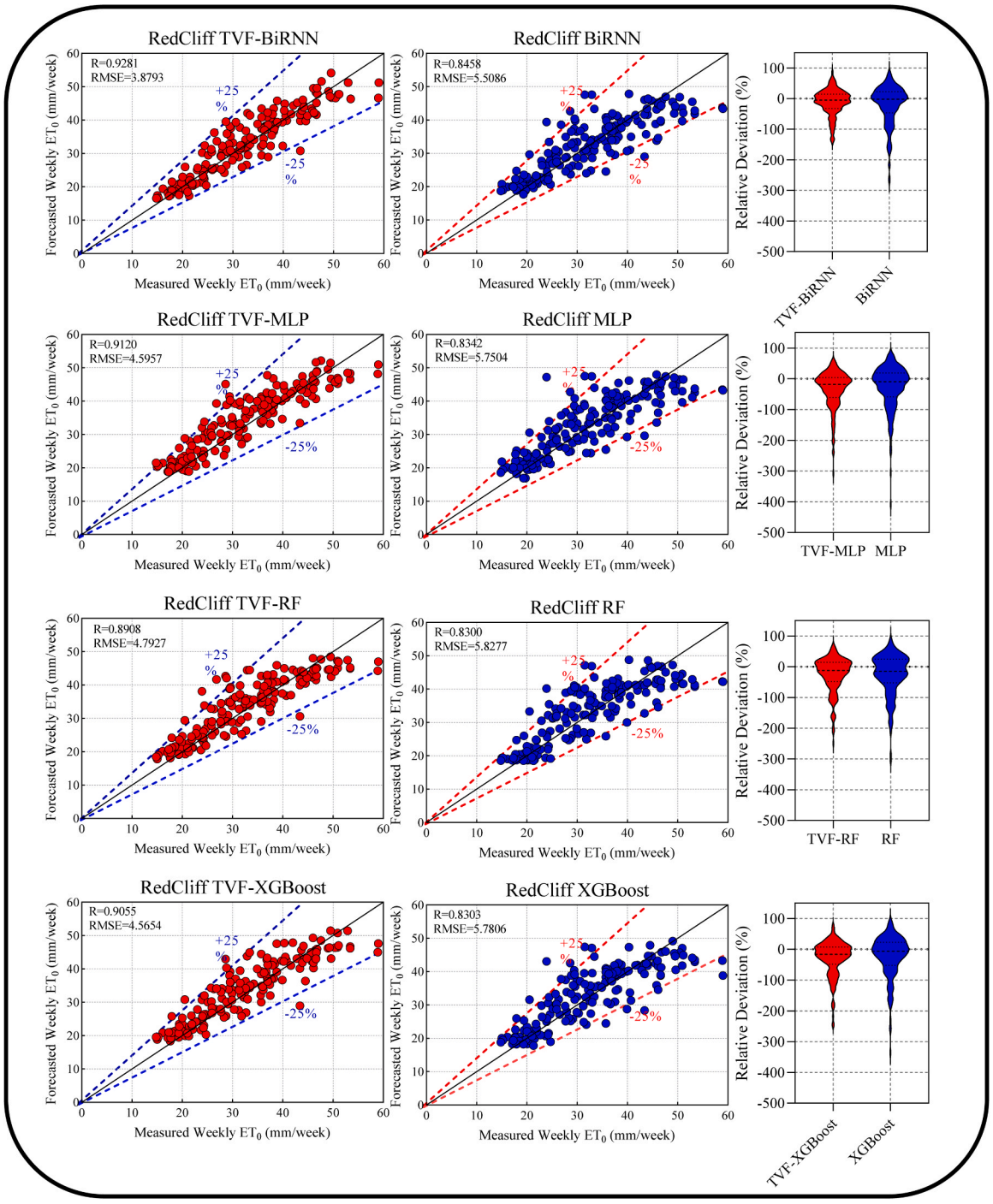


Fig. 7. Scatter plots of weekly ET_o simulated by single and TVF-EMD-based models for Redcliffe station.

n_{tree} denotes the number of trees, and z_i and $\hat{z}_i^{OOB}$ are the mean of all *OOB* predictions and the actual *OOB* value, respectively. The primary benefits of the random forest include: (1) RF can assess continuous and categorical variables with missing values for the original data. It is resilient to noise, and hence preprocessing of the data is unnecessary. (2) RF reduces computing time for training operations since each tree may be constructed in parallel, and pruning is not required. (3) For model performance, RF always converges. Thus, overfitting is never an issue, and errors are unbiased estimates of the unused bootstrap subset (Breiman, 1999). RF needs tuning of two parameters: the number of trees (n_{tree}) and the maximum depth of the trees (max_depth) (Genuer et al., 2010). The structure of the RF model is depicted in Fig. 2.

2.9. Extreme gradient boosting (XGBoost)

XGBoost (Chen and Guestrin, 2016) is one of the emerging ensemble machine learning approaches. It is an improved version of the gradient boosting decision obtained by boosting the gradient boosting decision tree (GBDT). The most important advantages are its flexibility, efficiency, and accuracy in solving non-linear regression and classification problems. This method is less exposed to overfitting due to the application of a regulator in the loss function (Bhagat et al., 2020). The main strategy in this method is to use all the samples in the training phase so that the weight changes of each sample are applied at each stage of training (Sharafati et al., 2020). Specifically, the error of the previous

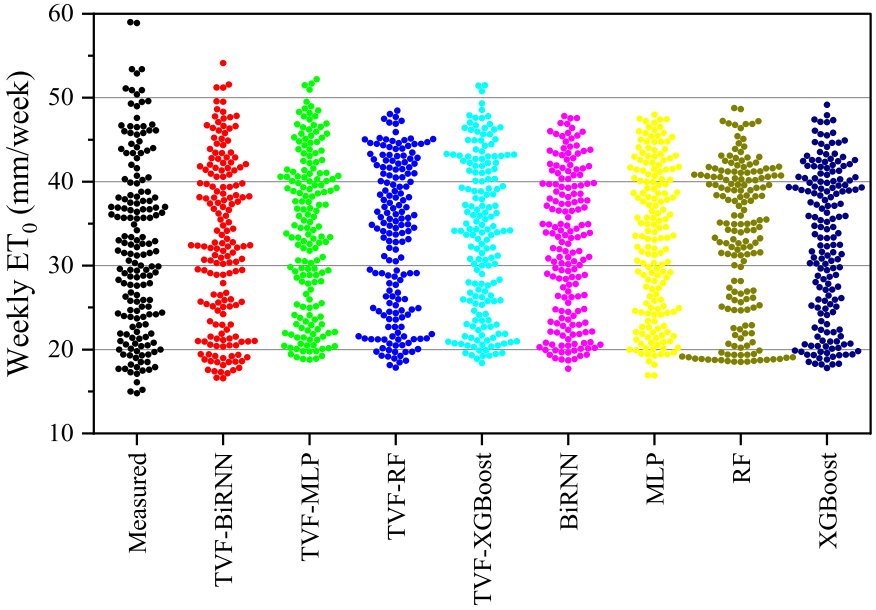


Fig. 8. Swarm plots of weekly ET_o – Comparison of the measured data and simulations from eight different models for Redcliffe station.

Table 5
Modeling results for Gold Coast station.

Model	Data	R	RMSE	MAPE	NSE	I_A	$U_{95\%}$
TVF-BiRNN	**Train**	**0.8978**	**4.0361**	**11.0330**	**0.7912**	**0.9357**	**11.0438**
	Test	**0.8717**	**4.1169**	**11.5408**	**0.7580**	**0.9243**	**11.4256**
TVF-MLP	Train	0.9054	3.7571	10.5862	0.8186	0.9466	10.4125
	Test	0.8645	4.3557	12.3673	0.7291	0.9201	11.8898
TVF-RF	Train	0.9851	1.6065	4.3743	0.9669	0.9910	4.4563
	Test	0.8594	4.4085	12.6664	0.7225	0.9116	12.0863
TVF-XGBoost	Train	0.9601	2.4861	6.7604	0.9208	0.9784	6.8939
	Test	0.8565	4.3209	11.8024	0.7334	0.9187	11.9924
BiRNN	Train	0.8535	4.7289	12.5443	0.7134	0.9107	12.9476
	Test	0.7951	5.1113	13.6822	0.6270	0.8847	14.1815
MLP	Train	0.8355	4.8911	13.8681	0.6925	0.8961	13.5559
	Test	0.7930	5.3358	14.8995	0.5935	0.8778	14.5446
RF	Train	0.8588	4.5386	12.5865	0.7360	0.9160	12.5892
	Test	0.7699	5.4686	14.8003	0.5730	0.8659	15.0623
XGBoost	Train	0.9663	2.5243	6.5829	0.9183	0.9761	6.9047
	Test	0.7620	5.4263	14.3437	0.5796	0.8592	15.0621

round is fitted by considering the criterion of a small threshold error value or a predetermined number of iterations for each stage of training. Afterward, the final prediction is obtained by using the weighted average of the prediction results (Jamei et al., 2022c). For this purpose, if $D = \{(x_i, y_i)\}_{i=1:N}$ is being the training sample including n samples and m features, the prediction after the t^{th} iteration can be defined as (Singh et al., 2022):

$$f^t(x) = f^{t-1}(x) + f_t(x) = \sum_{k=1}^{K} f_k, f_k \in F \tag{10}$$

where x is a multi-dimensional vector, $f_t(x)$ denotes the t^{th} increment, F is the regression trees space, k is the base model related to the additional model, and K denotes the tree number implemented to ensemble outcomes. A heuristic approach is always used in the non-deterministic polynomial process of model learning to identify the best answer. The objection function of XGBoost composed of the traditional loss function (L) and regularization function (Ψ) can be defined as (Li et al., 2019):

$$\text{Objective} \quad = \sum_{i=1}^{m} L(y_i, \hat{y}_i) + \sum_{k=1}^{t} \Psi(g_k) \tag{11}$$

$$\Psi(f) = \gamma T + \frac{1}{2}\lambda \parallel \psi \parallel^2 \tag{12}$$

where m denotes the total data; i is the sample number of datasets; m denotes the overall volume of data fed into the k^{th} tree; and y_i and $\hat{y}_i$ denote the observed and forecasted values, respectively. In Eq. (12) for the function of $\Psi(f)$, T represents the number of leaf nodes; γ is the minimum loss to split the leaf nodes, ψ represents the fraction of leaf nodes, and λ is the tuning parameter (Li et al., 2019). The structure of XGBoost is depicted in Fig. 2.

2.10. Model development

This research focused on developing a new expert framework comprised of an advanced preprocessing stage coupled with the Bi-RNN deep learning scheme to forecast weekly evapotranspiration for Redcliffe and Gold Coast in Australia over a period of ten years (from January 2012 to December 2021). The Boruta-extra tree combined with the TVF-EMD decomposition technique determines the significant predator's lags (antecedent information) and decomposes the non-stationary signals into sub-sequences of intrinsic mode functions

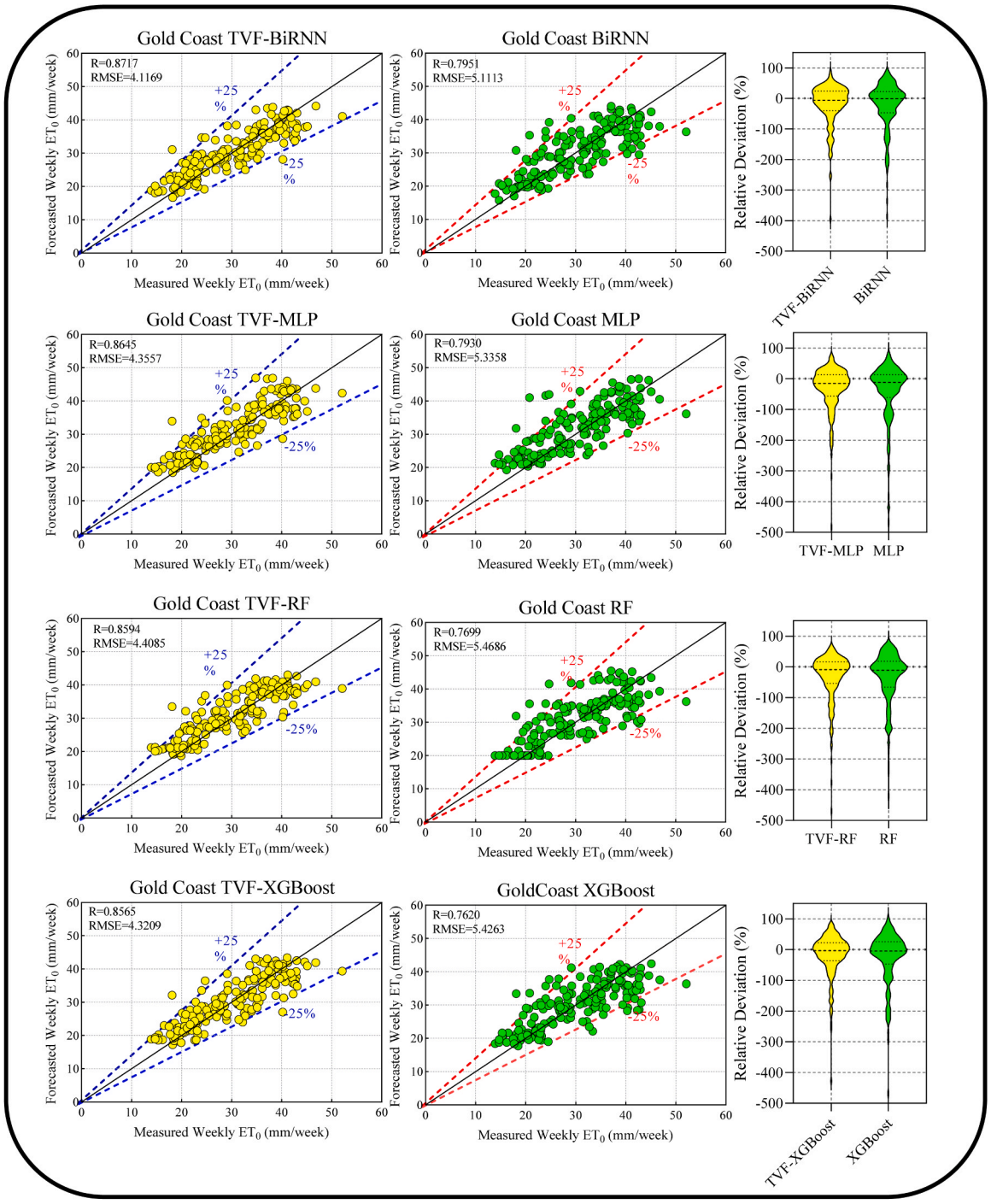


Fig. 9. Scatter plots of weekly ET_o simulated by single and TVF-EMD based models for Gold Coast station.

(IMFs). The predictor signals are Tmax (C°), Tmin (C°), RH (%), wind speed (m/s), and solar radiation (MJ/m^2). In addition, XGBoost, RF, and MLP are employed to validate the main model in both hybrid and standalone frameworks. TVF-BiRNN, TVF-MLP, TVF-RF, and TVF-XGBoost, along with their standalone counterparts, are applied to assess the potential of forecasting ET_o at the two stations. The TVF-EMD decomposition technique and MLP are performed in MATLAB R2019b environment, whereas the Boruta-extra tree, BiRNN, XGBoost, and RF are executed based on Scikit-learn, Tensorflow, and XGBoost libraries in the Python platform. The modeling steps are detailed as follows:

Step 1: determination of significant lag time of predictors.

In the model development, the significant time-lags associated with the predictors (i.e., Tmax, Tmin, RH, wind speed, and solar radiation) need to be determined. In this regard, the correlation between the antecedent time-lagged predictors and ET_o is computed for the first five lags using the cross-correlation function as depicted in Fig. 3. As shown in Fig. 3, the most influential antecedent components are the first two time lags (Lags 1 and 2).

Step 2: decomposition of the predictor's signal by applying TVF-EMD.

The performance of the multi-predictor time series-based models is totally dependent on the information extracted from the signals. Thus, data preprocessing is another stage that is extremely significant in enhancing the model's accuracy. Since the predictor signals are non-

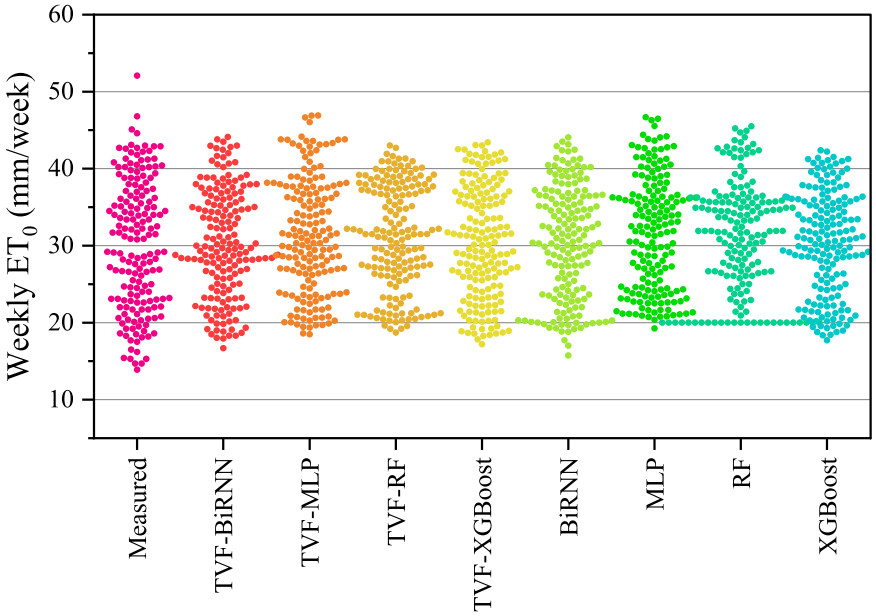


Fig. 10. Swarm plots of weekly ET_o – Comparison of the measured data and simulations from eight different models for Gold Coast station.

stationary and highly noisy, feeding the ML model using them without decomposition will yield a loss of accuracy and efficiency. TVD-EMD is one of the most efficient advanced extensions of the EMD algorithm. A time-varying filter decomposition technique inside the sifting procedure solves the defect of mode mixing in the classical EMD (Li et al., 2017). Therefore, it is resistant to noise interference, has higher stability even when sampling rates are low, and can increase frequency separation performance (Zhou et al., 2021). In this study, the TVF-EMD decomposition was carried out to decompose the IMFs and a residual component in an individual fashion for each forecasting horizon and sum up all the forecasting results for the two stations. The trial and error procedure was employed to gain the optimal setting of TVD-EMD and the results for Redcliffe and Gold Coast stations are listed in Table 2. It should be noted that the number of IMFs in each of the predictors varies between 10 and 15 in order to achieve the best accuracy. Thus, the total number of sub-sequences attained using the TVD-EMD for Redcliffe and Gold Coast stations are 108 and 102, respectively.

Step 3: significant sub-sequences filtering and feeding preparation.

The selection of features is a crucial stage for the ML-based models due to their capabilities in reducing the computational complexity and input dimensionality and improving accuracy (Bennasar et al., 2015). For these reasons, the Boruta extra tree feature selection was adopted to filter the redundant sub-sequences based on the max shadow benchmark. Fig. 4 illustrates the outcomes of the filtering procedure for Redcliffe and Gold Coast stations. In the box plot, the green color represents the components passed through the feature selection criteria. It is clear that Boruta-extra-tree efficiently removed 73% and 88% of the total sub-sequences pool and reduced the computational costs. The next step is feeding the machine learning models. In general, there is no specific standard for dividing all available datasets into training and testing subsets. In hydrological applications, 60–80% of the data are generally allocated for training and the remaining 20–40% of the data for testing (Jamei et al., 2022a). In the current research, to develop the AI-based models for ET_o forecasting, 70% and 30% of the entire dataset are respectively used for training and testing phases. Essentially, the difference in the input data scale leads variables with a greater size to inhibit the influence of the data with a smaller scale. To overcome this problem, normalization of the data between 0 and 1 is performed for the training period (Jamei et al., 2022b). One of the goals of this research is to express the robustness of the hybrid models compared to the corresponding individual models in ET_o forecasting. For this purpose, single ML models (i.e., BiRNN, MLP, RF, and XGBoost) without decomposition preprocessing are also employed to forecast ET_o for both stations..

Step 4: Optimal tuning of the hyperparameters of machine learning approaches.

Even with the best approaches for non-linear modeling situations, promising outcomes may not be attained if the hyperparameters are not effectively set. Choosing these factors optimally is crucial. Recently, trial-and-error (Rehamnia et al., 2021), grid search (Shahsavar et al., 2021), random search (Seyedzadeh et al., 2019), and Bayesian approaches (Zhang et al., 2020) have been employed to accomplish this. In this research, four machine learning approaches (i.e., BiRNN, MLP, RF, and XGBoost) are used in both hybrid and standalone frameworks, which are tuned using a grid search strategy. The BiRNN model, as the main ML approach, has five important regularization factors: the number of hidden layers, the number of neurons, the learning rate, the Epochs number, and the training algorithm. For the MLP model, the number of hidden layers, number of neurons, and training algorithm are the crucial factors. RF and XGBoost, as the tree-based ML approaches, are strongly influenced by the N_Estimatiors and Max-Depth parameters (Singh et al., 2022). In order to optimize the BiRNN model, it has defined a space interval for each significant hyperparameter (i.e., number of neurons, learning rate, number of neurons, and epoch number) to gain the best accuracy to forecast the ET_0 weekly. As can be seen, the optimal values of learning rate, neuron numbers, epoch values, and number of layers fall into the intervals of (0.000101–0.0141), (110−200), (44−70), and (fixed two layers), respectively. Table 3 summarizes the optimal settings of the ML approaches implemented for forecasting weekly ET_o. The modeling flowchart is depicted in Fig. 6.

2.11. Evaluation metrics

The performances of the developed machine learning models for forecasting weekly ET_o were evaluated by six statistical indicators, including RMSE (root mean square error), MAPE (mean absolute percent error), NSE (Nash-Sutcliffe efficiency), R (coefficient of correlation), I_A (index of agreement), and $U_{95\%}$ (uncertainty coefficient with 95% confidence level). Mathematically, RMSE (Ebtehaj et al., 2021), MAPE (Singh et al., 2022), NSE (Nash and Sutcliffe, 1970), R (Malik et al., 2021; Moriasi et al., 2015), I_A (Willmott, 1982), and $U_{95\%}$ (Patino and

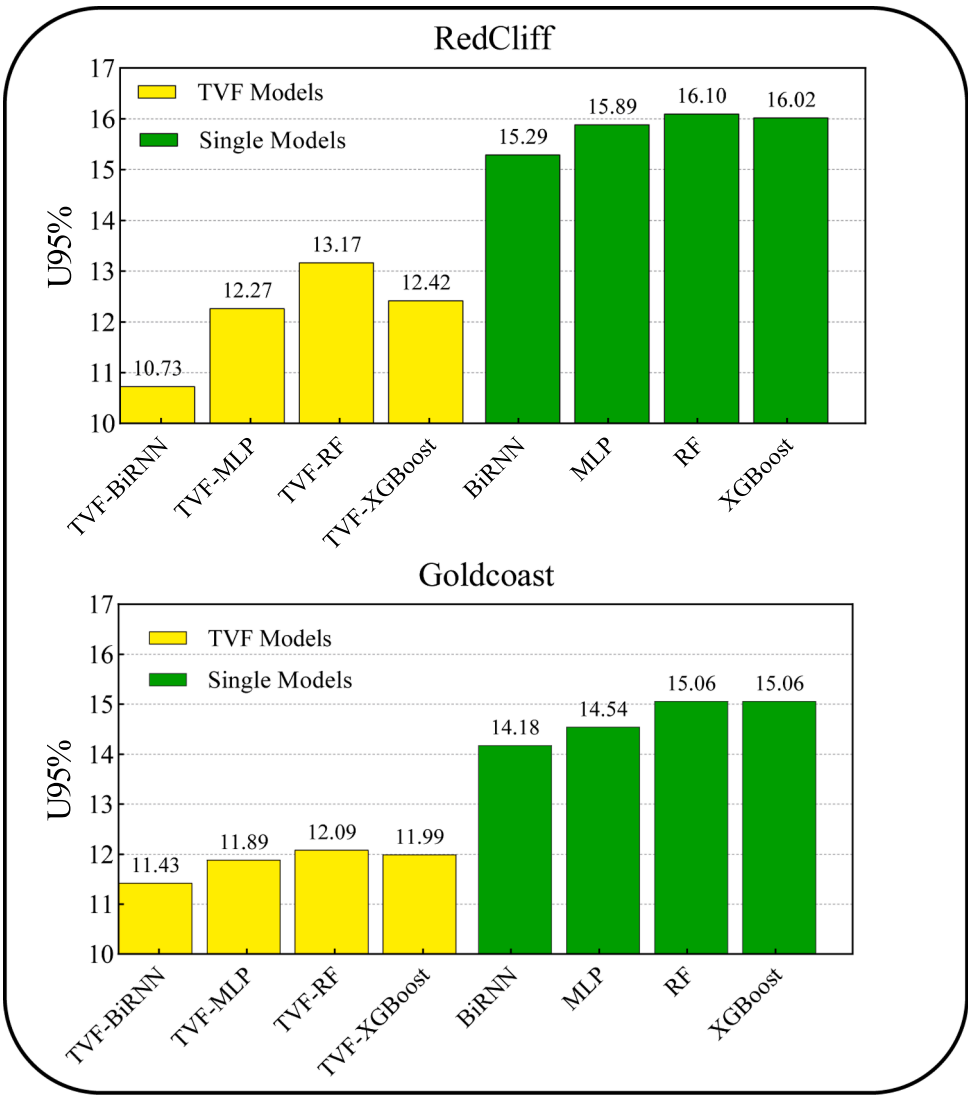


Fig. 11. Uncertainty variations of different models for Redcliffe and Gold Coast stations.

Ferreira, 2015) are respectively expressed as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (ET_o^{meas,i} - ET_o^{forc,i})^2} \quad (13)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{ET_o^{meas,i} - ET_o^{forc,i}}{ET_o^{meas,i}} \right| \quad (14)$$

$$NSE = 1 - \left[\frac{\sum_{i=1}^{N} (ET_o^{meas,i} - ET_o^{forc,i})^2}{\sum_{i=1}^{N} (ET_o^{meas,i} - \overline{ET_o^{meas}})^2} \right] \quad (15)$$

$$R = \frac{\sum_{i=1}^{N} (ET_o^{meas,i} - \overline{ET_o^{meas}}) \quad (ET_o^{forc,i} - \overline{ET_o^{forc}})}{\sqrt{\sum_{i=1}^{N} (ET_o^{meas,i} - \overline{ET_o^{meas}})^2 \quad \sum_{i=1}^{N} (ET_o^{forc,i} - \overline{ET_o^{forc}})^2}} \quad (16)$$

$$I_A = 1 - \left[\frac{\sum_{i=1}^{N} (ET_o^{forc,i} - ET_o^{meas,i})^2}{\sum_{i=1}^{N} (|ET_o^{forc,i} - \overline{ET_o^{meas}}| + |ET_o^{meas,i} - \overline{ET_o^{meas}}|)^2} \right] \quad (17)$$

$$U_{95\%} = 1.96 \sqrt{Error\ Standard\ deviation^2 + RMSE^2} \quad (18)$$

where, $ET_o^{meas,i}$ and $ET_o^{forc,i}$ are the measured and forecasted weekly ET_o values; $\overline{ET_o^{meas}}$, and $\overline{ET_o^{forc}}$ are the averages of the measured and forecasted weekly ET_o values; and N is the total number of observations.

3. Results and discussion

The proposed hybrid TVF-BiRNN model was designed to forecast weekly ET_o for Redcliffe and Gold Coast stations in Queensland, Australia. The performance of the TVF-BiRNN was compared against those of TVF-MLP, TVF-RF, TVF-XGBoost, BiRNN, MLP, RF, and XGBoost models based on the metrics of R, RMSE, MAPE, NSE, I_A, and $U_{95\%}$ during the training and testing periods.

Table 4 revealed that the proposed hybrid TVF-BiRNN model attained better accuracy in the forecast of weekly ET_o for Redcliffe station (R = 0.9281, RMSE = 3.8793, MAPE = 9.2010, NSE = 0.8588, I_A = 0.9621, $U_{95\%}$ = 10.7304), followed by TVF-MLP [R = 0.9120, RMSE = 4.5957, MAPE = 12.1399, NSE = 0.8018, I_A = 0.9442, $U_{95\%}$ = 12.2654], TVF-XGBoost [R = 0.9055, RMSE = 4.5654, MAPE = 11.5553, NSE = 0.8044, I_A = 0.9437, $U_{95\%}$ = 12.4234], and TVF-RF [R = 0.8908, RMSE = 4.7927, MAPE = 11.7583, NSE = 0.7845, I_A

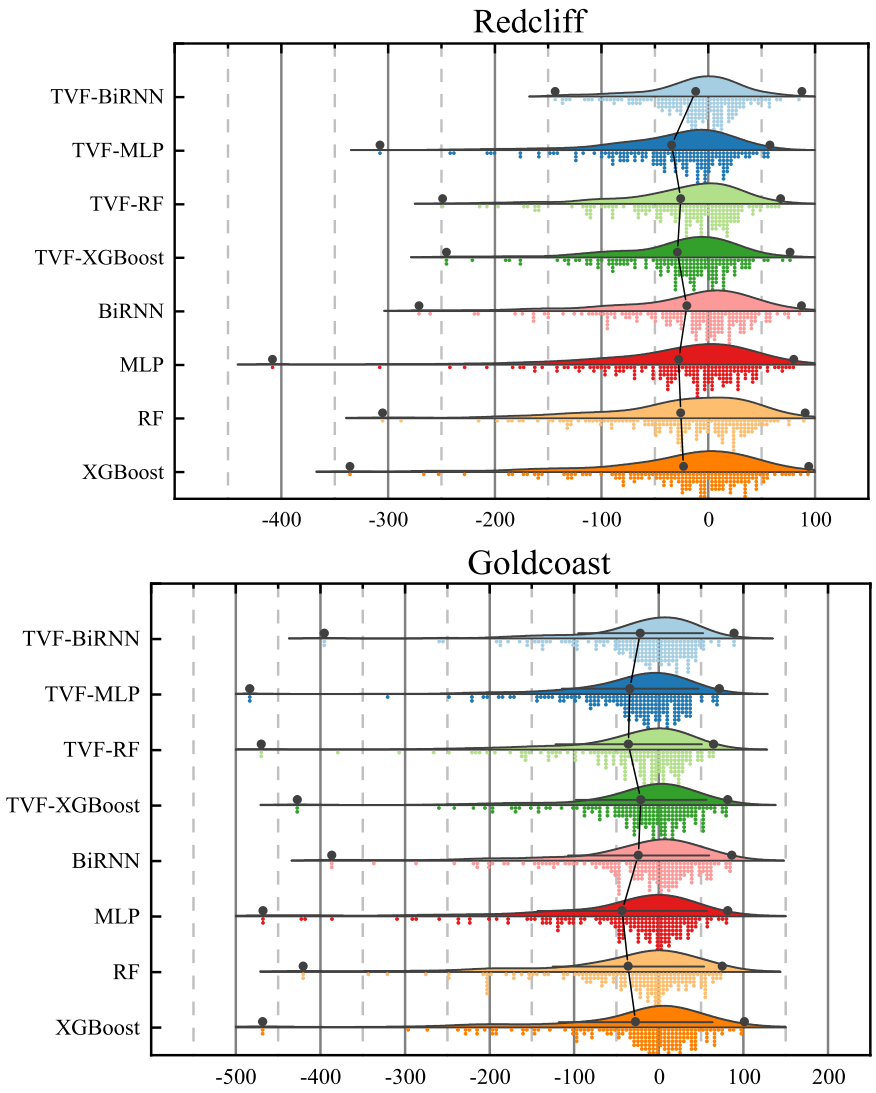


Fig. 12. Rain cloud graphs of relative deviations for Redcliffe and Gold Coast stations.

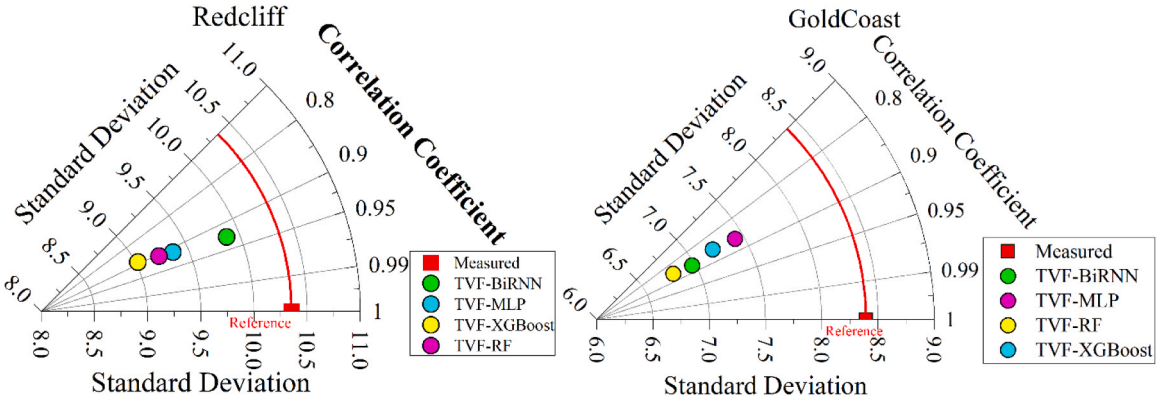


Fig. 13. Taylor diagrams for Redcliffe and Gold Coast stations.

= 0.9365, $U_{95\%}$ = 13.1700]. The performances of the standalone models (i.e., BiRNN, MLP, RF, and XGBoost) are lower than those of the TVF-BiRNN model (Table 4).

Fig. 7. shows the scatter (LHS) and violin (RHS) plots of the weekly ET_o simulated by the TVF-BiRNN, TVF-MLP, TVF-RF, TVF-XGBoost, BiRNN, MLP, RF, and XGBoost models for Redcliffe station. The results confirm that the TVF-BiRNN model (R = 0.9281 and RMSE = 3.8793) outperformed BiRNN (R = 0.8458; RMSE = 5.5086), as well as TVF-MLP and MLP; TVF-XGBoost and XGBoost, and TVF-RF and RF. The distributions of the violin plots also confirm the superiority of the TVF-BiRNN model over other benchmark comparing models. Overall, the TVF-BiRNN model is better than TVF-MLP, TVF-RF, TVF-XGBoost, BiRNN,

MLP, RF, and XGBoost.

Fig. 8 shows the swarm plot of the measured weekly ET_o and those simulated by TVF-BiRNN, TVF-MLP, TVF-RF, TVF-XGBoost, BiRNN, MLP, RF, and XGBoost for Redcliffe station. The proposed TVF-BiRNN model exhibited a swarm pattern very close to that of the measured data, followed by TVF-MLP, TVF-XGBoost, and TVF-RF. In contrast, the swarm plots from the standalone BiRNN, MLP, RF, and XGBoost indicate relatively poor performances in the forecast of weekly ET_o for this station.

Table 5 shows the modeling results for the Gold Coast station. The proposed hybrid TVF-BiRNN model again outperformed all other models in the forecast of weekly ET_o [R = 0.8717, RMSE = 4.1169, MAPE = 11.5408, NSE = 0.7580, I_A = 0.9243, $U_{95\%}$ = 11.4256]. The performance ranking of other models follows TVF-MLP, TVF-RF, and TVF-XGBoost. The performances of the BiRNN, MLP, RF, and XGBoost models are poor for the Gold Coast station. The modeling results again demonstrated the superior performance of the TVF-BiRNN model in the ET_o forecasting.

Fig. 9 displays the scatter (LHS) and violin (RHS) plots of the weekly ET_o simulated by the TVF-BiRNN, TVF-MLP, TVF-RF, TVF-XGBoost, BiRNN, MLP, RF, XGBoost models for Gold Coast station. The TVF-BiRNN model {R = 0.8717; RMSE = 4.1169} acquired a better degree of accuracy in terms of the violin distribution than BiRNN {R = 0.7951; RMSE = 5.1113}, followed by TVF-MLP, TVF-XGBoost, TVF-RF MLP, RF, and XGBoost.

Fig. 10 shows the swarm plot of the measured weekly ET_o and those simulated by TVF-BiRNN, TVF-MLP, TVF-RF, TVF-XGBoost, BiRNN, MLP, RF, and XGBoost. The TVF-BiRNN model displays a swarm distribution similar to that of the measured data, followed by TVF-MLP. The TVF-XGBoost, TVF-RF, and all standalone models (i.e., BiRNN, MLP, RF, and XGBoost) depict comparatively lower accuracy in the forecast of weekly ET_o for Gold Coast station.

Fig. 11 depicts the performances of the TVF-BiRNN, TVF-MLP, TVF-RF, TVF-XGBoost, BiRNN, MLP, RF, and XGBoost models in terms of uncertainty variation ($U_{95\%}$) in their forecasting of weekly ET_o for both Redcliffe and Gold Coast stations. The lower the $U_{95\%}$ value, the better the model's performance. It can be observed from the magnitudes of $U_{95\%}$ that TVF-BiRNN is the most accurate model for both stations (10.73 for Redcliffe station; 11.43 for Gold Coast station), followed by TVF-MLP, TVF-XGBoost, and TVF-RF. The hybrid models have fairly low values of $U_{95\%}$, compared to the standalone BiRNN, MLP, RF, and XGBoost models for both stations. Generally, the TVF-BiRNN model displays good efficiency with lower $U_{95\%}$ values.

The relative deviation is described in the form of rain cloud graphs (Fig. 12) for the TVF-BiRNN, TVF-MLP, TVF-RF, TVF-XGBoost, BiRNN, MLP, RF, and XGBoost models for both Redcliffe and Gold Coast stations. The TVF-BiRNN model apparently provided a better forecast with a range from − 100–100 for Redcliffe station, compared to TVF-RF and TVF-XGBoost [from − 250–100] and TVF-MLP [from − 350–100]. The MLP model had the worst performance according to the rain cloud distribution plots compared to other models. Similarly, the TVF-BiRNN with a range from − 440–100 is reasonably accurate against other benchmarking models for Gold Coast station.

Fig. 13 shows the Taylor diagrams of the measured weekly ET_o and simulations from TVF-BiRNN, TVF-MLP, TVF-RF, TVF-XGBoost, BiRNN, MLP, RF, and XGBoost for Redcliffe and Gold Coast stations. Taylor diagram (Taylor, 2001) provides a more conclusive evaluation among the models with reference ET_o based on standard deviation and correlation coefficient in a polar coordinate system. A model is precise by lying closely with the reference values. Fig. 13 verifies that TVF-BiRNN (green dot) is sitting close to the reference ET_o within the arc of 0.90–0.95 for Redcliffe station, and 0.8–0.9 for Gold Coast station, followed by TVF-MLP, TVF-RF, and TVF-XGBoost. Thus, TVF-BiRNN is the most accurate model for weekly ET_o forecasting for both stations.

The framework designed in this work was innovative in terms of TVF-BiRNN to weekly ET_o, but some limitations exist that can be considered to broaden the scope of future studies. The satellite-derived inputs can be another substitute instead of ground data to considerably enhance the TVF-BiRNN precision accuracy to forecast weekly ET_o. Additionally, synoptic mode indices are greatly influenced the ET_o and thus the utilization of southern oscillation index (SOI), sea surface temperatures (Nino3SST, Nino3.4SST, Nino4SST), pacific decadal oscillation (PDO), Indian Ocean dipole (IOD), El-Nino southern oscillation Modoki index (EMI), southern annular mode (SAM) can augment the forecasting ability of TVF-BiRNN model (Nicholls, 2004).

The black-box characteristics of the deep learning models have some limitations, which are difficult and challenging to understand and verify the complex relationship between the predictors and target for the learning process. Therefore, integration of deep learning with numerical weather prediction models can provide a new direction to overcome. Furthermore, the TVF-BiRNN model can be more diversified and improved by Bayesian Model Averaging (Sloughter et al., 2010) and ensemble techniques (Tiwari and Chatterjee, 2011) to handle uncertainties.

The data decomposition of TVF in the construction of the TVF-BiRNN model is helpful to upsurge the accuracy of the BiRNN by simultaneously handling the non-stationary and non-linearity issues mode (Rehman and Mandic, 2010). The better forecasting of the TVF-BiRNN against other comparing models is undoubtedly attributable to data transformation in a deterministic way and decomposed into several signals, which results in a lower error. The results affirmed that the TVF-BiRNN models perform well in forecasting weekly ET_o for both stations.

Additionally, by integrating the TVF with the BiRNN model, further enhancement in model accuracy was accomplished with the help of the Boruta-Extra Tree feature selection method to identify the highest-ranking input signals. The deep learning models are fully dependent on historical data, which can considerably impact the learning process. The outcomes of this research work ensure that an appropriate feature selection is compulsory (Badr and Fahmy, 2004; Singh et al., 2012; Sweetlin et al., 2017).

4. Conclusion

This research developed a new model, TVF-BiRNN, that combined the bidirectional recurrent neural network with TVF-EMD to forecast weekly ET_o for Redcliffe and Gold Coast stations in Australia. Further, the Boruta-Extra Tree method was also introduced to determine and select the best IMF signals. The main scheme of this research entailed the utilization of the TVF-EMD algorithm to demarcate the input predictors into IMFs and residuals. The Boruta-Extra Tree method was employed to select the IMF features, improving the models' accuracy. The selected IMFs were then used as inputs in BiRNN to develop the new TVF-BiRNN model. The MLP, RF, and XGBoost models were also adopted for comparison purposes to construct the TVF-MLP, TVF-RF, and TVF-XGBoost models. Moreover, the standalone models were also benchmarked against those hybrid counterpart models. The performances of the models were evaluated by using a set of statistical metrics (including R, RMSE, MAPE, NSE, I_A, and $U_{95\%}$), which demonstrated that TVF-BiRNN was the most precise model to forecast weekly ET_o for both stations. Specifically, these metrics for TVF-BiRNN were: R = 0.9281, RMSE = 3.8793, MAPE = 9.2010, NSE = 0.8588, I_A = 0.9621, $U_{95\%}$ = 10.7304 for Redcliffe station; R = 0.8717, RMSE = 4.1169, MAPE = 11.5408, NSE = 0.7580, I_A = 0.9243, $U_{95\%}$ = 11.4256 for Gold Coast station.

In addition to ET_o forecasting, the TVF-BiRNN modeling approach can also potentially be applied to other problems in the areas of hydrology, agriculture, climate change adaptation, and renewable and sustainable energy. The ET_o forecasting indicates the evaporated and transpired amount of water and is a crucial parameter in hydrology and agriculture that serves as a reference to gauge the water requirements of crops and plants. This assists farmers in determining crop water requirements to improve irrigation scheduling and guaranteeing effective

water usage to handle overwatering or under-watering issues. Moreover, forecasting ET_o is very useful for calculating the water demand in agriculture, urban water supply, and environmental conservation and helps in water resource management. This work will also help farmers determine crop planning and selection according to the specific region and season.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

References

Ahmed, A.A.M., Deo, R.C., Ghahramani, A., Raj, N., Feng, Q., Yin, Z., Yang, L., 2021. LSTM integrated with Boruta-random forest optimiser for soil moisture estimation under RCP4. 5 and RCP8. 5 global warming scenarios. Stoch. Environ. Res. Risk Assess. 1–31.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration: Guidelines for computing crop requirements. Irrig. Drain. Pap. No. 56, FAO. https://doi.org/10.1016/j.eja.2010.12.001.

Apaydin, H., Feizi, H., Sattari, M.T., Colak, M.S., Shamshirband, S., Chau, K.-W., 2020. Comparative analysis of recurrent neural network architectures for reservoir inflow forecasting. Water 12, 1500. https://doi.org/10.3390/w12051500.

Asadollah, S.B.H.S., Sharafati, A., Motta, D., Yaseen, Z.M., 2020. River water quality index prediction and uncertainty analysis: a comparative study of machine learning models. J. Environ. Chem. Eng. https://doi.org/10.1016/j.jece.2020.104599.

Ayaz, A., Rajesh, M., Singh, S.K., Rehana, S., 2021. Estimation of reference evapotranspiration using machine learning models with limited data. AIMS Geosci. 7, 268–290.

Babaeian, E., Paheding, S., Siddique, N., Devabhaktuni, V.K., Tuller, M., 2022. Short- and mid-term forecasts of actual evapotranspiration with deep learning. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2022.128078.

Badr, A., Fahmy, A., 2004. A proof of convergence for ant algorithms. Inf. Sci. (Ny.). 160, 267–279.

Bennasar, M., Hicks, Y., Setchi, R., 2015. Feature selection using joint mutual information maximisation. Expert Syst. Appl. 42, 8520–8532.

Bhagat, S.K., Tiyasha, T., Tung, T.M., Mostafa, R.R., Yaseen, Z.M., 2020. Manganese (Mn) removal prediction using extreme gradient model. Ecotoxicol. Environ. Saf. 204, 111059 https://doi.org/10.1016/j.ecoenv.2020.111059.

Bhattarai, A., Qadir, D., Sunusi, A.M., Getachew, B., Mallah, A.R., 2023. Dynamic sliding window-based long short-term memory model development for pan evaporation forecasting. knowledge-based. Eng. Sci. 4, 37–54.

Bouguettaya, A., Zarzour, H., Kechida, A., Taberkit, A.M., 2022. Deep learning techniques to classify agri. Cult. Crops UAV Imag.: a Rev. Neural Comput. Appl. 1 26.

Breiman, L., 1999. Randon Forests. Machinelearning202. Pbworks. Com. 1–35.

Brutsaert, W., 2013. Evaporation Into the Atmosphere: Theory, History and Applications. Springer Science & Business Media.

Chen, S., He, C., Huang, Z., Xu, X., Jiang, T., He, Z., Liu, J., Su, B., Feng, H., Yu, Q., 2022. Using support vector machine to deal with the missing of solar radiation data in daily reference evapotranspiration estimation in China. Agric. For. Meteorol. 316, 108864.

Chen, T., Guestrin, C., 2016. XGBoost. Proc. 22nd ACM SIGKDD. Int. Conf. Knowl. Discov. Data Min. https://doi.org/10.1145/2939672.2939785.

Danandeh Mehr, A., Rikhtehgar Ghiasi, A., Yaseen, Z.M., Sorman, A.U., Abualigah, L., 2022. A novel intelligent deep learning predictive model for meteorological drought forecasting. J. Ambient Intell. Humaniz. Comput. 1–15.

Dong, J., Zhu, Y., Jia, X., Han, X., Qiao, J., Bai, C., Tang, X., 2021. Nation-scale reference evapotranspiration estimation by using deep learning and classical machine learning models in China. J. Hydrol., 127207

Ebtehaj, I., Sammen, S.S., Sidek, L.M., Malik, A., Sihag, P., Al-Janabi, A.M.S., Chau, K.-W., Bonakdari, H., 2021. Prediction of daily water level using new hybridized GS-GMDH and ANFIS-FCM models. Eng. Appl. Comput. Fluid Mech. 15, 1343–1361. https://doi.org/10.1080/19942060.2021.1966837.

Fayer, G., Lima, L., Miranda, F., Santos, J., Campos, R., Bignoto, V., Andrade, M., Moraes, M., Ribeiro, C., Capriles, P., 2023. A temporal fusion transformer deep learning model for long-term streamflow forecasting: a case study in the funil reservoir, Southeast Brazil. knowledge-based. Eng. Sci. 4, 73–88.

Fu, M., Fan, T., Ding, Z., Salih, S.Q., Al-Ansari, N., Yaseen, Z.M., 2020. Deep learning data-intelligence model based on adjusted forecasting window scale: application in daily streamflow simulation. IEEE Access 8, 32632–32651. https://doi.org/10.1109/ACCESS.2020.2974406.

Genuer, R., Poggi, J.M., Tuleau-Malot, C., 2010. Variable selection using random forests. Pattern Recognit. Lett. 31, 2225–2236. https://doi.org/10.1016/j.patrec.2010.03.014.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Mach. Learn. 63, 3–42.

Goyal, P., Kumar, S., Sharda, R., 2023. A review of the Artificial Intelligence (AI) based techniques for estimating reference evapotranspiration: Current trends and future perspectives. Comput. Electron. Agric. https://doi.org/10.1016/j.compag.2023.107836.

Granata, F., Di Nunno, F., 2021. Forecasting evapotranspiration in different climates using ensembles of recurrent neural networks. Agric. Water Manag. https://doi.org/10.1016/j.agwat.2021.107040.

Granata, F., Di Nunno, F., 2023. Neuroforecasting of daily streamflows in the UK for short-and medium-term horizons: A novel insight. J. Hydrol. 624, 129888.

Granata, F., Di Nunno, F., de Marinis, G., 2022. Stacked machine learning algorithms and bidirectional long short-term memory networks for multi-step ahead streamflow forecasting: A comparative study. J. Hydrol. 613, 128431.

Hameed, M.M., Alomar, M.K., Mohd Razali, S.F., Kareem Khalaf, M.A., Baniya, W.J., Sharafati, A., Alsaadi, M.A., 2021. Application of artificial intelligence models for evapotranspiration prediction along the southern coast of Turkey. Complexity 2021, 1–20. https://doi.org/10.1155/2021/8850243.

Hargreaves, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. Appl. Eng. Agric. 1, 96–99. https://doi.org/10.13031/2013.26773.

Jaihuni, M., Basak, J.K., Khan, F., Okyere, F.G., Sihalath, T., Bhujel, A., Park, J., Lee, D. H., Kim, H.T., 2022. A novel recurrent neural network approach in forecasting short term solar irradiance. ISA Trans. 121, 63–74. https://doi.org/10.1016/j.isatra.2021.03.043.

Jamei, M., Karbasi, M., Alawi, O.A., Kamar, H.M., Khedher, K.M., Abba, S.I., Yaseen, Z. M., 2022a. Earth skin temperature long-term prediction using novel extended Kalman filter integrated with Artificial Intelligence models and information gain feature selection. Sustain. Comput. Inform. Syst. 35, 100721.

Jamei, M., Karbasi, M., Malik, A., Abualigah, L., Islam, A.R.M.T., Yaseen, Z.M., 2022b. Computational assessment of groundwater salinity distribution within coastal multi-aquifers of Bangladesh. Sci. Rep. 12 (1), 28.

Jamei, Mehdi, Karbasi, M., Malik, A., Jamei, Mozhdeh, Kisi, O., Yaseen, Z.M., 2022c. Long-term multi-step ahead forecasting of root zone soil moisture in different climates: Novel ensemble-based complementary data-intelligent paradigms. Agric. Water Manag. 269, 107679.

Keshtegar, B., Abdullah, S.S., Huang, Y.F., Saggi, M.K., Khedher, K.M., Yaseen, Z.M., 2022. Reference evapotranspiration prediction using high-order response surface method. Theor. Appl. Climatol. https://doi.org/10.1007/s00704-022-03954-4.

Kisi, O., 2010. Fuzzy genetic approach for modeling reference evapotranspiration, 136, 175–183.

Kişi, Ö., 2007. Streamflow forecasting using different artificial neural network algorithms. J. Hydrol. Eng. 12, 532–539. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:5(532).

Kursa, M.B., Jankowski, A., Rudnicki, W.R., 2010. Boruta – a system for feature selection. Fundam. Inform. 101, 271–285. https://doi.org/10.3233/FI-2010-288.

Kurth, T., Treichler, S., Romero, J., Mudigonda, M., Luehr, N., Phillips, E., Mahesh, A., Matheson, M., Deslippe, J., Fatica, M., 2018. Exascale deep learning for climate analytics. in: SC18: International Conference for High Performance Computing, Networking, Storage and Analysis. IEEE,, pp. 649–660.

Lang, D., Zheng, J., Shi, J., Liao, F., Ma, X., Wang, W., Chen, X., Zhang, M., 2017. A comparative study of potential evapotranspiration estimation by eight methods with FAO Penman–Monteith method in southwestern China. Water 9, 734.

Li, H., Li, Z., Mo, W., 2017. A time varying filter approach for empirical mode decomposition. Signal Process. 138, 146–158. https://doi.org/10.1016/j.sigpro.2017.03.019.

Li, Y., Liang, Z., Hu, Y., Li, B., Xu, B., Wang, D., 2019. A multi-model integration method for monthly streamflow prediction: modified stacking ensemble strategy. J. Hydroinformatics 22, 310–326. https://doi.org/10.2166/hydro.2019.066.

Malik, A., Tikhamarine, Y., Al-Ansari, N., Shahid, S., Sekhon, H.S., Pal, R.K., Rai, P., Pandey, K., Singh, P., Elbeltagi, A., Sammen, S.S., 2021. Daily pan-evaporation estimation in different agro-climatic zones using novel hybrid support vector regression optimized by Salp swarm algorithm in conjunction with gamma test. Eng. Appl. Comput. Fluid Mech. 15, 1075–1094. https://doi.org/10.1080/19942060.2021.1942990.

Martí, P., González-Altozano, P., López-Urrea, R., Mancha, L.A., Shiri, J., 2015. Modeling reference evapotranspiration with calculated targets. Assessment and implications. Agric. Water Manag. 149, 81–90. https://doi.org/10.1016/j.agwat.2014.10.028.

Mehr, A.D., Jabarnejad, M., Nourani, V., 2019. Pareto-optimal MPSA-MGGP: a new gene-annealing model for monthly rainfall forecasting. J. Hydrol. 571, 406–415.

Merufinia, E., Sharafati, A., Abghari, H., Hassanzadeh, Y., 2023. On the simulation of streamflow using hybrid tree-based machine learning models: a case study of Kurkursar basin, Iran. Arab. J. Geosci. https://doi.org/10.1007/s12517-022-11045-x.

Monteith, J.L., 1965. Evaporation and Environment. In: Symposia of the Society for Experimental Biology. Cambridge University Press (CUP), Cambridge, pp. 205–234.

Moriasi, D.N., Gitau, M.W., Pai, N., Daggupati, P., 2015. Hydrologic and water quality models: performance measures and evaluation criteria. Trans. ASABE 58, 1763–1785. https://doi.org/10.13031/trans.58.10715.

Naghibi, S.A., Ahmadi, K., Daneshi, A., 2017. Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping. Water Resour. Manag. 31, 2761–2775.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I - a discussion of principles. J. Hydrol. 10, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Nawandar, N.K., Cheggoju, N., Satpute, V., 2021. ANN-based model to predict reference evapotranspiration for irrigation estimation. In: Proceedings of International

Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications. Springer, pp. 671–679.
Nicholls, N., 2004. The changing nature of Australian droughts. Clim. Change 63, 323–336.
Pashazadeh, A., Javan, M., 2020. Comparison of the gene expression programming, artificial neural network (ANN), and equivalent Muskingum inflow models in the flood routing of multiple branched rivers. Theor. Appl. Climatol. 139, 1349–1362.
Patino, C.M., Ferreira, J.C., 2015. Confidence intervals: a useful statistical tool to estimate effect sizes in the real world. J. Bras. Pneumol. https://doi.org/10.1590/s1806-37562015000000314.
Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of the surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100, 81–92.
Rasp, S., Pritchard, M.S., Gentine, P., 2018. Deep learning to represent subgrid processes in climate models. Proc. Natl. Acad. Sci. 115, 9684–9689.
Rehamnia, I., Benlaoukli, B., Jamei, M., Karbasi, M., Malik, A., 2021. Simulation of seepage flow through embankment dam by using a novel extended Kalman filter based neural network paradigm: Case study of Fontaine Gazelles Dam, Algeria. Meas. J. Int. Meas. Confed. https://doi.org/10.1016/j.measurement.2021.109219.
Rehman, N., Mandic, D.P., 2010. Multivariate empirical mode decomposition. Proc. R. Soc. A Math. Phys. Eng. Sci. https://doi.org/10.1098/rspa.2009.0502.
Rezaeianzadeh, M., Tabari, H., Arabi Yazdi, A., Isik, S., Kalin, L., 2014. Flood flow forecasting using ANN, ANFIS and regression models. Neural Comput. Appl. 25, 25–37.
Roy, D.K., Lal, A., Sarker, K.K., Saha, K.K., Datta, B., 2021. Optimization algorithms as training approaches for prediction of reference evapotranspiration using adaptive neuro fuzzy inference system. Agric. Water Manag. https://doi.org/10.1016/j.agwat.2021.107003.
Schuster, M., Paliwal, K.K., 1997. Bidirectional recurrent neural networks. IEEE Trans. Signal Process. https://doi.org/10.1109/78.650093.
Seyedzadeh, S., Rahimian, F.P., Rastogi, P., Glesk, I., 2019. Tuning machine learning models for prediction of building energy loads. Sustain. Cities Soc. 47, 101484.
Shahsavar, A., Jamei, M., Karbasi, M., 2021. Experimental evaluation and development of predictive models for rheological behavior of aqueous Fe3O4 ferrofluid in the presence of an external magnetic field by introducing a novel grid optimization based-Kernel ridge regression supported by sensitivity. Powder Technol. https://doi.org/10.1016/j.powtec.2021.07.037.
Sharafati, A., Asadollah, S.B.H.S., Hosseinzadeh, M., 2020. The potential of new ensemble machine learning models for effluent quality parameters prediction and related uncertainty. Process Saf. Environ. Prot.
Sharifani, K., Amini, M., 2023. Machine learning and deep learning: a review of methods and applications. World Inf. Technol. Eng. J. 10, 3897–3904.
Singh, G., Kumar, N., Kumar Verma, A., 2012. Ant colony algorithms in MANETs: a review. J. Netw. Comput. Appl. https://doi.org/10.1016/j.jnca.2012.07.018.
Singh, U.K., Jamei, M., Karbasi, M., Malik, A., Pandey, M., 2022. Application of a modern multi-level ensemble approach for the estimation of critical shear stress in cohesive sediment mixture. J. Hydrol., 127549
Sit, M., Demiray, B.Z., Xiang, Z., Ewing, G.J., Sermet, Y., Demir, I., 2020. A comprehensive review of deep learning applications in hydrology and water resources. Water Sci. Technol.
Sloughter, J.M., Gneiting, T., Raftery, A.E., 2010. Probabilistic wind speed forecasting using ensembles and Bayesian model averaging. J. Am. Stat. Assoc. 105, 25–35.
Song, C., Chen, X., Wu, P., Jin, H., 2021. Combining time varying filtering based empirical mode decomposition and machine learning to predict precipitation from nonlinear series. J. Hydrol. 603, 126914 https://doi.org/10.1016/j.jhydrol.2021.126914.
Sweetlin, J.D., Nehemiah, H.K., Kannan, A., 2017. Feature selection using ant colony optimization with tandem-run recruitment to diagnose bronchitis from CT scan images. Comput. Methods Prog. Biomed. https://doi.org/10.1016/j.cmpb.2017.04.009.
Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106, 7183–7192. https://doi.org/10.1029/2000JD900719.
Tiwari, M.K., Chatterjee, C., 2011. A new wavelet–bootstrap–ANN hybrid model for daily discharge forecasting. J. Hydroinformatics 13, 500. https://doi.org/10.2166/hydro.2010.142.
Torres, A.F., Walker, W.R., McKee, M., 2011. Forecasting daily potential evapotranspiration using machine learning and limited climatic data. Agric. Water Manag. 98, 553–562.
Wang, K., Fu, W., Chen, T., Zhang, B., Xiong, D., Fang, P., 2020. A compound framework for wind speed forecasting based on comprehensive feature selection, quantile regression incorporated into convolutional simplified long short-term memory network and residual error correction. Energy Convers. Manag. 222, 113234 https://doi.org/10.1016/j.enconman.2020.113234.
Wang, X., Gao, B., Wang, X.S., 2022. Investigating the ability of deep learning on actual evapotranspiration estimation in the scarcely observed region. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2022.127506.
Willmott, C., 1982. Some comments on the evaluation of model performance. Bull. Am. Meteorol. Soc. https://doi.org/10.1175/1520-0477(1982)063<1309:SCOTEO>2.0.CO;2.
Yan, X., Yang, N., Ao, R., Mohammadian, A., Liu, J., Cao, H., Yin, P., 2023. Deep learning for daily potential evapotranspiration using a HS-LSTM approach. Atmos. Res. https://doi.org/10.1016/j.atmosres.2023.106856.
Yaseen, Z.M., 2021. An insight into machine learning models era in simulating soil, water bodies and adsorption heavy metals: Review, challenges and solutions. Chemosphere 277, 130126. https://doi.org/10.1016/j.chemosphere.2021.130126.
Yin, J., Deng, Z., Ines, A.V.M., Wu, J., Rasu, E., 2020. Forecast of short-term daily reference evapotranspiration under limited meteorological variables using a hybrid bi-directional long short-term memory model (Bi-LSTM). Agric. Water Manag. 242, 106386 https://doi.org/10.1016/j.agwat.2020.106386.
Zhang, G.P., Patuwo, E.B., Michael, Y., H., 1998. Forecasting with artificial neural networks: The state of the art. Int. J. Forecast. 14, 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7.
Zhang, W., Wu, C., Zhong, H., Li, Y., Wang, L., 2020. Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. Geosci. Front. https://doi.org/10.1016/j.gsf.2020.03.007.
Zhang, X., Liu, Z., Miao, Q., Wang, L., 2018. An optimized time varying filtering based empirical mode decomposition method with grey wolf optimizer for machinery fault diagnosis. J. Sound Vib. 418, 55–78. https://doi.org/10.1016/j.jsv.2017.12.028.
Zheng, Z., Ali, M., Jamei, M., Xiang, Y., Karbasi, M., Yaseen, Z.M., Farooque, A.A., 2023. Design data decomposition-based reference evapotranspiration forecasting model: A soft feature filter based deep learning driven approach. Eng. Appl. Artif. Intell. 121, 105984.
Zhou, J., Qiu, Y., Armaghani, D.J., Zhang, W., Li, C., Zhu, S., Tarinejad, R., 2021. Predicting TBM penetration rate in hard rock condition: a comparative study among six XGB-based metaheuristic techniques. Geosci. Front. 12, 101091.