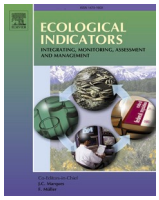
Ecological Indicators 155 (2023) 111030

Contents lists available at ScienceDirect

Ecological Indicators

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



New achievements on daily reference evapotranspiration forecasting: Potential assessment of multivariate signal decomposition schemes

Mumtaz Ali [a,c,h,*], Mehdi Jamei [b,h,*], Ramendra Prasad [d], Masoud Karbasi [e,h], Yong Xiang [f], Borui Cai [f], Shahab Abdulla [a,c], Aitazaz Ahsan Farooque [g,h], Abdulhaleem H. Labban [i]

[a] UniSQ College, University of Southern Queensland 4350 QLD, Australia
[b] Faculty of Engineering, Shohadaye Hoveizeh Campus of Technology, Shahid Chamran University of Ahvaz, Dashte Azadegan, 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 Science, School of Science and Technology, The University of Fiji, Saweni, Lautoka, Fiji
[e] Water Engineering Department, Faculty of Agriculture, University of Zanjan, Zanjan, Iran
[f] School of Information Technology, Deakin University, VIC 3125, Australia
[g] Faculty of Sustainable Design Engineering, University of Prince Edward Island, Charlottetown, PE, Canada
[h] Canadian Centre for Climate Change and Adaptation, University of Prince Edward Island, St Peters, PE, Canada
[i] Centre of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

ARTICLE INFO

Keywords:
Reference evapotranspiration
Multivariate variational mode decomposition
Multivariate empirical mode decomposition
Boosted regression tree

ABSTRACT

Reference evapotranspiration (ET_o) is a vital climate parameter affecting plants' water use. ET_o can generate large deficits in soil moisture and runoff in different regions and seasons, leading to uncertainties in drought warning systems. A novel multivariate variational mode decomposition integrated with a boosted regression tree (i.e., MVMD-BRT) is constructed to forecast daily ET_o. Firstly, the correlation matrix based on cross-correlation was computed to investigate the significant input predictor lags of daily ET_o. Secondly, the MVMD technique decomposes the significant input lags into signals called intrinsic mode functions (IMFs). Thirdly, the IMFs were then employed in the BRT to build the MVMD-BRT model for daily ETo forecasting. A comparative assessment of MVMD against multivariate empirical mode decomposition (MEMD) was also performed on the same lines to develop the MEMD-BRT model. The MVMD-BRT model is compared against the random forest (RF) and hybrid MVMD-RF, MEMD-RF, extreme learning machine (ELM), and hybrid MVMD-ELM, MEMD-ELM, and cascaded feedforward neural network (CFNN) along with its hybrid MVMD-CFNN models for two stations in Queensland, Australia using a set of goodness-of-fit metrics. The results prove that the MVMD-BRT provide accurate daily ET_o forecasting against the benchmark models. The MVMD-BRT model yielded the highest accuracy in terms of ($WI_E = 0.9070$, $NS_E = 0.8421$, $LM_E = 0.6529$, $KGE = 0.8792$) and ($WI_E = 0.8966$, $NS_E = 0.8396$, $LM_E = 0.6521$, $KGE = 0.8803$) for Brisbane and Gympie stations against the comparing models.

1. Introduction

The rapid onset drought or flash drought is harmful for crop production and the whole agricultural sector and water resources. Flash drought occurs in weeks, caused by rapid intensification from near-normal soil moisture to drought conditions (IPCC, 2021). There is an

Abbreviations: ET_o, Reference evapotranspiration; MVMD, Multivariate variational mode decomposition; MEMD, Multivariate empirical mode decomposition; BRT, Boosted regression tree; ELM, Extreme learning machine; RF, Random forest; CFNN, Cascaded feedforward neural network; AIMFs, Intrinsic mode functions; Radn, Solar radiation; WS, Wind speed; RH_{min}, Minimum relative humidity; RH_{max}, Maximum relative humidity; T_{min}, Minimum temperature; T_{max}, Maximum temperature; Rain, Rainfall; R, Correlation coefficient; MSE, Mean squared error; RMSE, Root mean squared error; MAE, Mean Absolute error; WI_E, Willmott's Index of agreement; NS_E, Nash-Scuttle estimator; LM_E, Legates and MacCabe's; KGE, Kling-Gupta efficiency; RRMSE, Relative root mean squared percentage error; RMAE, Relative mean absolute percentage error; ML, Machine learning; NN, Neural network; MRA, Multi-resolution analysis; EMD, Empirical mode decomposition; VMD, Variational mode decomposition.

* Corresponding authors at: UniSQ College, University of Southern Queensland 4350 QLD, Australia (M. Ali); Canadian Centre for Climate Change and Adaptation, University of Prince Edward Island, St Peters, PE, Canada (M. Jamei).

E-mail addresses: mumtaz.ali@usq.edu.au (M. Ali), M.jamei@shhut.ac.ir (M. Jamei).

https://doi.org/10.1016/j.ecolind.2023.111030
Received 1 May 2023; Received in revised form 15 September 2023; Accepted 28 September 2023

increase in agricultural and ecological drought that can be attributed to human influence with medium confidence (IPCC, 2013). In addition, comparing the highly vulnerable regions with deficient vulnerability regions, the human mortality from events such as floods, droughts and storms was 15 times higher between 2010 and 2020 (IPCC, 2013). The flash drought can occur in Australia in all seasons. All droughts emanate from lower-than-normal rainfall and intensify with higher temperatures, wind speed and incoming solar radiation. However, during flash droughts, with higher temperatures, the evapotranspiration increases. Reference evapotranspiration (ET_o) is an essential component of drought that acts as a conduit for water transfer from the soil, water and vegetation surfaces and transpiration from plants into the atmosphere. ET_o drives the soil moisture, near-surface relative humidity, moisture loss from plants, and plant stress, the surface water-energy–momentum balance or terrestrial latent heat fluxes (Douville et al., 2021).

ET_o can produce large deficits in soil moisture and runoff in different regions and seasons (Douville et al., 2021) and large uncertainties in drought early warning systems. The magnitude and intensity of droughts cannot be forecasted with precision if the land–atmosphere coupling, i.e., evapotranspiration information, is missing (Meng, Evans et al. 2014). If the evapotranspiration is not forecasted early enough, it can cause extensive damage to the crops, agricultural production, ecosystems, ecological damage and the economy. In addition, global warming has contributed to increased terrestrial evapotranspiration (medium confidence), and due to increased atmospheric demand, increased evapotranspiration has been conjectured for south-western Australia with a high confidence level (IPCC, 2021). Hence, it is imperative to develop precise forecasting models for evapotranspiration forecasts for drought mitigation, water resource management, agricultural sustainability, drought and heat waves detection, urban heat islands climate and cloud formation (Meng et al., 2014).

Forecasting ET_0 has been a challenge since there is a multitude of unresolved critical Earth-System-Science challenges regarding the complex nature of ET_o (Fisher et al., 2017). For the current application purposes, the daily ET_0 is largely hindcasted or estimated using the adapted Penman-Monteith equation with relative humidity and wind speed data and daily solar radiation data as inputs (Webb, 2010). In order to know the magnitude of the ET_0 at a point in time, the magnitudes of other parameters need to be known. This only allows for hindcasting, yet for decision support systems and decision-making, the knowledge of future ET_0 is essential, which the physical models are unable to capture properly. As a result, advanced machine learning and artificial intelligence models can be used to provide further insights and produce accurate forecasts of this importantland–atmosphere coupling variable.

Machine Learning (ML) algorithm provides a more direct and effective solution (Granata and Di Nunno, 2021). ET_o forecasting with machine learning models is the least explored hydrological variable in literature and the prediction of ET_0 by machine learning was developed late. The classical autoregressive integrated moving average (ARIMA) (Landeras et al., 2009) and artificial neural network (ANN) (Trajkovic et al., 2003, Landeras et al., 2009, Kisi et al., 2015) models were trialled primarily. Recently, the classical k-Nearest Neighbor algorithm (KNN) (Feng and Tian, 2020) and support vector machine (SVM) (Chia et al., 2020) were also applied. When comparing ANN with adaptive neuro-fuzzy inference system (ANFIS) with grid partition (GP) and ANFIS with subtractive clustering (SC) and gene expression programming (GEP), Kisi et al. (2015) found that the GEP model provided the worst estimates while the overall accuracies of ANN, ANFIS-GP and ANFIS-SC models were similar. To get more precise forecasts, advanced modelling approaches are necessary. The newer standalone modelling approaches, including extreme learning machine (ELM), backpropagation neural networks optimized by genetic algorithm (GANN) and wavelet neural networks (WNN) models, were applied by Feng et al. (2016) in estimating evapotranspiration in a Southwest China region. They evaluated the performances of ELM, GANN and WNN against two temperature-based (Hargreaves and modified Hargreaves) models and three radiation-based (Makkink, Priestley–Taylor and Ritchie) ET_o models. The results showed that the proposed models, ELM and GANN models recorded better performances with the competing WNN model, and the Hargreaves and modified Hargreaves models and the Makkink, Priestley–Taylor and Ritchie models (Feng et al., 2016). In another study, the random forest (RF) and gene-expression programming (GEP) methods were used by Wang et al. (2019) in estimating ET_0 using the data from 24 weather stations in a karst region of southwest China. They found that RF-based models were suitable for water balance research, while the GEP-based models were more suited for agricultural irrigation applications. In another study, the newer approaches namely, the multilayer perceptron (MLP), Generalized Regression Neural Network (GRNN), ELM, Support Vector Machines (SVM), RF and XGBoost were evaluated and the study revealed that ELM was able register better performances for ET_o estimates (Bellido-Jiménez et al., 2021).

In many of these studies in predicting ETo, the prediction accuracy is still not ideal, mainly due to the limitation of these classical machine learning models. The non-linearity, uncertain and stochastic nature of the ET_o makes it difficult for these models to capture the embedded features sufficiently. The forecasting accuracy of the models is severely affected if all pertinent features/patterns, such as trends, seasonality, cyclic behavior, outliers and abrupt changes in the time series are not properly captured. Hence, to unveil and present the embedded features within the data series, a multiresolution analysis (MRA) tool becomes necessary. To extract the underlying sub-frequencies, sequential decomposition is commonly used whereby each input series is decomposed one at a time despite having multiple inputs (Prasad et al., 2019b). The frequently used methods include Fourier spectra analysis (Soman et al., 2015); discrete wavelet transformation (Mallat, 1989, Mallat, 1998, Nourani et al., 2009, Krishna et al., 2011, Nourani et al., 2014, Deo et al., 2016a; Deo et al., 2016b); Empirical Mode Decomposition (EMD) (Huang et al., 1998); Ensemble EMD (EEMD) (Wu and Huang, 2009), complete ensemble EMD with adaptive noise (CEEMDAN) (Torres et al., 2011) and improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) (Colominas et al., 2014). To overcome sequential decomposition and conduct parallel decomposition simultaneously on multiple predictor inputs, the variational mode decomposition (VMD) methods are preferred. VMD, an adaptive and non-recursive decomposition tool, can concurrently decompose and simultaneously resolve the embedded sub-frequency components in multiple input predictors without losing any information. In addition, it is purely data dependent and requires minimal external involvement during MRA process (Dragomiretskiy and Zosso, 2014). Consequently, Fu et al. (2021) developed a hybrid model by integrating VMD method with the grey wolf optimizer (GWO) algorithm and SVM, leading to the VMD-GWO-SVM model. They found that the VMD-GWO-SVM model had the best performance in estimating ET without regional meteorological monitoring in the southeastern margins of the Tengger Desert, China. With the need for better MRA tools, (ur Rehman and Aftab, 2019) proposed a cutting-edge MRA tool, i.e., the Multivariate variational mode decomposition (MVMD). The MVMD is inherently different and superior to VMD, which can only obtain univariate oscillations from a single time series. The MVMD builds on to its predecessor (i.e., VMD), having better capability in handling multichannel data. The key advantage of MVMD is its ability to directly obtain multivariate modulated oscillations from the multivariate input data by precisely looking into multidimensional space where the multivariate signal resides (Rehman and Aftab, 2019). In addition, the MVMD preserves mode-alignment between similar frequency content across multivariate input data, which is essential to glean meaningful joint information related to nonstationary multivariate data (Rehman and Aftab, 2019). Capitalizing on these benefits, the MVMD has been successfully applied for concurrent decomposition MRA in mechanical and electrical engineering applications (Gu et al., 2020, Rahul et al., 2021) and for underwater acoustic signal predictions (Yang et al., 2020). Yet,

this advanced concurrent decomposition MRA tool has not been used for energy and wave power applications in particular.

Until now, MVMD-based concurrent decomposition MRA tools for modelling evapotranspiration forecasting problems had never been. In addition, another novelty of the paper is the development of a Boosted regression tree (BRT) for evapotranspiration forecasting. Literature shows that its predecessor, the RF, has been successfully applied. However, as mentioned above, the RF-based models for forecasting ETo are very few. This study extends the previous one by developing and applying the BRT models, a non-parametric algorithm combining regression trees with a boosting method (Friedman, 2001). This is beneficial since the BRT model does not take any assumptions about prior knowledge between the connections of independent and dependent variables (Saha et al., 2021). Yet, the BRT model considers each predictor's effect after accounting for the effects of other predictors. In addition, the boosting method utilized by the BRT model assigns weights to input data in subsequent trees to ensure that the data poorly modelled in previous trees are prioritised to be modelled in the newer trees and the sequential process continues. With this boosting approach, the BRT model can continuously improve the forecasting accuracy. This paper aims to appraise two MRA utilities, including the advanced MVMD and Multivariate Empirical Mode Decomposition (MEMD) incorporated into forecasting models that include BRT, Cascaded forward neural network (CFNN), Extreme Learning Machine (ELM) and Random Forest (RF). Integrating the MVMD with BRT led to the development of the MVMD-BRT model, while MEMD led to the MEMD-BRT model. The proposed MVMD-BRT model is benchmarked with similar MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF models. These models are developed based on data collected in two varied stations, i.e., Brisbane and Gympie in Australia. The research also aims to evaluate the effectiveness of models under different climatic conditions of the two sites under study. The next section of the paper outlines the study area and data, followed by a brief description of the modelling and development approach. Then, the results are presented with discussions and conclusions.

2. Materials and methods

2.1. Study Area and data description

The acquired meteorological data consisted of the input predictors' daily rainfall (i.e., Rain), minimum temperature (T_{mix}), maximum temperature (T_{max}), minimum relative humidity (RH_{min}), maximum relative humidity (RH_{max}), wind speed (WS), and solar radiation (Radn), and reference evapotranspiration (ET_o) as the objective/target variable. The time-step-interval of the dataset is daily. The Bureau of Meteorology, Australia (Webb, 2010) supplied the dataset for Station 1: Brisbane and Station 2: Gympie in Queensland State, Australia. The datasets were from 1st January 2009 to 29th May 2022, the most recent dataset for these sites. The missing values in the dataset were substituted by averaging values of the corresponding daily values to overcome this issue. Table 1 describes the summary of each input predictor along with the two stations' objective variable (i.e., ETo).

Further, the statistical summary in terms of minimum, maximum, range, mean, standard deviation, skewness, and kurtosis provides some basic analysis of the dataset for both station 1 and station 2 in Table 1. It can be noted that the daily ET_o has a minimum value of 0.6 and 0.7 for station 1 and station 2, respectively. For more details, refer to Table 1. Also, Fig. 1 shows the map of the study stations 1 and 2.

Also, in Fig. 1, the location of the studied stations is specified. Further analyses were carried out using the cross-correlation function. The relationship between the input predictor variables (i.e., Rain, T_{min}, T_{max}, RH_{max}, RH_{min}, WS, Radn) and ET_o was investigated to determine the correlation between the antecedent time-lagged inputs based on the cross-correlation function (CCF). The CCF is used to determine which information is more significant to contribute to the ET_o forecasting. The input variables may be related to past lags of the $ET_{o,}$ and the CCF can help identify the lags of the ET_o that might be valuable inputs. In other words, the goal may be to determine which variable is leading and which is lagging. The correlation matrix in Fig. 2 shows the significance of each input with daily reference evapotranspiration (ETo) for station 1 and station 2. The numbers marked with an asterisk represent the significant values of the cross-correlation of the inputs. Fig. 2 reveals that the one-lag at (t-1) associated with all the input predictors can be more effective than the other antecedent lags.

2.2. FAO-56 Penman-Monteith (FAO-PM56) equation

The reference ET_o can be computed by the FAO-PM56 relationship (Allen et al., 1998), which is a convenient method in different geographic and climatic conditions. The Food and Agricultural Organization (FAO) proposed the FAO-PM56 method using the following equation to calculate the ET_o values (Allen et al., 1998).

$$ET_O = \frac{0.484\Delta(Radn - G) + \gamma(900/(T_a + 273))U_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34U_2)} \quad (1)$$

Here, ET_0 is the reference evapotranspiration, Δ represents the slope of the saturation vapor pressure $\left(kPa^{\circ}C^{-1}\right)$, Radn represents the net solar radiation ($MJ/m^2/day$), G is soil heat flux density $(MJ/m2/day)$ γ denotes the psychrometric constant $\left(kPa^{\circ}C^{-1}\right)$, T is the daily average temperature $(^{\circ}C)$, U_2 is the mean value of wind speed (m/s), and e_a and e_s are the actual and saturation vapor pressure (kPa).

Table 1
Geographic coordinates and statistical description of the data at respective stations.

	Statistical index	Rain (mm)	T_{max} (°C)	T_{min} (°C)	RH_{max} (%)	RH_{min} (%)	WS (m/s)	Radn (W/m^2)	ET_o (mm)
Station 1: Brisbane	Minimum	0	12.6	2.6	27	6	0.21	0.53	0.6
	Maximum	228.4	41.2	28	100	98	7.12	37.1	8.6
	Range	228.4	28.6	25.4	73	92	6.91	36.57	8
	Mean	3.22	26.73	16.71	87.69	46.90	1.65	18.77	3.72
	Std. Deviation	12.39	3.78	4.52	8.22	13.49	0.64	6.59	1.54
	Skewness	8.64	−0.07	−0.29	−1.82	0.10	1.11	0.14	0.32
	Kurtosis	106.02	−0.36	−0.81	5.93	0.75	2.87	−0.57	−0.91
	Latitude: 27.4705°S Longitude: 153.0260°E								
Station 2: Gympie	Minimum	0	5	−1.8	36	4	0	1.13	0.7
	Maximum	239.2	42.4	25.4	992	100	6.14	37.18	9.6
	Range	239.2	37.4	27.2	956	96	6.14	36.05	8.9
	Mean	3.02	27.35	14.21	96.05	46.20	1.51	19.15	3.66
	Std. Deviation	11.02	4.33	5.75	13.63	15.75	0.79	6.45	1.61
	Skewness	8.76	0.12	−0.48	57.97	0.30	1.03	0.08	0.42
	Kurtosis	117.01	−0.13	−0.67	3816.47	0.21	1.55	−0.53	−0.67
	Latitude: 26.1836°S Longitude: 152.6624°E								

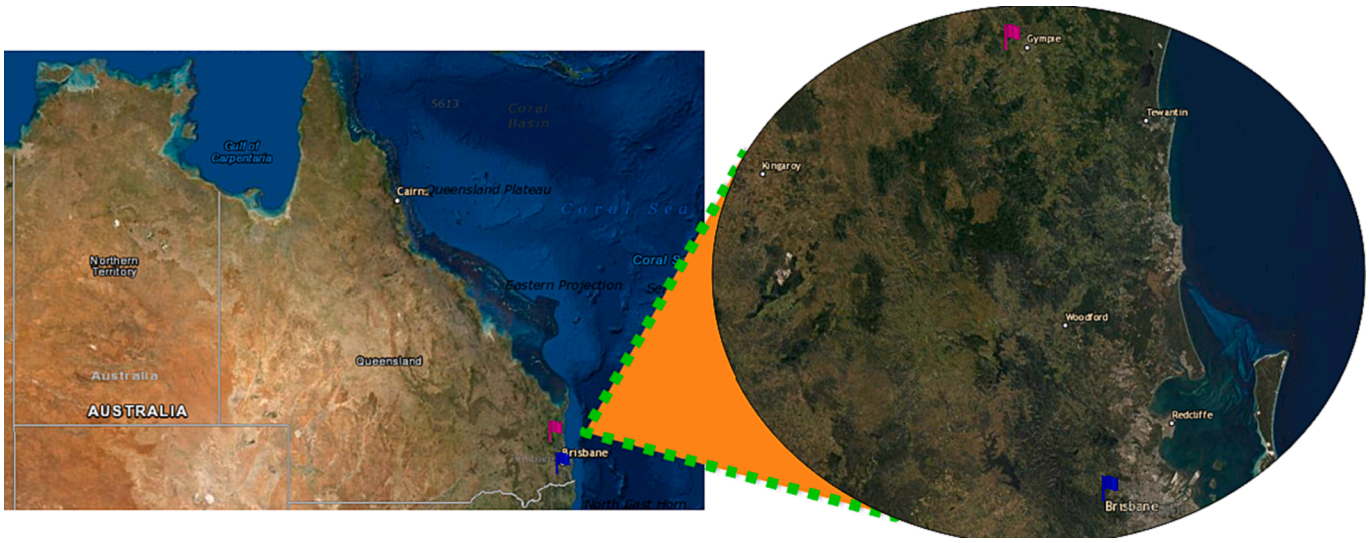


Fig. 1. Map of the stations.

2.3. Multivariate variational mode decomposition (MVMD)

The MVMD algorithm is a generic version of the variational mode decomposition (VMD) algorithm for multivariate or multichannel data sets, which for the first time presented, was developed by (ur Rehman and Aftab, 2019). The mechanism of classical VMD, as a univariate non-recursive decomposition technique, is decomposing a complex signal $x(t)$ into subset of modes (sub-signals) with fixed bandwidth called intrinsic mode decomposition (IMF) $IMF_k(t)$ and residual component (Dragomiretskiy and Zosso, 2013). Research shows that MVMD can capture the non-stationary and non-linearity of a multichannel signal simultaneously and, compared to bivariate empirical mode decomposition (BEMD), can avoid mixing mode (Gao and Shao, 2022). This pre-processing scheme can decompose a subset of signals, including C number time series $x(t) = \sum_{k=1}^{K} IMF_K(t) = [x_1(t); x_2(t); x_3(t); \cdots; x_k(t)]$ into predefined K number of multivariate $IMF_k(t) = [IMF_1(t); IMF_2(t); IMF_3(t); \cdots; IMF_k]$. In the aim of optimization of the MVMD, the mode number (K) can be selected by minimizing the bandwidths summation of the modes as follow (ur Rehman and Aftab, 2019):

Minimizing

$$\{u_{k,c}\}, \{\psi_k\} \left\{ \sum_k \sum_c \partial_t \left[u_+^{k,c}(t) e^{-j\psi_k t} \right]_2^2 \right\} \tag{2}$$

Subject to

$$\sum_k u_{k,c}(t) = x_c(t), c = 1, 2, C \tag{3}$$

where $u_+^{k,c}$ denotes the analytical modulation signal, $\{\psi_k\}$ is the set of center frequencies, ∂_t is the partial derivative operation, $\{u_{k,c}\}$ denotes the sets of all the number of modes (K), c is the channel, and k is the mode. The above optimization problem can be solved via the alternate direction method of multipliers (ADMOM) algorithm in the following steps (Gao and Shao, 2022):

1- Initializing the

$$\left\{\widehat{u}_{k,C}^1\right\}, \{\psi_k^1\} \text{and } \widehat{\Gamma}^1{}_C \tag{4}$$

Updating the Mode using the below relationship:

$$u\frac{n+1}{kc}(\psi) = \frac{\widehat{x}_c(\psi) - \sum i = k\widehat{u}_{i,c}(\psi) + \frac{\Gamma_c(\psi)}{2}}{1 + 2\alpha(\psi - \psi_k^n)^2}, c = 1, 2, \ldots, C, k = 1, 2, \ldots, K \tag{5}$$

in which α denotes the parameter of equilibrium associated with the needed data loyalty constraint and Γc is the Lagrange multiplier. Also, $\widehat{x}_c$ and $\widehat{u}_{i,c}$ denote the Fourier transform of the $x_c(t)$ and $u_{i,c}$, respectively.

2- Updating the center frequency (ω_k)

$$\psi_k^{n+1} = \frac{\sum c \int 0 \psi u \frac{n+1}{kc}(\psi)^2 d\psi}{\sum c \int 0 u \frac{n+1}{kc} \psi 2 d\psi}, k = 1, 2, \ldots, K \tag{6}$$

3- Updating the Lagrangian multipliers operators.

$$\Gamma_c^{n+1}(\psi) = \Gamma_c^n(\psi) + \omega\left(\widehat{x}_c(\psi) - \sum k u_{k,c}^{n+1}(\psi)\right), c = 1, 2, \ldots, C \tag{7}$$

where ω denotes the update parameter.

4- Repeating steps 2–4 till gaining the convergence.

2.4. Multivariate empirical mode decomposition (MEMD)

The empirical mode decomposition (EMD) for the first time was proposed by (Huang et al., 1998), which is widely used to forecast several engineering ML-based investigations (Ji et al., 2019, Jicheng et al., 2021). EMD algorithm, as a data-driven approach, was proposed to multiscale decompose the time series $x(t)$ and time–frequency analysis in a linear framework including intrinsic mode functions (IMFs). Basically, EMD decomposition method suffers from the mixing mode in high non-linear problem and the ensemble EMD can solve this issue to some extent (Rehman and Mandic, 2010). However, both schemes are only capable of decompose a univariate signal. Recently, a new self-adaptive extension of EMD scheme, namely multivariate EMD (MEMD) has been proposed by (Rehman and Mandic, 2010) which can simultaneous decompose the multiple signals and solving the mode mixing problem employing the white Gaussian noises (Prasad et al., 2019a). The structure of MEMD to handle the non-linear and non-stationary multivariate time series can be mathematically defined as (ur Rehman and Aftab, 2019):

$$x(t) = \sum_{i=1}^{l} C_i(t) + Res(t) \tag{8}$$

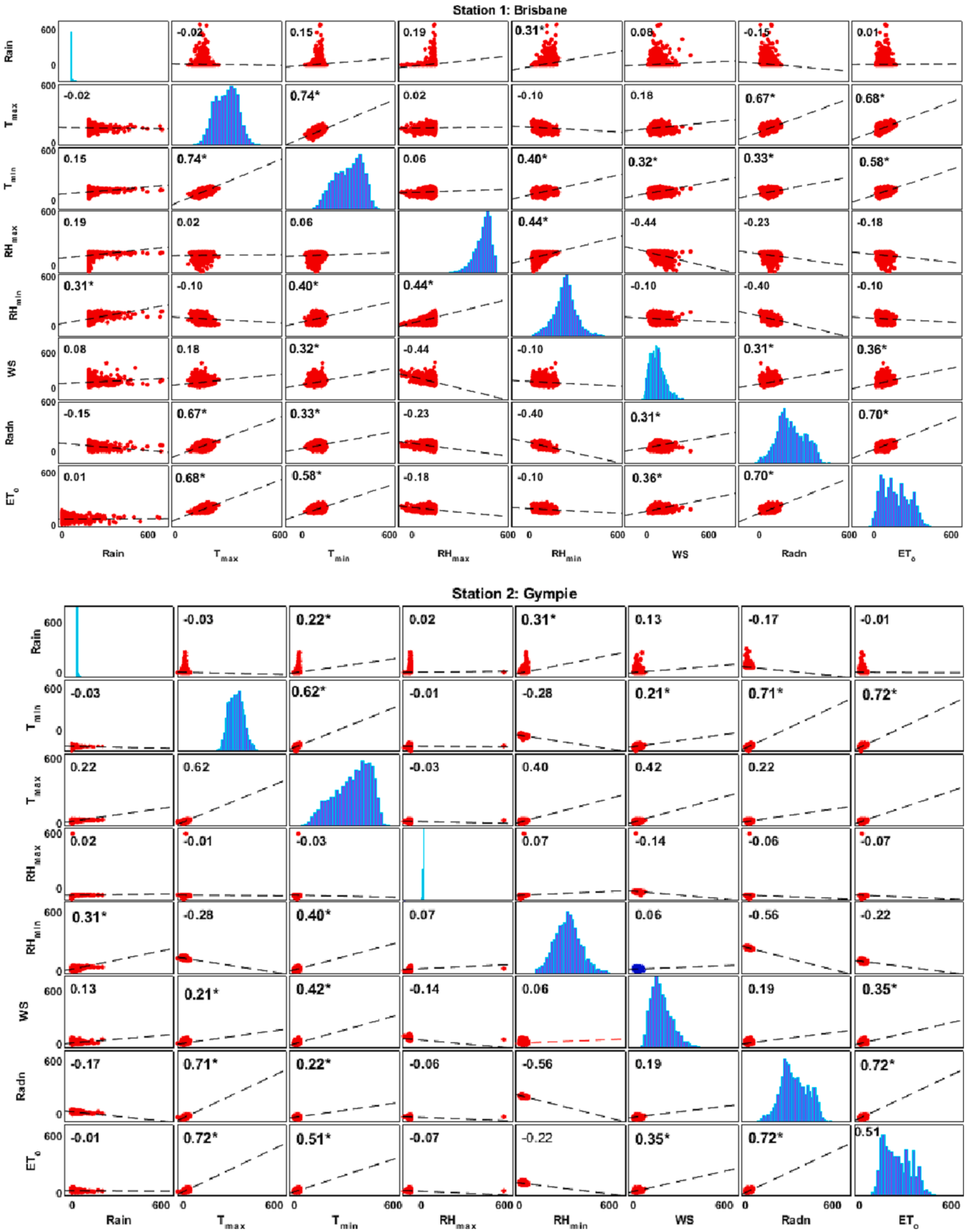


Fig. 2. Correlation matrix based on cross correlation indicates the significance of each input with daily evapotranspiration (ET_o) for station 1 and station 2. The (*) represents the significant numbers.

in which l is the modes number $C_i(t)$ denotes the set of l.

IMFs$\{C_i(k)\}_{i=1}^{l}$ and Res(t) denotes the monotonic residual. The local extrema identification, envelopes creation, and determining a local mean in higher dimensions are the significant problem in multivariate signals decomposition (Andersson et al., 2017). To solve this problem in MEMD technique, projection of the signal across several directions and averaging over the resulting envelopes are considered (Andersson et al., 2017). The MEMD by fixing $j = 1, 2, \cdots, \gamma$ direction vectors and create the multivariate envelope $e^j(t)$ computes the average value of the envelope curve based on the following formula (Andersson et al., 2017):

$$M(t) = \frac{1}{\gamma}\sum_{j=1}^{t} e^j(t) \tag{9}$$

In the next stage, similar to EMD, the detail $d(t) = x(t) - M(t)$ is computed. Afterward, the residual $\mathrm{Res}(t) = x(t) - d(t)$ are estimated after satisfying the stoppage criterion and repeat this process until all IMFs have been taken out of the signal (Andersson et al., 2017).

2.5. Boosted regression tree (BRT)

The BRT is a non-parametric model which does not presume prior information between the input and objective variables (Saha, Arabameri et al. 2021), and rather integrates boosting and regression trees (Friedman, 2001). The BRT is a type model which improved the performance accuracy by several individual models (Faskari et al., 2022). The BRT approach is mainly based on (*a*) CART regression tree, (*b*) the construction and assimilation of a series of models *via* boosting procedure which leads to a more precise and robust model. The BRT method solves the problem of the single decision tree's, which creates just the initial tree from the training data while the rest of the data is utilized to build the succeeding trees (Elith et al., 2008). Boosting techniques are employed to enhance the regression tree's forecasting ability. It resembles to the model averaging where the averaged results of several models are used, except that the boosting operation in a step-by-step manner to fit the models to a subset of the training set (Naghibi and Pourghasemi, 2015). The efficacy of the BRT is deeply reliant on two regularization parameters: (i) the number of additive terms or tress (n_t) and (ii) learning rate (LR). The LR parameter is used to minimize the impact of every single tree in the model which ranges 0.1——0.0001. The smaller LR value leads to a decreased loss function; but, this requires the presence of additional tress (n_t) to the model (Carty et al., 2015). This methodology poses several advantages, including the capacity to assess rapidly the large dataset which is less susceptible to overfitting (Westreich et al., 2010). Fig. 3 demonstrates the framework of the BRT model.

2.6. Cascaded forward neural network (CFNN)

The CFNN model, as described by Fahlman and Lebiere (Fahlman and Lebiere, 1989), is an artificial neural network (ANN) model variant. It employs a parallel information processing system consisting of three layers of neurons: input, hidden, and output. CFNN has a similar architecture to FFNN, except the input signal is coupled to each concealed layer behind it through a weight matrix. The distinction lies in the neurons of their hidden layer. A new hidden neuron is added to these networks at each successive stage. Each new neuron takes information from the input neurons and all previously hidden neurons before exiting to the input of each output neuron. In addition to the interactions

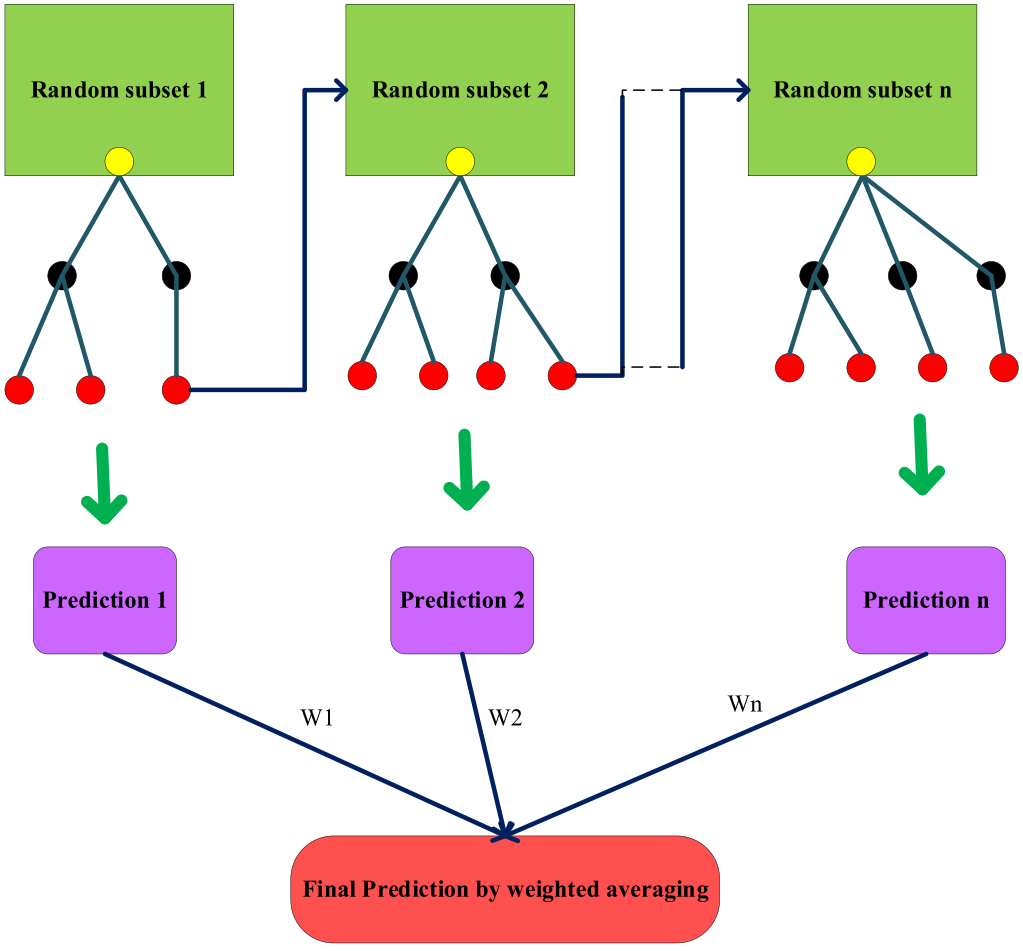


Fig. 3. The structure of the BRT method.

between hidden neurons, the input and output neurons are connected (Fig. 4).

Except for the first hidden layer, all hidden layers in CFNN contain at least two weight matrices, which are used to regulate the output signal of the top layer and the input signal of the network, respectively. This topology can provide more degrees of freedom to the training procedure, hence enhancing the network's non-linear mapping capability. The BP learning algorithm is used to optimize the weight matrix and bias matrix of a CFNN during training. Its objective is to bring the actual output of the network as near to the predicted output as possible, as measured by the mean square error. Before modeling, the network topology, including the number of hidden layers and neurons in each layer, must be defined for conventional neural networks such as MLP. Therefore, reliable detection of optimal design is frequently difficult and typically requires trial and error (Dharma et al., 2017). In the initial step, cascade networks are trained using input and output neurons, similar to classical networks. Training will terminate if the error is acceptable after a predetermined number of repetitions. If not, the model will be re-executed at each stage by adding a new neuron and appropriately training the network to decrease residual error (Fahlman and Lebiere, 1989). This training will continue until the error rate falls below the target threshold or the rate of change is slowed (Mohammadi et al., 2021). Fig. 4 depicts the topology of the CFNN model. The following is the governing equation for CFNN:

$$Out_{CFNN}(k) = f_{act}\left(\sum_{k=0}^{N}\left[HN(j)\times W_j(j,k) + I(i)\times W_i(i,k)\right] + b(k)\right) \tag{10}$$

where $Out_{CFNN}(k)$ is the output neuron, $W_j(j,k)$ and $W_i(i,k)$ are the vectors of weights, b(k) the bias weight, f_{act} is the activation function, $I(i)$is the input value, and $HN(j)$ is the hidden neuron.

2.7. Extreme learning Machine (ELM)

SLFN, also known as a feedforward artificial neural network with a single hidden layer, is the most prevalent type of ANN model and holds the capacity for universal approximation (Hornik et al., 1989). Training the SLFN is the most crucial aspect of the success of ANN models, and a great deal of work has been done to improve the SLFN model's learning process. Huang et al. presented ELM, a novel training method for the SLFN (Huang et al., 2006). Using the typical backpropagation training method, the training process is accomplished by updating weights and biases between the input and hidden layers and between the hidden and output layers. One of Huang et al. innovative ELM approaches is to split the training process into two phases: weights between input neurons and biases are generated randomly. In contrast, output weights can be analyzed analytically using Moore–Penrose generalized inverse matrix, which increases performance and decreases training times for large volumes of data (Huang et al., 2011). The ELM has been successfully applied in several emerging areas (Ali et al., 2018, Ali and Prasad, 2019). The numerical expression of a single feed-forward neural network (SLFN) is as follows:

$$\sum_{i=1}^{L} B_i g_i(\alpha_i x_t + \beta_i) = z_t, i = 1, 2.\cdots, N \tag{11}$$

Where L is the number of the hidden node, $g_i(\alpha_i x_t + \beta_i)$ is the hidden layer output function, α_i and β_i are hidden node parameters that are randomly determined and $i = 1, 2, 3, \cdots L$. B_i is the weight factor connecting the ith hidden nodes and the output node, and z_t is ELM model output. The weight vector elements are:

$$\alpha_i = [\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{in}]^T \tag{12}$$

The ELM model's hidden node parameters can be created randomly without knowing the training data or repeatedly adjusting the hidden layer neurons one by one for the lowest mean square error. The above equation can be written concisely:

$$H \times B = z \tag{13}$$

$$B = [B_1, B_2, \cdots, B_L]^T \tag{14}$$

$$z = [z_1, z_2, \cdots, z_N]^T \tag{15}$$

$$H(\alpha_1, \cdots, \alpha_L, x_1, \cdots, x_N, \beta_1, \cdots, \beta_L) = \begin{bmatrix} g(\alpha_1.x_1 + \beta_1) & \cdots & (\alpha_L.x_1 + \beta_L) \\ \vdots & \ddots & \vdots \\ (\alpha_1.x_N + \beta_1) & \cdots & (\alpha_L.x_N + \beta_L) \end{bmatrix}_{N\times L} \tag{16}$$

Examining the above relationships, it can be seen that all parameters except the output weight B_i are fixed in the extreme learning machine model. Therefore, the goal is to find a solution for the output weight vector using linear equation (12). Although, in many cases, H is not a square matrix, there is no vector like B_i to apply to Equation (12). The conventional method for solving this problem is to find the minimum square value of $\widehat{B}$by the following equation:

$$min\|z - HB\| \tag{17}$$

The optimal solution with respect to the minimum norm is expressed in the following form:

$$B = Hz \tag{18}$$

His the Moore-Penrose generalized inverse of the H matrix. Since the number of training samples is usually greater than the number of hidden layer neurons, Equation (15) can be rewritten as follows:

$$\widehat{B} = \left(H^T H\right)^{-1} H^T z \tag{19}$$

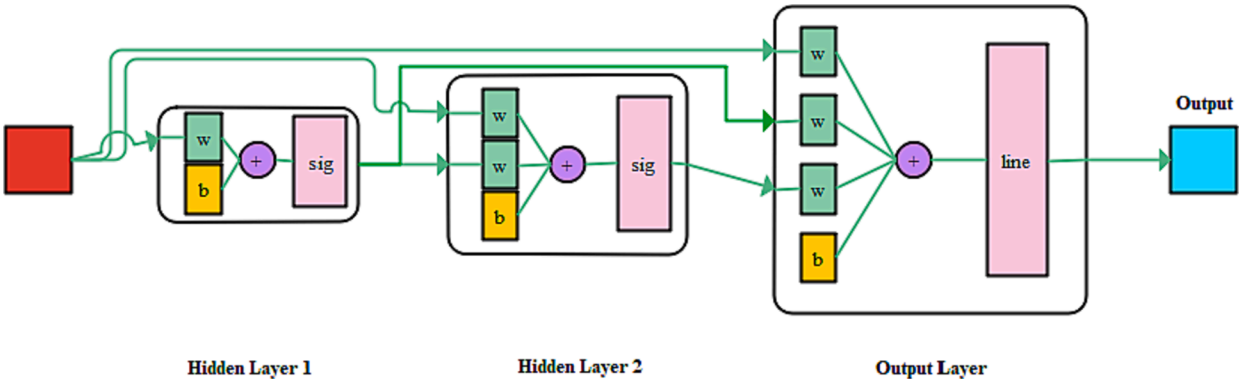


Fig. 4. Topology of CFNN model.

2.8. Random forest (RF)

Random Forest (RF) is a bagging (bootstrap aggregating) approach that creates a large number of regression trees separately created using a bootstrap sample of the dataset. Random Forest is one of the ensemble machine learning methods (Breiman, 2001). A regression tree divides the predictor space into nonoverlapping regions. The locations in the tree where the predictor space is partitioned are referred to as nodes, and the terminal nodes are called leaves. The collection of trees is known as a forest. Random forest is different from previous bagging methods because the regression tree is built on a new set of data and a random selection of predictors is applied to each node (Breiman, 1996).

Fig. 5 depicts the architecture of the RF method, where the input matrix X contains N samples and M input variables (sample set $S = [(x_i, z_i), i = 1, 2, ..., N], (X, Z) \in R^M \times R$). The bootstrap method is used to generate n sample tree sets from the sample set S. At each bootstrap sample, one-third of dataset S was utilized as out-of-the-bootstrap (OOB) data and the rest as in-bag data. For each sample set, the regression tree is modeled. All individual trees in the RF algorithm yield a prediction outcome. The ultimate prediction value is derived from the average performance of all individual trees. The definition of the prediction error is as follows (Liaw and Wiener, 2002):

$$MSE_{OOB} = \frac{\sum_{i=1}^{n_{tree}} \left(z_i - \hat{z}_i^{OOB}\right)^2}{n_{tree}} \tag{20}$$

$MSE_{OOB}$is the mean square error of the *OOB* data prediction, n_{tree} represents the number of trees, and y_i and $\hat{z}_i^{OOB}$ represent the actual value of the *OOB* data and the mean of all *OOB* forecasts, respectively. The RF approach has the highest capacity among ensemble methods for addressing classification and regression issues since it combines many basic regression trees to maximize prediction (Zaklouta and Stanciulescu, 2012, Ali et al., 2020).

2.9. Model performance evaluation

Performance evaluation is a crucial element in the creation of a model. It involves comparing the estimated values of models with their actual values using statistical measurement to see how well the suggested model simulates the actual output. In the current study, the forecasting values are compared to the actual values computed by the following indices: R (Correlation Coefficient) RMSE (Root Mean Square Error), MAE (Mean Absolute Error), WI_E (Willmott's Index of agreement) (Willmott, 1982), NS_E (Nash-Scuttle estimator) (McCuen et al., 2006), LM_E (Legates and MacCabe's)(Legates and McCabe, 1999), KGE (Kling-Gupta efficiency) (Gupta et al., 2009), RRMSE (Relative root mean squared percentage error), and RMAE (Relative mean absolute percentage error). The following equations describe the indices:

$$R = \frac{\sum_{i=1}^{N}\left(ObservedET_{o,i} - \overline{ObservedET_o}\right)\left(ForecastedET_{o,i} - \overline{ForecastedET_o}\right)}{\sqrt{\sum_{i=1}^{N}\left(ObservedET_{o,i} - \overline{ObservedET_o}\right)^2 \sum_{i=1}^{N}\left(ForecastedET_{o,i} - \overline{ForecastedET_o}\right)^2}} \tag{21}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(ObservedET_{o,i} - ForecastedET_{o,i}\right)^2} \tag{22}$$

$$WI_E = 1 - \frac{\sum_{i=1}^{N}\left(ObservedET_{o,i} - ForecastedET_{o,i}\right)^2}{\sum_{i=1}^{N}\left(\left|ObservedET_{o,i} - \overline{ObservedET_o}\right| + \left|ObservedET_{o,i} - \overline{ObservedET_o}\right|\right)^2} \tag{23}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|ObservedET_{o,i} - ForecastedET_{o,i}\right| \tag{24}$$

$$NS_E = 1 - \frac{\sum_{i=1}^{N}\left(ObservedET_{o,i} - ForecastedET_{o,i}\right)^2}{\sum_{i=1}^{N}\left(ObservedET_{o,i} - \overline{ObservedET_o}\right)^2} \tag{25}$$

$$KGE = 1 - \sqrt{(R-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \tag{26}$$

$$LM_E = 1 - \left[\frac{\sum_{i=1}^{N}\left|ForecastedET_{o,i} - ObservedET_{o,i}\right|}{\sum_{i=1}^{N}\left|ObservedET_{o,i} - \overline{ObservedET_o}\right|}\right] \tag{27}$$

$$RRMSE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\left(ForecastedET_{o,i} - ObservedET_{o,i}\right)}{ObservedET_{o,i}}\right| \times 100 \tag{28}$$

$$RMAE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\left(ForecastedET_{o,i} - ObservedET_{o,i}\right)}{ObservedET_{o,i}}\right| \times 100 \tag{29}$$

Where $ForecastedET_{o,i}$ is the forecasted evapotranspiration value and $ObservedET_{o,i}$is the actual evapotranspiration value. $\overline{ForecastedET_o}$ is the average of the forecasted results. $\overline{ObservedET}$ is the average of calculated values. N denotes the total number of samples collected. Moreover, α shows the relative variability of the forecasted and actual values, whereas β is the ratio between the forecasted and actual mean values. The WI_E displays the differences between forecasted and calculated means and variances, which reflect sensitivity to outliers in the observation data and insensitivity to additional and proportional variances between expected and calculated values. The value of LM_E and WI_E ranges from 0 to + 1, with + 1 being the ideal value. The NS_E is used to compare model performance (range from $-\infty$ to + 1), and its best value is 1. Regarding this metric, the performance of the model is scored as follows: great ($NS_E > 0.75$), good ($0.65 < NS_E < 0.75$), satisfactory ($0.50 < NS_E < 0.65$), acceptable ($0.40 < NS_E < 0.50$), and inadequate

($NS_E < 0.4$). KGE fluctuates between $-\infty$ and 1, and values near to 1 suggest reliable model predictions.

3. Model development

In this study, eight newly hybridized models MVMD-BRT, MVMD-CFNN, MVMD-RF, MVMD-ELM, MEMD-BRT, MEMD-CFNN, MEMD-RF, and MEMD-ELM were developed in the MATLAB R2019b environment to daily ET_o forecasting in Brisbane, and Gympie of Australia. All the input predictors are the Rain, T_{min}, T_{max}, RH_{min}, RH_{max}, Ws, and Radn, and the target variable is the ET_0, which all the signals in both stations are collected form 01/01/2009 to 25/05/2022. All the models were executed using an Intel Core i5-8400, 2.80 GHz CPU series, and 8 GB RAM. The workflow of the daily ET_0 forecasting in two regions of Australia is shown in Fig. 4. Further steps of the model development are described in detail as follows:

Step 1: Determination of lagged-time components

This step involved determining the correlation between the antecedent time-lagged input data (i.e., Rain, T_{min}, T_{max}, RH_{max}, RH_{min}, Ws, and Radn) and ET_o using the cross-correlation function (CCF). The results statistically demonstrated that the one-lag (t-1) associated with all the predictors in both stations can be more effective than the other antecedent information.

Decomposition of predictors using the MEMD and MVMD

This is the primary pre-processing step of the proposed comparative hybrid modelling approaches. The MEMD and MVMD decomposition techniques are capable of simultaneously decomposing the input predictors in opposite to the classical EMD and VMD techniques, which are necessary to make separate forecasting for each of the IMF or residual and finally, all the forecasts are added together (Ali et al., 2020, Jamei et al., 2022). The number of modes (IMFs) in this study was optimized based on a trial-and-error process. The optimal mode number (K) for Brisbane and Gympie stations was obtained equal to ten. All the components for each predictor, by taking into account the residual were MVMD = 11 (IMFS) and MEMD = 11 (IMFS), and all the subsequences for each station (considering 7 predictors) have attained MVMD (7@($11 \times IMF_S + Res$) = 77), and MEMD (7@($11 \times IMF_S + Res$) = 77). The setting parameters of the MVMD and MEMD schemes are summarized in Table 2.

Preparations of models feeding

The decomposed subsequences (i.e., IMFs and Res) are fed directly into the ML models. For this aim, the rate of splitting datasets is 70 % for training and 30 % for the testing phase, respectively, to create ET_o forecasting AI-based models (Fijani et al., 2019). The normalization and denormalization procedures for the training and testing sets were also adopted to unify the data scale and speed up the convergence of the ML-based frameworks. Since the main objective of this research study is to explore how the pre-processing procedure affects enhancing the accuracy of daily forecasting of the ET_o, it is crucial to compare the performance of the proposed hybrid models with the counterpart standalone

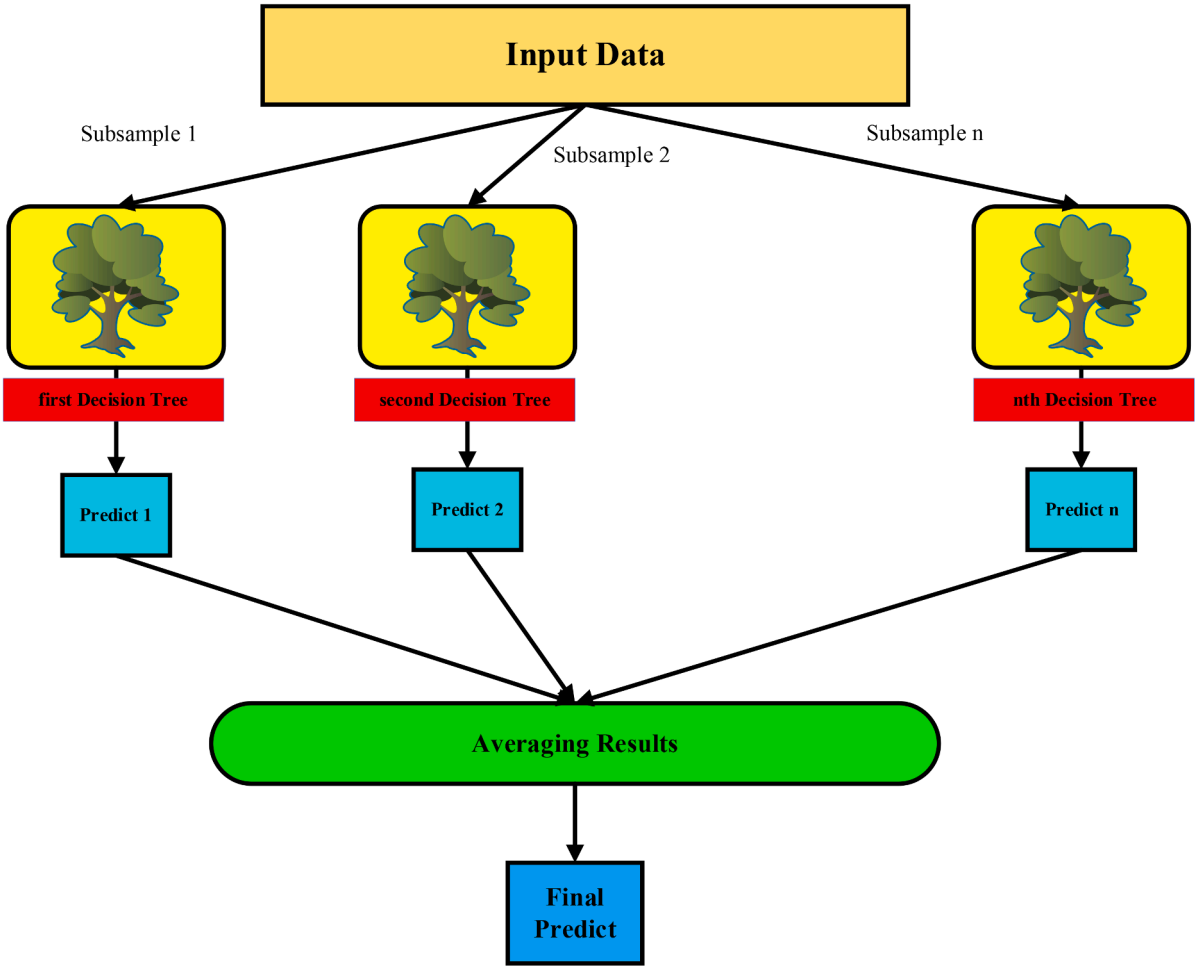


Fig. 5. Basic diagram of the RF model.

Table 2
Design parameters of MEMD and MVMD methods involved in decomposing the data into IMFs and residuals for each station.

Inputs	Station 1: Brisbane				Station 2: Gympie			
	No. of total projections	Stop vector		No. of IMFs & Residuals	No. of total projections	Stop vector		No. of IMFs & Residuals
		tolerance values	threshold			tolerance values	threshold	
MEMD method								
Rain	8	[0.8 0.8]	0.8	11	8	[0.7 0.7]	0.7	11
T_{max}	8	[0.8 0.8]	0.8	11	8	[0.7 0.7]	0.7	11
T_{min}	8	[0.8 0.8]	0.8	11	8	[0.7 0.7]	0.7	11
RH_{max}	8	[0.8 0.8]	0.8	11	8	[0.7 0.7]	0.7	11
RH_{min}	8	[0.8 0.8]	0.8	11	8	[0.7 0.7]	0.7	11
WS	8	[0.8 0.8]	0.8	11	8	[0.7 0.7]	0.7	11
Radn	8	[0.8 0.8]	0.8	11	8	[0.7 0.7]	0.7	11

	Station 1: Brisbane						Station 2: Gympie					
	α	τ	DC	Init	tol	k	α	τ	DC	Init	tol	k
MVMD method												
Rain	2000	0	0	1	1e-7	11	2000	0	0	1	1e-7	11
T_{max}	2000	0	0	1	1e-7	11	2000	0	0	1	1e-7	11
T_{min}	2000	0	0	1	1e-7	11	2000	0	0	1	1e-7	11
RH_{max}	2000	0	0	1	1e-7	11	2000	0	0	1	1e-7	11
RH_{min}	2000	0	0	1	1e-7	11	2000	0	0	1	1e-7	11
WS	2000	0	0	1	1e-7	11	2000	0	0	1	1e-7	11
Radn	2000	0	0	1	1e-7	11	2000	0	0	1	1e-7	11

Table 3
Parameter setting for the ML model forecasting of the ETo.

Stations	Models	Tuning parameter models
Brisbane	BRT	Hybrid and Standalone Structure • Learn rate = 0.194, Method = LSBoost, • Combine weight = Weighted Sum, • Learner name = Tree,
	ELM	Hybrid and Standalone Structure • Number of hidden neurons = 49; 59, • Activation functions = Sigmoidal,
	RF	Hybrid and Standalone StructureNumber of trees (ntree) = 1000,Prediction tree (mtry) = 2; 14,
	CFNN	Hybrid Structure: 78-9-1; Standalone Structure: 7-9-1 Epoch = 12 iterations, Validation checks = 6, Mu = 0.001, Training = Levenberg-Marquadt
Gympie	BRT	Hybrid and Standalone Structure • Learn rate = 0.1940, Method = LSBoost, • Combine weight = Weighted Sum, • Learner name = Tree,
	ELM	Hybrid and Standalone Structure • Number of hidden neurons = 40; 49; 50, • Activation functions = Sigmoidal; Sine,
	RF	Hybrid and Standalone StructureNumber of trees (ntree) = 1000,Prediction tree (mtry) = 2; 14,
	CFNN	Hybrid Structure: 78-9-1; Standalone Structure: 7-9-1 Epoch = 11 iterations, Validation checks = 6, Mu = 0.001, Training = Levenberg-Marquadt

The training accuracy of the novel MVMD-BRT against benchmarking models for both stations has been presented in Table 4 using *R* and *MSE* metrics. The MVMD-BRT attained the values of ($R = 0.9792$, $MSE = 0.100$) for station 1 and ($R = 0.9802$, $MSE = 0.105$) for station-2 as compared to other models to forecast daily ET_0.

models (CFNN, ELM, BRT, and RF). Therefore, it is necessary to mention that for developing the standalone schemes, the normalized antecedent information of the predictors on the previous day (t-1) can be used to feed the ML methods.

Tuning machine learning approaches

one of the most crucial stages in making an ML-based prediction/

Table 4
Training accuracy of the MVMD-BRT, MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF models based on *MSE* and *R* metrics.

Station 1: Brisbane			Station 2: Gympie	
	MSE	***R***	***MSE***	***R***
MVMD-RF	0.066	0.9883	0.077	0.9876
MEMD-RF	0.071	0.9869	0.081	0.9864
RF	0.151	0.9724	0.169	0.9713
MVMD-CFNN	0.378	0.9179	0.314	0.9388
MEMD-CFNN	0.362	0.9224	0.347	0.9320
CFNN	0.764	0.8245	0.780	0.8397
MVMD-ELM	0.525	0.8832	0.485	0.9035
MEMD-ELM	0.494	0.8905	0.574	0.8848
ELM	0.764	0.8244	0.778	0.8400
MVMD-BRT	0.100	0.9792	0.105	0.9802
MEMD-BRT	0.129	0.9730	0.145	0.9727
BRT	0.459	0.8997	0.472	0.9071

forecasting model is adjusting the hyperparameters, which significantly affects their results and accuracy. There are several approaches to gain the optimal setting (hyper) parameters, including traditional trial and error procedure (Rehamnia et al., 2020), grid search (Shahsavar et al., 2021), random search and *meta*-heuristic algorithm. The current research found the best hyperparameters using the grid search methodology with the RMSE as the convergence criterion in the MATLAB environment. Details on setting up the parameters during the multi-step forecasting of daily ET_o are summarized in Table 3. According to Table 3, the BRT, as the main model, is a tree-based model, in which within the model, the Learn rate value and ensemble strategy (LSBoost) are the significant adjustments, whereas, in the ELM and CFNN models, the neurons number in the hidden layer are the important hyper parameters. Also, in the active function, ELM models (hybrid and standalone) are "sigmoidal" and "sine" (Karbasi et al., 2021). In contrast, in the CFNN model, the best training algorithm for all the models is "Levenberg-Marquadt" (Zhang et al., 2021). Finally, RF model based on the Number of trees (ntree = 1000) and Prediction tree (mtry = 2 and 14 for standalone and hybrid models, respectively) achieved the best results. Fig. 6 shows the schematic view of the modelling framework designed in this work.

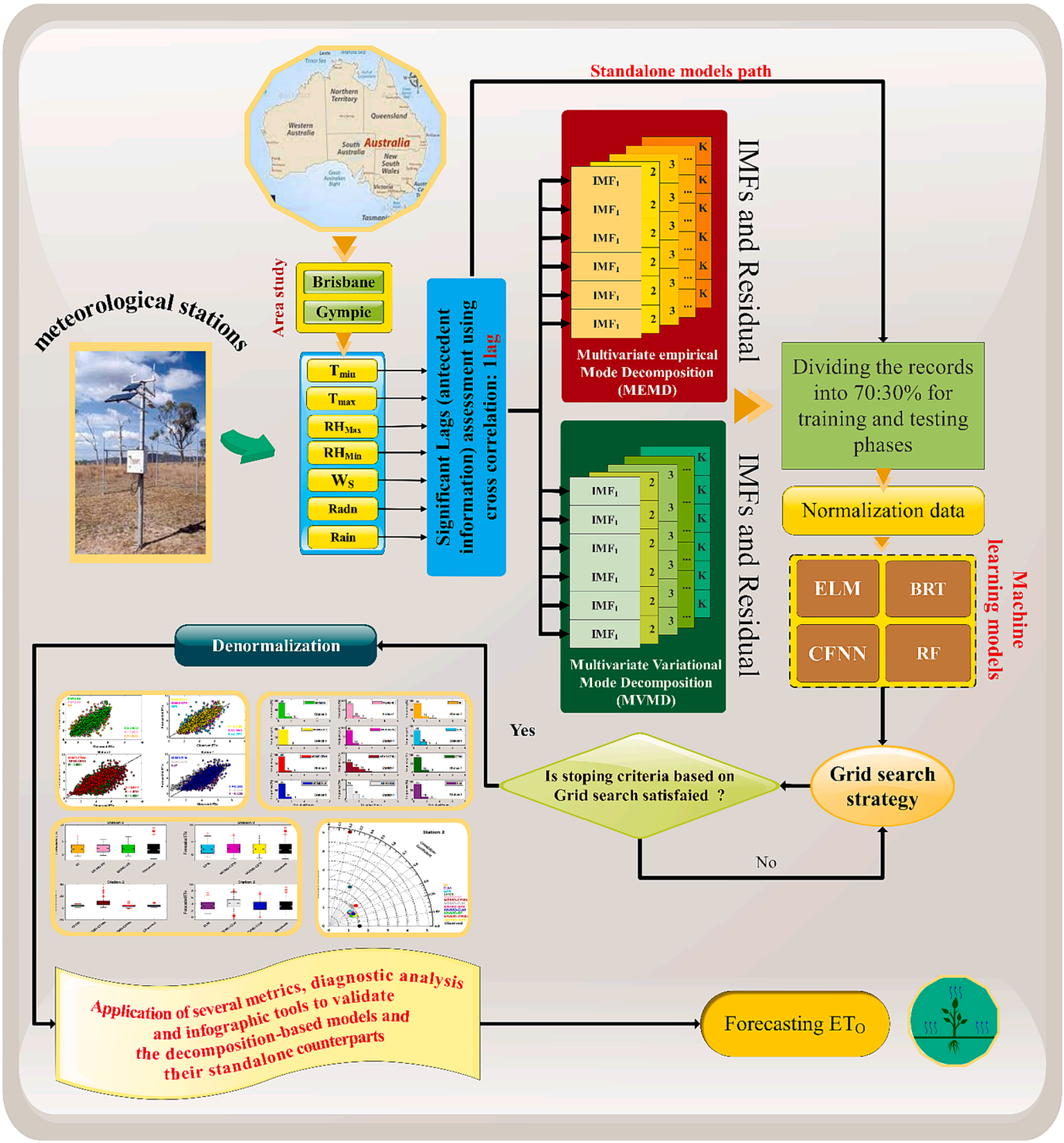


Fig. 6. Schematic diagram of the modelling strategy.

4. Application results and analysis

The proposed data decomposition-based MVMD-BRT is benchmarked in contrast to the MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF models using r, RMSE, MAE, WI_E, NS_E, LM_E, KGE, RMAE (%) and RRMSE (%) assessment metrics and diagnostic plots to forecast daily ETo for Station 1: Brisbane and Station 2: Gympie in the state of Queensland, Australia.

(a) Interpretation *via* tabular form

The proposed data decomposition based MVMD-BRT model is better than MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD- BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF models for both stations in terms of highest *R* and lowest *RMSE* and *MAE* errors. For Station 1, these metrics are: MVMD-BRT ($R = 0.9181$; $MAE = 0.605$ mm; $RMSE = 0.449$ mm), followed by MVMD-CFNN ($R = 0.9148$; $MAE = 0.616$ mm; $RMSE = 0.448$ mm) and MVMD-RF ($R = 0.9022$; $MAE = 0.674$ mm; $RMSE = 0.501$ mm), MEMD-BRT ($R = 0.8961$; $MAE = 0.706$ mm; $RMSE = 0.537$ mm), MVMD-ELM ($R = 0.8833$; RMSE = 0.716 mm; MAE = 0.548 mm) and so on. Table 5 results show that MVMD-based models (i.e., MVMD-BRT, MVMD-CFNN, etc.) perform better than

Table 5
Testing performance of the **MVMD-BRT** against MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF models using *R*, *RMSE (mm),* and *MAE (mm)*. The 30% data was used to test the models.

Station 1				Station 2		
	R	***RMSE (mm)***	***MAE (mm)***	***R***	***RMSE***	***MAE***
MVMD-RF	0.9022	0.674	0.501	0.9020	0.692	0.488
MEMD-RF	0.8693	0.792	0.609	0.8764	0.835	0.648
RF	0.8217	0.868	0.633	0.8289	0.884	0.619
MVMD-CFNN	0.9148	0.615	0.448	0.7888	1.100	0.534
MEMD-CFNN	0.6735	1.532	1.186	0.2109	7.785	5.957
CFNN	0.8235	0.864	0.630	0.4599	2.150	0.672
MVMD-ELM	0.8833	0.716	0.548	0.8907	0.718	0.534
MEMD-ELM	0.8016	1.009	0.793	0.7375	1.277	1.032
ELM	0.8249	0.860	0.625	0.8296	0.882	0.621
MVMD-BRT	**0.9181**	**0.605**	**0.449**	**0.9109**	**0.652**	**0.464**
MEMD-BRT	0.8961	0.706	0.537	0.8642	0.799	0.596
BRT	0.7972	0.921	0.657	0.8169	0.912	0.634

MEMD-based models in Station 1 to forecast daily ETo.

For Gympie Station 2, the MVMD-BRT again turned out to be the most accurate model by acquiring (R = 0.9109; RMSE = 0.652 mm; MAE = 0.464 mm) with MVMD-RF (R = 0.9020; RMSE = 0.692 mm; MAE = 0.488 mm) as the second-best model, following by MVMD-ELM (R = 0.8907; RMSE = 0.718 mm; MAE = 0.534 mm) and so on. For more details, we refer to Table 5, which proves that the MVMD-BRT model is reasonably better in forecasting daily ETo as compared to MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF models. Table 5 also proves that the MVMD significantly improved the performance accuracy of the models, especially when hybridized with BRT to construct MVMD-BRT models compared to the MEMD-based models. Further, the MVMD-BRT model is superior to the standalone counterpart models (i.e., BRT, CFNN, RF, and ELM) to forecast daily ETo. Overall, the MVMD-BRT model is better for forecasting ETo for Stations 1 and 2.

The model's performance was further examined in Table 6 using the WI_E, NS_E, LM_E, and KGE assessment metrics. The better forecasting accuracy yielded by the MVMD-BRT model for station 1 (WI_E = 0.9070, NS_E = 0.8421, LM_E = 0.6529, KGE = 0.8792), and station 2 (WI_E = 0.8966, NS_E = 0.8396, LM_E = 0.6521, KGE = 0.8803) against the benchmark comparing MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF models. Here again, the analysis in Table 6 confirmed that the hybrid version of MVMD is better in precision than the MEMD-based models and standalone counterparts. Based on WI_E, NS_E, LM_E and KGE metrics, it is apparent that the MVMD-BRT model has improved analytical capacities to perform accurate forecasts. This is justified by the statement that the models documented an $NS_E \leq 0.800$ are supposed to be 'unsatisfactory'. At the same time, NS_E ranges between 0.800——0.900, the models are considered 'fairly good', and the models registering $NS_E \geq$ 0.900 are deemed 'very satisfactory' (Shamseldin, 1997). Therefore, based on these assessment metrics, the proposed MVMD-BRT model can be classified as 'fairly good and satisfactory' at both stations in forecasting daily ETo.

A clear difference in the precision is being examined and computed in Tables 5 and 6. However, an essential limitation of these metrics is their deficits in comparing the models at geographically diverse locations, whereby the relative error metrics are the best choice. The relative error metrics in percentage (Table 7) confirmed that the MVMD-BRT model acquired the lowest *RRMSE (%)* against MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF,

Table 7
Geographic comparison of the accuracy of the **MVMD-BRT** against MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF models based on RRMSE (%) and RMAE (%). Note that the best model is boldfaced (black). The 30% data was used to test the models.

Station 1			Station 2	
	RRMSE %	***RMAE %***	***RRMSE %***	***RMAE %***
MVMD-RF	18.30	16.30	19.34	16.53
MEMD-RF	21.49	21.11	23.33	23.29
RF	23.57	20.65	24.72	20.71
MVMD-CFNN	16.70	14.02	30.75	17.04
MEMD-CFNN	41.59	42.56	217.66	209.90
CFNN	23.44	21.13	60.11	22.08
MVMD-ELM	19.43	17.82	20.08	17.85
MEMD-ELM	27.40	25.95	35.71	38.40
ELM	23.36	20.68	24.66	20.91
MVMD-BRT	**16.42**	**13.73**	**18.22**	**14.91**
MEMD-BRT	19.15	18.05	22.34	19.94
BRT	25.01	21.32	25.49	21.00

Table 6
The performance of **MVMD-BRT** *vs.* benchmarking comparing models based on WI_E, NS_E, LM_E, and KGE assessment metrics. Note that the best model is boldfaced (black). The 30% data was used to test the models.

Station 1					Station 2			
	WI_E	NS_E	LM_E	KGE	WI_E	NS_E	LM_E	KGE
MVMD-RF	0.8710	0.8038	0.6121	0.7793	0.8733	0.8081	0.6336	0.8017
MEMD-RF	0.7933	0.7294	0.5291	0.6900	0.7753	0.7206	0.5134	0.6412
RF	0.7911	0.6744	0.5105	0.7339	0.7956	0.6865	0.5356	0.7450
MVMD-CFNN	0.9048	0.8366	0.6534	0.8898	0.7530	0.5146	0.5990	0.7591
MEMD-CFNN	0.2409	−0.0137	0.0825	0.5938	−0.2346	−23.3174	−3.4703	−1.9120
CFNN	0.7872	0.6781	0.5122	0.7466	0.4040	−0.8546	0.4959	0.2699
MVMD-ELM	0.8587	0.7788	0.5763	0.8208	0.8759	0.7930	0.5993	0.8495
MEMD-ELM	0.8066	0.5601	0.3861	0.7871	0.5414	0.3454	0.2254	0.6931
ELM	0.7949	0.6804	0.5165	0.7512	0.7966	0.6879	0.5337	0.7523
MVMD-BRT	**0.9070**	**0.8421**	**0.6529**	**0.8792**	**0.8966**	**0.8296**	**0.6521**	**0.8803**
MEMD-BRT	0.8495	0.7851	0.5848	0.8251	0.8335	0.7439	0.5525	0.8274
BRT	0.7735	0.6334	0.4917	0.7392	0.7888	0.6666	0.5241	0.7551

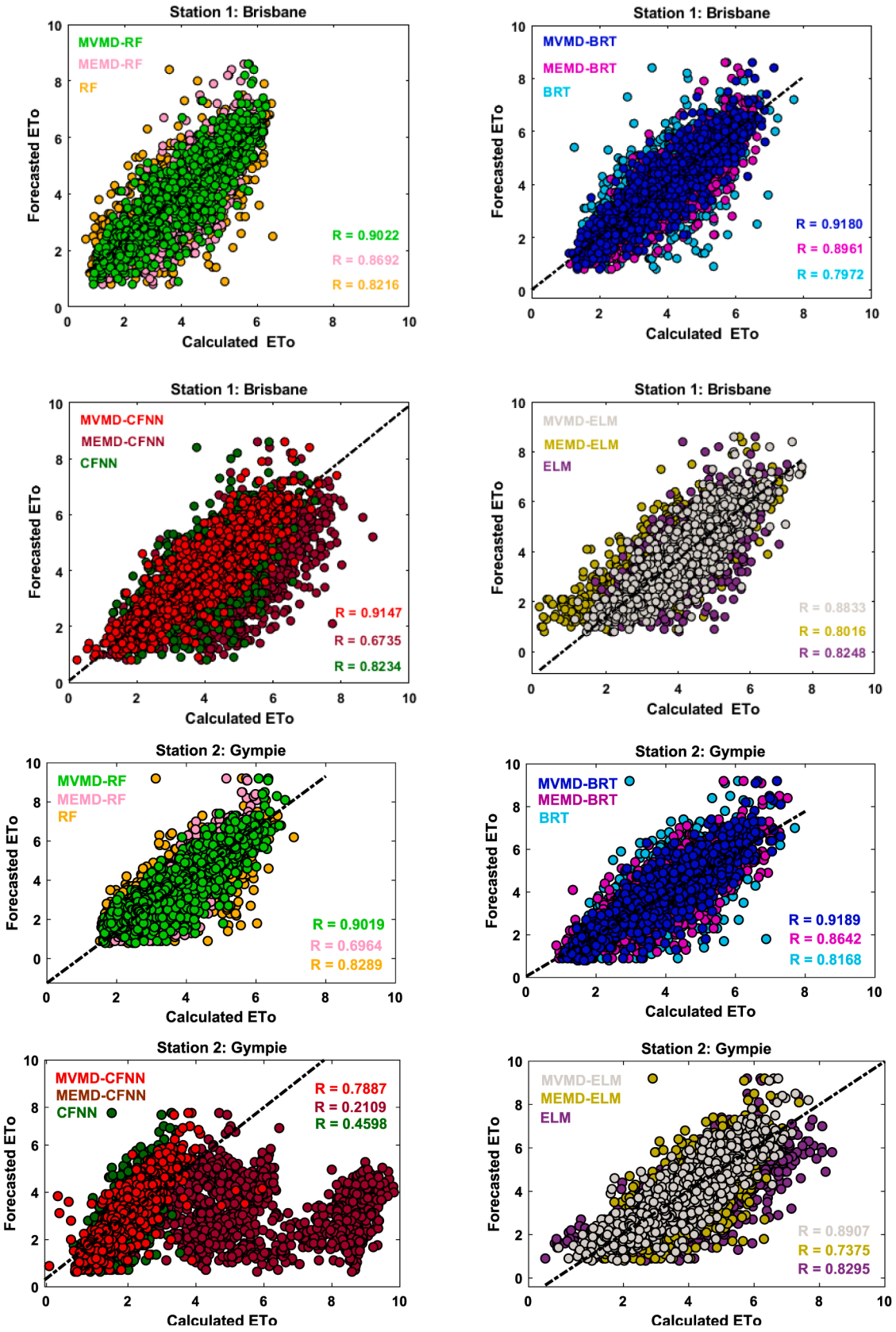


Fig. 7. Scatter chart between the forecasted and calculated ET_o using MVMD-BRT and benchmark comparing models.

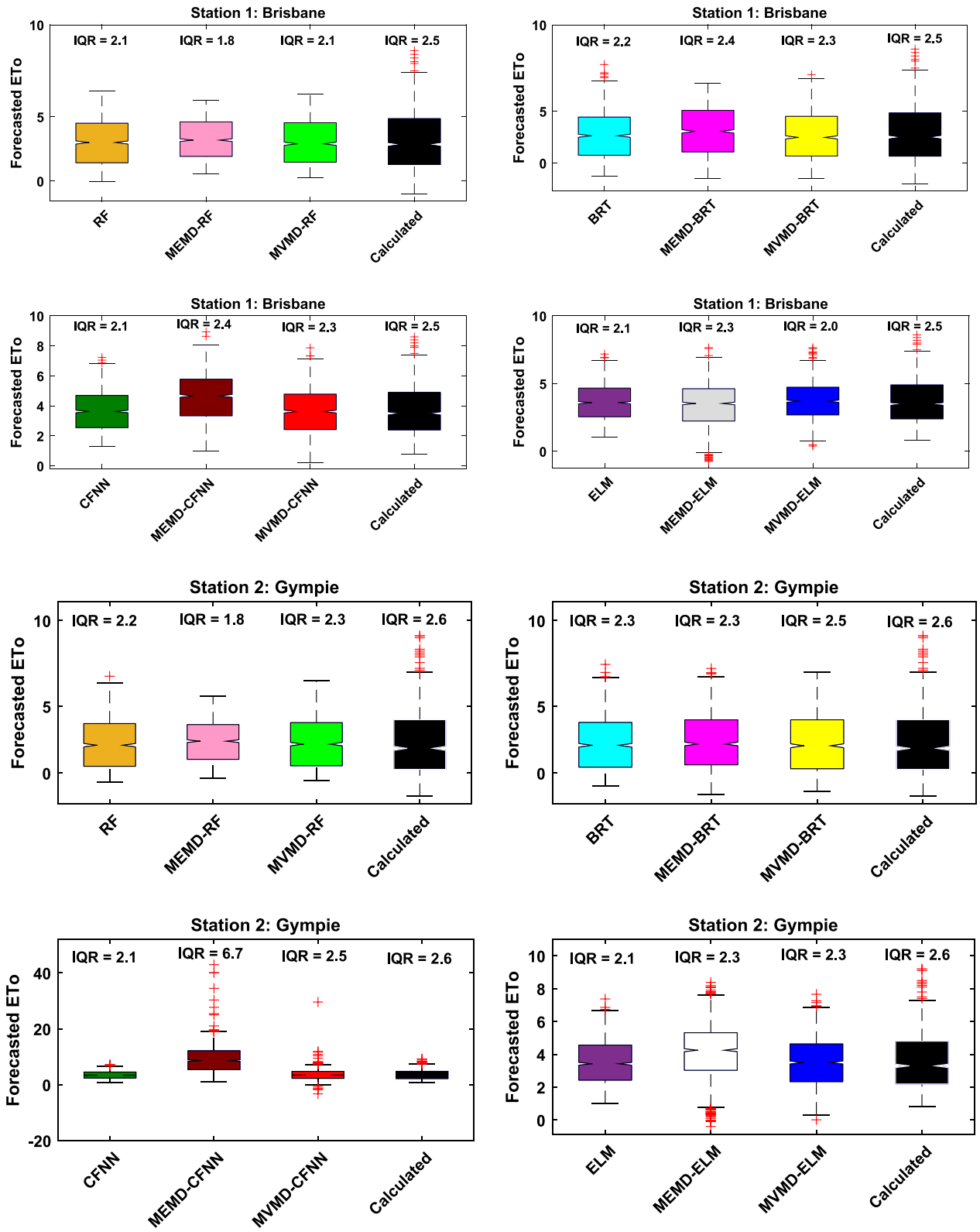


Fig. 8. Boxplot of the calculated and forecasted ET_o generated by the MVMD-BRT *vs.* MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF models. Here the IQR represents the interquartile range.

BRT, CFNN, ELM, and RF models for Station 1 and Station 2. More precisely, the *RRMSE (%)* and *RMAE (%)* when comparing the MVMD-BRT model with the least accurate counterpart, MEMD-CFNN, in the grouping [MVMD-BRT: MEMD-CFNN] were as follows: Station 1: [16.42 %, 13.73 %: 41.59 %, 42.56 %]; and Station 2: [18.22 %, 14.91 %: 217.66 %, 209.90 %]. Overall, the standalone RF, CFNN, ELM, BRT and hybrid models MEMD-RF, MEMD-CFNN, MEMD-ELM etc., appeared to have lower forecasting accuracy. Therefore, the *RRMSE (%)* and *RMAE (%)* depict that the MVMD-BRT model accomplished the highest accuracy for Station 1, followed by Station 2.

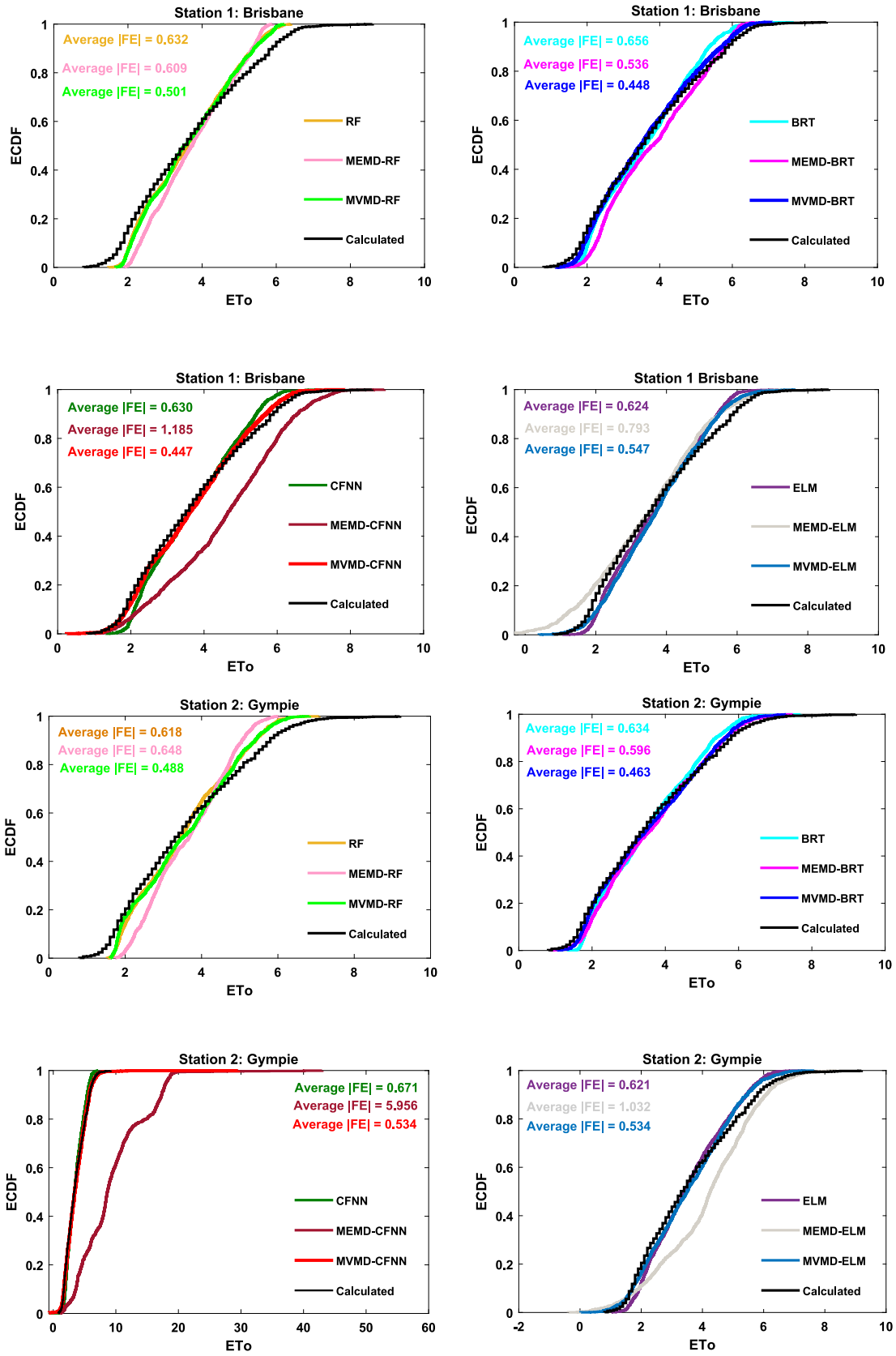


Fig. 9. Empirical cumulative distribution function (ECDF) of the forecasted and calculated ET_o generated by the MVMD-BRT *vs.* other benchmarking models for each station. The Average |FE| indicates the average forecasting errors between calculated and forecasted ET_o.

(b) Analysis *via* diagnostic plots and graphs

The scatter plots between the daily forecasted and calculated/actual ET_o using MVMD-BRT and benchmarking models (i.e., MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF) is presented in Fig. 7. For station 1, the forecasted and calculated ET_o for MVMD-BRT (yellow colour) with R = 0.9180 appeared to be the best model among all the other models. Similarly, for station 2, the MVMD-BRT model shows a reasonable degree of accuracy (R = 0.9189), followed by MVMD-RF (green colour), MVMD-ELM (blue colour), MEMD-BRT (pink colour) and so on. This again proves the MVMD-based hybrid (i.e., MVMD-BRT) is better in forecasting daily ET_0 as MVMD indeed enhances the forecasting capabilities by overcoming the non-linearity and non-stationarity issues in the data, which is also following Table 5, 6, and 7. Thus, for both station 1 and station 1, MVMD-BRT shows a higher precision accuracy than the benchmark comparing models.

An apparent discrepancy in efficiency of the MVMD-BRT model

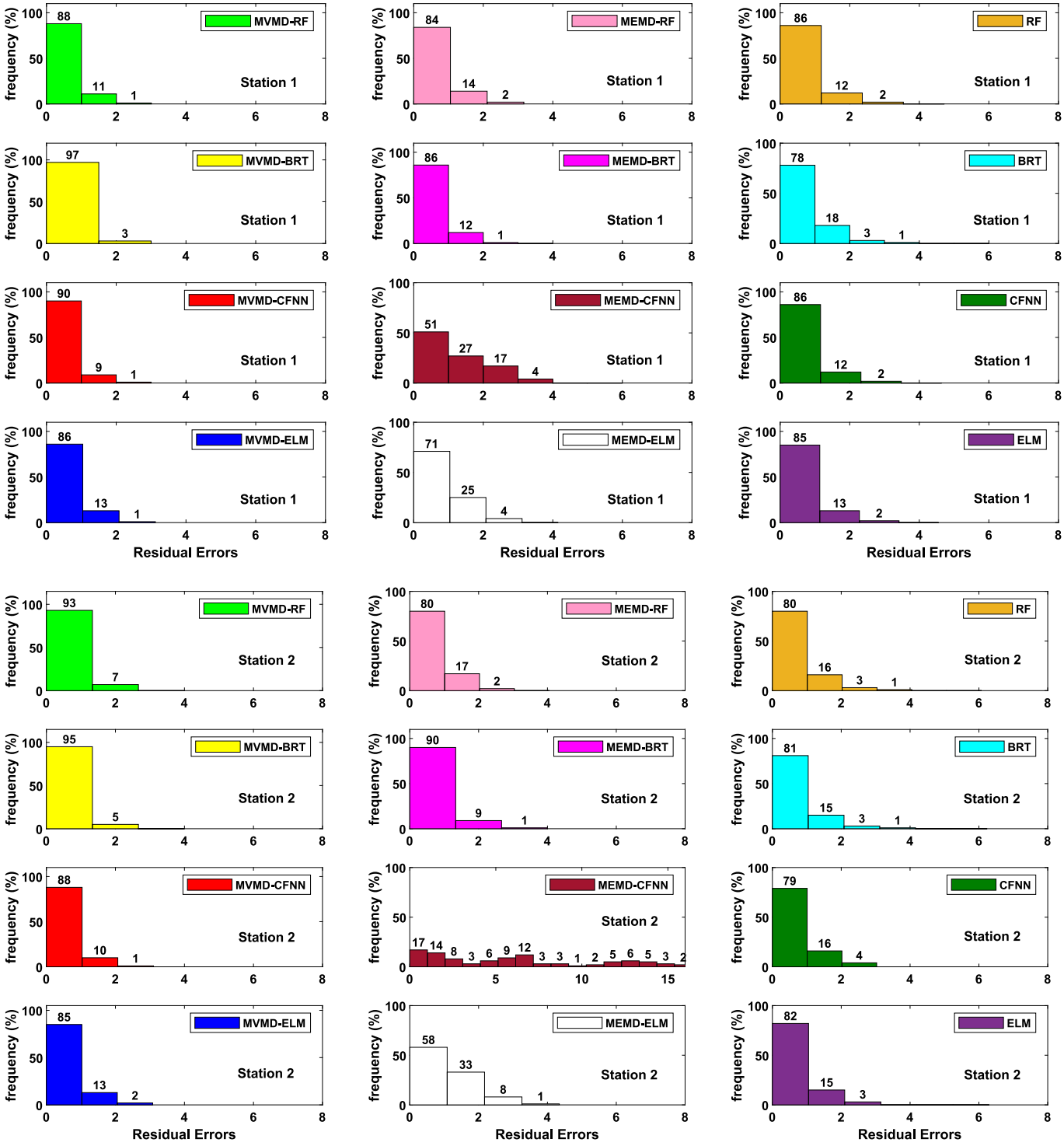


Fig. 10. Forecasted and calculated ET_o generated by the MVMD-BRT *vs.* MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF models using histogram frequency distribution plot.

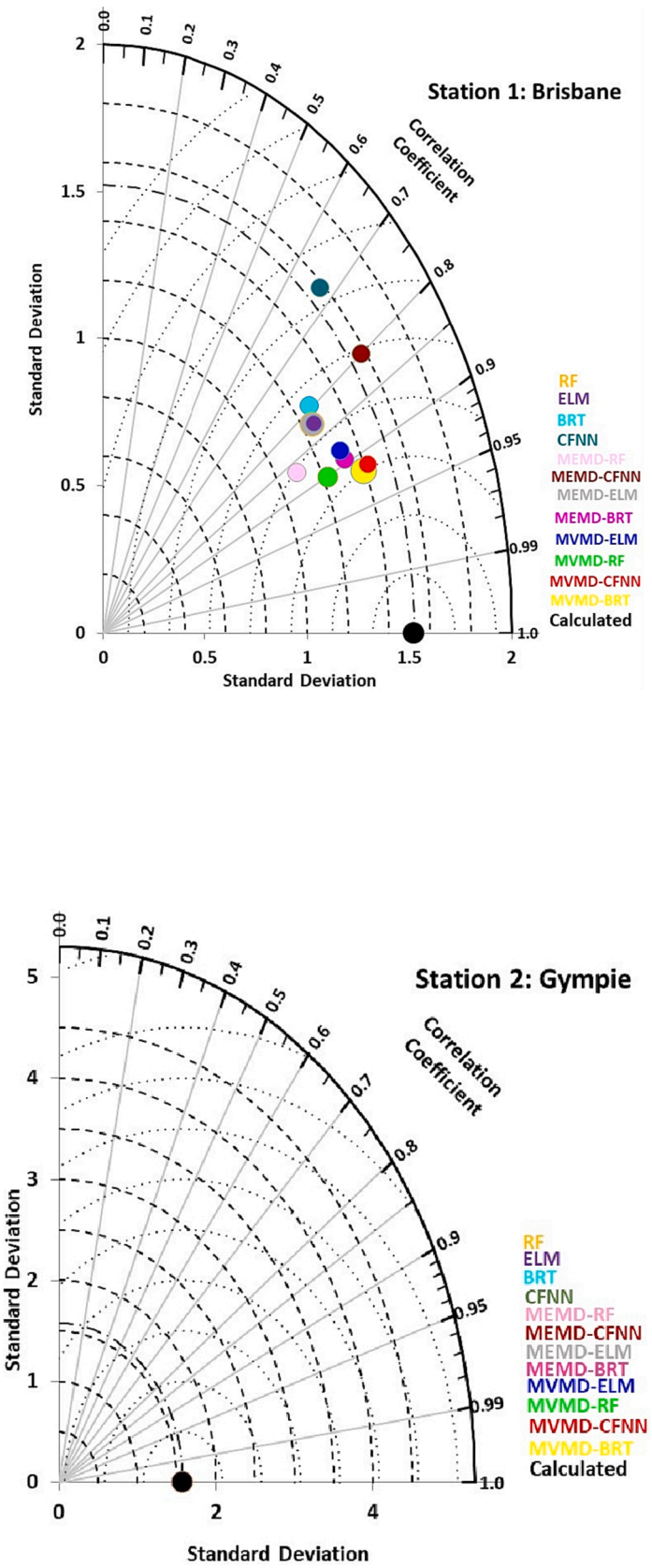


Fig. 11. Taylor diagram of the daily forecasted and calculated ET_o generated by the MVMD-BRT *vs.* the benchmarking models.

based on the boxplots in Fig. 8 expresses that the distribution calculated against forecasted ET_o yielded by MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF comparing models for station 1 and station 2 were dispersed exhibiting outliers. However, the boxplot distribution of the MVMD-BRT model has a very accurate and consistent representation of the forecasted and calculated ET_o along with the interquartile range IQR values = 2.3 (forecasted) and 2.5 (calculated) as compared to MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF models. Similarly, the boxplot distribution Alease appeared more accurate with MVMD-BRT (IQR = 2.5-forecasted, 2.6-calculated) ET_o than other models. Thus, the boxplots (Fig. 8) confirm the improved forecasting accuracy of the MVMD-BRT model compared to the benchmarking counterparts.

The empirical cumulative distribution function (ECDF) of forecasted and calculated ET_o maps in Fig. 9 represents the model's performance and the average forecasting errors, i.e., Average |FE| between estimated and forecasted ET_o. For both stations 1 and 2, the ECDF distribution of MVMD-BRT demonstrated a very close profile against MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF comparing models. The MVMD-BRT also produces a lower Average |FE| = 0.448 for Brisbane and 0.463 for Gympie stations compared to other models. Therefore, the ECDF plots (Fig. 9) further validate and verify the MVMD-BRT model's precise performance to forecast daily ET_0 in station 1 and station 2.

The histogram frequency distribution diagram in Fig. 10 plots the residual errors between the forecasted and calculated ET_o using the MVMD-BRT model against the MVMD-CFNN, MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, CFNN, ELM, and RF comparing models for both stations 1 and 2. Frequency distribution of residuals for MVMD-BRT were in a smaller error bracket varying between 0 and 2 mm with 97 % (at bin 1) and 3 %(at bin 2) frequency in station 1 while 95 % and 5 % for station 2 to forecast daily ET_o. Fig. 10 yet again endorses the proposed MVMD-BRT model's overall efficiency.

A Taylor diagram provides an orderly, systematic and comprehensive examination of the models' forecasting ability in a broader spectrum (Xu et al., 2016). Fig. 11 portrays a more quantifiable and convincing link between the forecasted and calculated ET_o based on correlation coefficient and standard deviation values. It is astonishingly proven that the MVMD-BRT model was lying close to the estimated ET_o data, suggesting the forecasting accuracy was noticeably better at both stations. In contrast, the CFNN and MEMD-CFNN models were parting in a far-off corner, confirming poor performance. For more details on other comparing models, refer to the Fig. 11, which affirms that the MVMD-BRT is reasonably accurate to forecast daily ET_o as compared to MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, ELM, and RF models.

5. Further discussion

In this research, the proposed hybrid MVMD-BRT model has been developed to forecast daily ET_o at station 1: Brisbane and station 2: Gympie in Queensland, Australia. The forecasting efficiency of the newly developed MVMD-BRT model is compared against MVMD-ELM, MVMD-RF, MEMD-BRT, MEMD-CFNN, MEMD-ELM, MEMD-RF, BRT, ELM, and RF models, which confirms its superiority to forecast daily ET_o, thus proving that the MVMD-BRT is a well-designed data intelligent model. The result of MVMD-BRT confirmed that the approach was effective in decomposing the input predictors by the MVMD process, boosting the BRT model's forecasting precision.

The novel MVMD-BRT model was an efficient approach to forecast daily ET_o using some standard good-of-fitness metrics and diagnostic plots, but some recommendations can be considered in the future. Here, only the historical meteorological input data were employed to design the MVMD-BRT model; however, the satellite-derived time-series predictors can be applied that can significantly enhance the forecasting ability of MVM-BRT. Thus, another approach based on satellite-derived data could be a potential strategy to include more physical data to forecast daily ET_o.

AI-based data-driven models are turned out to be powerful tools. However, some limitations of their black-box nature also restrict their ability. The black box characterization is hard to understand and validate complex relationships if the predictor data is in a learning activity. So, integrating ML with NWP models can be an emerging area of interest for the research community. Moreover, the MVMD-BRT model can be optimized by hybridizing Bayesian Model Averaging (Sloughter et al., 2010) and bootstrapping techniques (Tiwari and Chatterjee, 2011) to enhance the forecasting capability.

Furthermore, the MVMD helps to enhance the accuracy of the BRT model as it has a great capacity to simultaneously capture the non-stationary and non-linearity in multivariate data by overcoming the mode mixing issues (Gao and Shao, 2022) as compared to MEMD (Rehman and Mandic, 2010). The MVMD helps solve the problem of adaptive selection of mode parameters using scale segmentation. Moreover, the MVMD is advantageous for dealing with the input data's multivariate oscillatory nature. The MVMD offers mode separability and avoids any predefined wavelet filter bank boundaries. The MVMD employs multivariate modulated oscillations based on a joint or common frequency component among all input data channels. The results proved that the MVMD-based hybrid models (especially MVMD-BRT) perform exceptionally well in forecasting daily ET_o compared to the MEMD based hybrid version (see Tables 5 and 6).

A newly designed MVMD-BRT model is helpful in forecasting daily reference evapotranspiration (ETo) at stations 1: Brisbane and station 2: Gympie compared to other benchmarking models. Based on the obtained goodness-of-fitness metrics, it is established that MVMD-BRT could be a viable AI model in hydrological sciences. It can provide helpful information to the Govt on water resource management to design better crop strategies, climate change scenarios, and hydrology.

6. Conclusion

This paper aims to design a multivariate variational mode decomposition integrated with a boosted regression tree model (i.e., MVMD-BRT) to forecast daily evapotranspiration (ET_o). Another vital novelty aspect is comparing the MVMD pre-processing data decomposition method with multivariate empirical model decomposition (MEMD). The results confirm that the MVMD-based hybrid models, especially MVMD-BRT, are significantly better in forecasting the daily ETo for both Brisbane and Gympie stations than the MEMD-based hybrid version of the models. Further, the MVMD-BRT model also provides better precision against the standalone version of the BRT, RF, CFNN, and ELM models where several assessment indicators such as R, MAE, RMSE, E_{WI}, E_{NS}, KGE, RRMSE and RMAE were used to measure the performance of MVMD-BRT against all comparing models.

The MVMD-BRT approach established in this study was innovative by presenting the MVMD method hybridized with the BRT model, which significantly improves the predictive performance and handles the non-stationarity and non-linearity caused by the stochasticity and chaotic nature of the meteorological drivers. Further, a comparison between MVMD *vs.* MEMD was presented for the first time in a study to establish and prove that the MVMD-based hybrid models are outstanding in forecasting. To broaden the scope, a future work can be directed to implement the proposed hybrid MVMD-BRT model in other areas of research interest, such as environment and energy, hydrology, agriculture, etc., that will surely help the Government and other stake holders to better cope with energy issues, climate change situations, and agriculture crop optimizations.

CRediT authorship contribution statement

Mumtaz Ali: Conceptualization, Formal analysis, Visualization, Software. **Mehdi Jamei:** Investigation, Visualization, Validation, Supervision. **Ramendra Prasad:** Investigation, Validation. **Masoud Karbasi:** Formal analysis, Writing – review & editing. **Yong Xiang:** Validation, Methodology. **Borui Cai:** Validation, Investigation. **Shahab Abdulla:** Investigation. **Aitazaz Ahsan Farooque:** Investigation, Supervision. **Abdulhaleem H. Labban:** Investigation.

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

The authors do not have permission to share data.

Acknowledgements

The authors thank the Bureau of Meteorology, Australia, for providing the reference evapotranspiration and other predictor data.

References

Ali, M., Deo, R.C., Downs, N.J., Maraseni, T., 2018. Multi-stage hybridized online sequential extreme learning machine integrated with Markov Chain Monte Carlo copula-Bat algorithm for rainfall forecasting. Atmos. Res. 213, 450–464.

Ali, M., Prasad, R., 2019. Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition. Renew. Sustain. Energy Rev. 104, 281–295.

Ali, M., Prasad, R., Xiang, Y., Yaseen, Z.M., 2020. Complete Ensemble Empirical Mode Decomposition Hybridized with Random Forest and Kernel Ridge Regression Model for Monthly Rainfall Forecasts. J. Hydrol. 584, 124647.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Fao, Rome 300 (9), D05109.

Andersson, L. E., M. F. Aftab, F. Scibilia and L. Imsland (2017). Forecasting using multivariate empirical mode decomposition—Applied to iceberg drift forecast. 2017 IEEE Conference on Control Technology and Applications (CCTA), IEEE.

Bellido-Jiménez, J.A., Estévez, J., García-Marín, A.P., 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 245, 106558.

Breiman, L.J.M., 1996. Bagging Predictors. 24 (2), 123–140.

Breiman, L.J.M., 2001. Random Forests. 45 (1), 5–32.

Carty, D.M., Young, T.M., Zaretzki, R.L., Guess, F.M., Petutschnigg, A., 2015. Predicting and correlating the strength properties of wood composite process parameters by use of boosted regression tree models. For. Prod. J. 65 (7–8), 365–371.

Chia, M.Y., Huang, Y.F., Koo, C.H., 2020. Support vector machine enhanced empirical reference evapotranspiration estimation with limited meteorological parameters. Comput. Electron. Agric. 175, 105577.

Colominas, M.A., Schlotthauer, G., Torres, M.E., 2014. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. Biomed. Signal Process. Control 14, 19–29.

Deo, R.C., Tiwari, M.K., Adamowski, J.F., Quilty, J.M., 2016a. Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. Stoch. Env. Res. Risk A.

Deo, R.C., Wen, X., Qi, F., 2016b. A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset. Appl. Energy 168, 568–593.

Dharma, S., Hassan, M.H., Ong, H.C., Sebayang, A.H., Silitonga, A.S., Kusumo, F., Milano, J., 2017. Experimental study and prediction of the performance and exhaust emissions of mixed Jatropha curcas-Ceiba pentandra biodiesel blends in diesel engine using artificial neural networks. J. Clean. Prod. 164, 618–633.

Douville, H., K. Raghavan, J. Renwick, R. P. Allan, P. A. Arias, M. Barlow, R. Cerezo-Mota, A. Cherchi, T. Y. Gan, J. Gergis, D. Jiang, A. Khan, W. P. Mba, D. Rosenfeld, J. Tierney and O. Zolina, Eds. (2021). Water Cycle Changes. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. United Kingdom and New York, Cambridge University Press.

Dragomiretskiy, K., Zosso, D., 2013. Variational mode decomposition. IEEE Trans. Signal Process. 62 (3), 531–544.

Dragomiretskiy, K., Zosso, D., 2014. Variational Mode Decomposition. IEEE Trans. Signal Process. 62 (3), 531–544.

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77 (4), 802–813.

Fahlman, S. and C. Lebiere (1989). "The cascade-correlation learning architecture." Advances in neural information processing systems **2**.

Faskari, S.A., Ojim, G., Falope, T., Abdullahi, Y.B., Abba, S., 2022. A Novel Machine Learning based Computing Algorithm in Modeling of Soiled Photovoltaic Module. Knowl.-Based Eng. Sci. 3 (1), 28–36.

Feng, Y., Cui, N., Zhao, L., Hu, X., Gong, D., 2016. Comparison of ELM, GANN, WNN and empirical models for estimating reference evapotranspiration in humid region of Southwest China. J. Hydrol. 536, 376–383.

Feng, K., Tian, J., 2020. Forecasting reference evapotranspiration using data mining and limited climatic data. Eur. J. Remote Sensing 54 (sup2), 363–371.

Fijani, E., Barzegar, R., Deo, R., Tziritis, E., Skordas, K., 2019. Design and implementation of a hybrid model based on two-layer decomposition method coupled with extreme learning machines to support real-time environmental monitoring of water quality parameters. Sci. Total Environ. 648, 839–853.

Fisher, J.B., Melton, F., Middleton, E., Hain, C., Anderson, M., Allen, R., McCabe, M.F., Hook, S., Baldocchi, D., Townsend, P.A., Kilic, A., Tu, K., Miralles, D.D., Perret, J., Lagouarde, J.-P., Waliser, D., Purdy, A.J., French, A., Schimel, D., Famiglietti, J.S., Stephens, G., Wood, E.F., 2017. The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources. Water Resour. Res. 53 (4), 2618–2626.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232.

Fu, T., Li, X., Jia, R., Feng, L., 2021. A novel integrated method based on a machine learning model for estimating evapotranspiration in dryland. J. Hydrol. 603, 126881.

Gao, F., Shao, X., 2022. A novel interval decomposition ensemble model for interval carbon price forecasting. Energy 243, 123006.

Granata, F., Di Nunno, F., 2021. Forecasting evapotranspiration in different climates using ensembles of recurrent neural networks. Agric Water Manag 255, 107040.

Gu, C., Qiao, X., Jin, Y., Liu, Y., 2020. A Novel Fault Diagnosis Method for Diesel Engine Based on MVMD and Band Energy. Shock Vib. 2020, 1–17.

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377 (1–2), 80–91.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Netw. 2 (5), 359–366.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. **A** 454, 903–995.

Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., 2006. Extreme learning machine: Theory and applications. Neurocomputing 70 (1–3), 489–501.

Huang, G.-B., Wang, D.H., Lan, Y., 2011. Extreme learning machines: a survey. Int. J. Mach. Learn. Cybern. 2 (2), 107–122.

IPCC (2013). Climate Change 2023 Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)].

IPCC, Ed. (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, Cambridge University Press.

Jamei, M., Ali, M., Karbasi, M., Xiang, Y., Ahmadianfar, I., Yaseen, Z.M., 2022. Designing a multi-stage expert system for daily ocean wave energy forecasting: A multivariate data decomposition-based approach. Appl. Energy 326, 119925.

Ji, N., Ma, L., Dong, H., Zhang, X., 2019. EEG signals feature extraction based on DWT and EMD combined with approximate entropy. Brain Sci. 9 (8), 201.

Jicheng, L., Gu, Y., Chou, Y., Gu, J., 2021. Seismic data random noise reduction using a method based on improved complementary ensemble EMD and adaptive interval threshold. Explor. Geophys. 52 (2), 137–149.

Karbasi, M., Jamei, M., Ahmadianfar, I., Asadi, A., 2021. Toward the accurate estimation of elliptical side orifice discharge coefficient applying two rigorous kernel-based data-intelligence paradigms. Sci. Rep. 11 (1), 19784.

Kisi, O., Sanikhani, H., Zounemat-Kermani, M., Niazi, F., 2015. Long-term monthly evapotranspiration modeling by several data-driven methods without climatic data. Comput. Electron. Agric. 115, 66–77.

Krishna, B., Rao, Y.R.S., Nayak, P.C., 2011. Time Series Modeling of River Flow Using Wavelet Neural Networks. J. Water Resour. Prot. 03 (01), 50–59.

Landeras, G., Ortiz-Barredo, A., López, J.J., 2009. Forecasting Weekly Evapotranspiration with ARIMA and Artificial Neural Network Models. J. Irrig. Drain. Eng. 135, 323–334.

Legates, D.R., McCabe, G.J., 1999. Evaluating the Use of "goodness-of-Fit" Measures in Hydrologic and Hydroclimatic Model Validation. Water Resour. Res. 35 (1), 233–241.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News 2 (3), 18–22.

Mallat, S.G., 1989. A theory for multiresolution signal decomposition: the wavelet representation. Pattern Analysis and Machine Intelligence, IEEE Transactions on 11 (7), 674–693.

Mallat, S.G., 1998. A wavelet tour of signal processing. Academic, New York.

McCuen, R.H., Knight, Z., Cutter, A.G., 2006. Evaluation of the Nash-Sutcliffe efficiency index. J. Hydrol. Eng. 11 (6), 597–602.

Meng, X.H., Evans, J.P., McCabe, M.F., 2014. The Impact of Observed Vegetation Changes on Land-Atmosphere Feedbacks During Drought. J. Hydrometeorol. 15 (2), 759–776.

Mohammadi, M.-R., Hemmati-Sarapardeh, A., Schaffie, M., Husein, M.M., Ranjbar, M., 2021. Application of cascade forward neural network and group method of data handling to modeling crude oil pyrolysis during thermal enhanced oil recovery. J. Pet. Sci. Eng. 205, 108836.
Naghibi, S.A., Pourghasemi, H.R., 2015. A comparative assessment between three machine learning models and their performance comparison by bivariate and multivariate statistical methods in groundwater potential mapping. Water Resour. Manag. 29 (14), 5217–5236.
Nourani, V., Komasi, M., Mano, A., 2009. A multivariate ANN-wavelet approach for rainfall–runoff modeling. Water Resour. Manag. 23 (14), 2877–2894.
Nourani, V., Baghanam, A.H., Adamowski, J., Kisi, O., 2014. Applications of hybrid wavelet–Artificial Intelligence models in hydrology: A review. J. Hydrol. 514, 358–377.
Prasad, R., Deo, R.C., Li, Y., Maraseni, T., 2019b. Weekly soil moisture forecasting with multivariate sequential, ensemble empirical mode decomposition and Boruta-random forest hybridizer algorithm approach. Catena 177, 149–166.
Prasad, R., M. Ali, P. Kwan and H. J. A. E. Khan (2019). "Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation." **236**: 778-792.
Rahul, S., R. Sunitha and V. Akhil (2021). Oscillation Mode Assessment in Power System Using Multivariate Variational Mode Decomposition. IECON 2021–47th Annual Conference of the IEEE Industrial Electronics Society, IEEE.
Rehamnia, I., Benlaoukli, B., Heddam, S., 2020. Modeling of seepage flow through concrete face rockfill and embankment dams using three heuristic artificial intelligence approaches: a comparative study. Environmental Processes 7 (1), 367–381.
Rehman, N.u., Aftab, H., 2019. Multivariate variational mode decomposition. IEEE Trans. Signal Process. 67 (23), 6039–6052.
Rehman, N., Aftab, H., 2019. Multivariate Variational Mode Decomposition. IEEE Trans. Signal Process. 67 (23), 6039–6052.
Rehman, N., Mandic, D.P., 2010. Multivariate empirical mode decomposition. Proc. r. Soc. a. 466 (2117), 1291–1302.
Saha, S., Arabameri, A., Saha, A., Blaschke, T., Ngo, P.T.T., Nhu, V.H., Band, S.S., 2021. Prediction of landslide susceptibility in Rudraprayag, India using novel ensemble of conditional probability and boosted regression tree-based on cross-validation method. Sci. Total Environ. 764, 142928.
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 analysis. Powder Technol. 393, 1–11.
Shamseldin, A.Y., 1997. Application of a neural network technique to rainfall runoff. J. Hydrol. 199, 272–294.
Sloughter, J.M., Gneiting, T., Raftery, A.E., 2010. Probabilistic wind speed forecasting using ensembles and Bayesian model averaging. J. Am. Stat. Assoc. 105 (489), 25–35.
Soman, K.P., Poornachandran, P., Athira, S., Harikumar, K., 2015. Recursive Variational Mode Decomposition Algorithm for Real Time Power Signal Decomposition. Procedia Technol. 21, 540–546.
Tiwari, M.K., Chatterjee, C., 2011. A new wavelet–bootstrap–ANN hybrid model for daily discharge forecasting. J. Hydroinf. 13 (3), 500–519.
Torres, M. E., M. A. Colominas, G. Schlotthauer and P. Flandrin (2011). A complete ensemble empirical mode decomposition with adaptive noise. 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).
Trajkovic, S., Todorovic, B., Stankovic, M., 2003. Forecasting of Reference Evapotranspiration by Artificial Neural Networks. J. Irrig. Drain. Eng. 129 (6), 454–457.
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.
Webb, C., 2010. Bureau of meteorology reference evapotranspiration calculations. Climate Services Centre, Queensland Regional Office, Bureau of Meteorology.
Westreich, D., Lessler, J., Funk, M.J., 2010. Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. J. Clin. Epidemiol. 63 (8), 826–833.
Willmott, C.J., 1982. Some Comments on the Evaluation of Model Performance. Bull. Am. Meteo. Soc. 63 (11), 1309–1313.
Wu, Z., Huang, N.E., 2009. Ensemble empirical mode decomposition: A noise-assisted data analysis method. Adv. Adaptive Data Analysis 1 (1), 1–41.
Xu, Z., Hou, Z., Han, Y., Guo, W., 2016. A diagram for evaluating multiple aspects of model performance in simulating vector fields. Geosci. Model Dev. 9 (12), 4365–4380.
Yang, H., Gao, L., Li, G., 2020. Underwater acoustic signal prediction based on MVMD and optimized kernel extreme learning machine. Complexity 2020, 1–17.
Zaklouta, F., Stanciulescu, B., 2012. Real-time traffic-sign recognition using tree classifiers. IEEE Trans. Intell. Transp. Syst. 13 (4), 1507–1514.
Zhang, C., H. Wang, H. Zeng, G. Chen, W. Kang and B. Zhang (2021). BP neural network for the signal recognition of micro-energy devices. Journal of Physics: Conference Series, IOP Publishing.