Deep Multi-Stage Reference Evapotranspiration Forecasting Model: Multivariate Empirical Mode Decomposition Integrated With the Boruta-Random Forest Algorithm

W. J. M. LAKMINI PRARTHANA JAYASINGHE[ID]1, (Member, IEEE),
RAVINESH C. DEO[ID]2, (Senior Member, IEEE), **AFSHIN GHAHRAMANI**[ID]3,
SUJAN GHIMIRE4, **AND NAWIN RAJ**[ID]2

[1]School of Sciences, University of Southern Queensland, Toowoomba, QLD 4350, Australia
[2]School of Sciences, University of Southern Queensland, Springfield, QLD 4300, Australia
[3]Centre for Sustainable Agricultural Systems, Institute for Life Sciences and the Environment, University of Southern Queensland, Toowoomba, QLD 4350, Australia
[4]School of Sciences, University of Southern Queensland, Toowoomba, QLD 4350, Australia

Corresponding authors: Ravinesh C. Deo (ravinesh.deo@usq.edu.au) and W. J. M. Lakmini Prarthana Jayasinghe
(lakmini.mudiyanselage@usq.edu.au)

ABSTRACT Evapotranspiration, as a combination of evaporation and transpiration of water vapour, is a primary component of global hydrological cycles. It accounts for significant loss of soil moisture from the earth to the atmosphere. Reliable methods to monitor and forecast evapotranspiration are required for decision-making. Reference evapotranspiration, denoted as ET, is a major parameter that is useful in quantifying soil moisture in a cropping system. This article aims to design a multi-stage deep learning hybrid Long Short-Term Memory (LSTM) predictive model that is coupled with Multivariate Empirical Mode Decomposition (MEMD) and Boruta-Random Forest (Boruta) algorithms to forecast ET in the drought-prone regions (*i.e.*, Gatton, Fordsdale, Cairns) of Queensland, Australia. Daily data extracted from NASA's Goddard Online Interactive Visualization and Analysis Infrastructure (GIOVANNI) and Scientific Information for Land Owners (SILO) repositories over 2003–2011 are used to build the proposed multi-stage deep learning hybrid model, *i.e.*, MEMD-Boruta-LSTM, and the model's performance is compared against competitive benchmark models such as hybrid MEMD-Boruta-DNN, MEMD-Boruta-DT, and a standalone LSTM, DNN and DT model. The test MEMD-Boruta-LSTM hybrid model attained the lowest Relative Root Mean Square Error ($\leq 17\%$), Absolute Percentage Bias ($\leq 12.5\%$) and the highest Kling-Gupta Efficiency (≥ 0.89) relative to benchmark models for all study sites. The proposed multi-stage deep hybrid MEMD-Boruta-LSTM model also outperformed all other benchmark models in terms of predictive efficacy, demonstrating its usefulness in the forecasting of the daily ET dataset. This MEMD-Boruta-LSTM hybrid model could therefore be employed in practical environments such as irrigation management systems to estimate evapotranspiration or to forecast ET.

INDEX TERMS Reference evapotranspiration forecasting, deep learning, multivariate empirical mode decomposition, boruta-random forest algorithm, long short-term memory network.

I. INTRODUCTION

Evapotranspiration estimation is involved in water resource management, hydrological studies, irrigation scheduling,

The associate editor coordinating the review of this manuscript and approving it for publication was Haiyong Zheng[ID].

crop modelling, and computing drought indices. Reference evapotranspiration (ET) and crop coefficient [1] are mostly used to estimate evapotranspiration related to a particular crop. ET can be directly measured by using the lysimeter method. Several empirical methods such as the Hargreaves equation, Priestley–Taylor equation, Ritchie equation, and

the PMF-56 equation have also been developed to estimate *ET* using climatic data. Among them, the PMF-56 equation is widely used due to its accuracy and stability [2]. Other than empirical methods, many researchers have developed data-driven Artificial Intelligence (AI) models to forecast *ET* and these models have shown superior performances despite non-linear behaviour of *ET* [3]. For instance, Nourani, *et al.* [4] employed ensemble Multiple Linear Regression (MLR), Support Vector Regression (SVR), Adaptive Neuro-Fuzzy Inference System, Artificial Neural Network (ANN), and Multiple Linear Regression (MLR) models for *ET* forecasting and the ensemble MLR model has shown the best performance. Tikhamarine, *et al.* [5] examined the comparative potential of ANN-Embedded Grey Wolf Optimizer, Multi-Verse Optimizer, Particle Swarm Optimizer, Whale Optimization Algorithm and Ant Lion Optimizer to predict monthly *ET* in India and Algeria.

Deep Learning (DL) techniques such as the Temporal Convolution Network [6] and the ensemble of Convolutional Neural Networks (CNN) [7] which are comparatively more advanced and precise than the above traditional machine learning methods have also been recently employed to predict *ET*. The Long Short-Term Memory (LSTM) network is a DL neural technique that has been used to predict hydrological variables like water quality [8], solar radiation [9], and streamflow [10], and rainfall-runoff [11]. Several recent studies have shown the exceptional performance of LSTM model in predicting hydrological time series [67]–[70]. The key advantage of the LSTM model is its ability in using sequential data as inputs instead of independent training samples and this feature ensures the model's capability in dealing with more extended historic hydrologic observations with temporal dependence [71], which is a common characteristic related with many types of hydrological time series [66]. However, less research has been carried out to predict *ET* using LSTM based models. Yin, *et al.* [12] proposed a new hybrid bi-directional LSTM model to forecast short term daily *ET* in data scarce regions.

In recent years, use of AI models have become more popular in resolving problems related to many various hydrological aspects [72]. For instance, DT model has been used to map the flood susceptible areas in Kelantan, Malaysia which performed with greater accuracy in comparison with frequency ratio (FR) and logistic regression (LR) methods [74]. The DNN model has been employed in water resource management e.g. development of spatial-temporally continuous evapotranspiration model [75], development of model for mapping suitable groundwater extraction location [76] and shown better performances compared to benchmark models. LSTM is also extensively used for flood forecasting [77], and predicting water table [8], etc.

To further enhance the forecasting model capabilities, hybrid models have been developed in the recent past by many researchers. Ferreira and da Cunha [13] developed a DL multi-step *ET* forecasting model with hybrid CNN-LSTM and assessed in comparison with standalone LSTM, CNN and traditional machine learning models (ANN and RF). According to the performance analysis, the hybrid CNN-LSTM model outperformed all the comparison models.

In addition, two-phase hybrid models which are capable of yielding high performances with relatively low errors are explored [14]. For example, Prasad, *et al.* [15] developed a two-phase hybrid Extreme Learning Machine (ELM) model to forecast soil moisture coupled with the Ensemble Empirical Mode Decomposition (EEMD) data pre-processing method and the Boruta-random forest optimizer (Boruta) feature selection method. This model was superior to the other comparative models and yielded a relatively accurate performance with a small number of errors.

The Multivariate Empirical Mode Decomposition (MEMD) is a data pre-processing method that is an improved extension of standard Empirical Mode Decomposition (EMD) for multichannel data [16] and works efficiently in time series nonlinear and nonstationary signal data pre-processing [17]. For instance, Prasad, *et al.* [15] and Ali, *et al.* [18] proposed new multi-stage models coupled with MEMD to forecast solar radiation and drought thereby showing superior performance when compared with other models.

Boruta-random forest (Boruta) is a feature selection technique [19] that can identify significant input parameters using a comparison with real features to those of random probes [20]. Boruta has been utilized successfully as a feature selection technique in hybrid models to forecast soil moisture [20], [21], streamflow [22], [23], and air quality [24].

However, *ET* forecasting based on multi-stage deep neural networks is yet to be explored. To address this research gap, this study is focused on developing a novel multi-stage MEMD-Boruta-LSTM deep neural network to forecast daily *ET* based on satellite and ground data. DT and DNN models which have been widely employed in prediction of various hydrological parameters are selected for model performance comparison with target model in this study.

II. THEORETICAL OVERVIEWS
In this section, the MEMD, Boruta, and LSTM are described in detail. The models used for comparison purposes in this study: Deep Neural Network (DNN) [9] and Decision Tree (DT) [25] are not explained in detail as they are well-known algorithms.

A. MULTIVARIATE EMPIRICAL MODE DECOMPOSITION METHOD
The MEMD is an advanced version of EMD proposed by Rehman and Mandic [16] which is capable of dealing with multivariate signals and resolved the mode mixing issue by using white Gaussian noises [26]. The MEMD method can be described as follows [16]:

I. Generate a suitable number of direction vectors.
II. Calculate projections of the multiple inputs along with different directions in an *n*-dimensional space.

III. Identify local maxima projections and obtain multi-variate envelope curves through them and subsequently calculate the mean.

IV. Extract the detail using the difference of the mean envelope curve and original signal until the stopping criteria is satisfied for a multivariate Intrinsic Mode Function (IMF) [27].

The mathematical formulae of the MEMD can be found elsewhere [16], [28].

B. FEATURE SELECTION: BORUTA-RANDOM FOREST OPTIMIZER ALGORITHM

The algorithm can be briefly explained as follows [19], [29]:

Let $x_t \in R^n$ be the group of predictors for the set of T and $y_t \in R$ be the target for n number of inputs, where $t = 1, 2, \ldots, T$.

I. Create a randomly ordered duplicated (shadow) variable, x'_t for x_t and then predict the target y_t.

II. Calculate Mean Decrease Accuracy (MDA) for every x_t and x'_t over all trees, m_{tree} (=500 in this study) [1], [30]:

$$
\begin{aligned}
&MDA \\
&= \frac{1}{m_{tree}} \sum\nolimits_{m=1}^{m_{tree}} \\
&\quad \times \frac{\sum_{t \epsilon OOB} I\left(y_t = f\left(x_t\right)\right) - \sum_{t \epsilon OOB} I\left(y_t = f\left(x_t^n\right)\right)}{|OOB|},
\end{aligned}
$$

(1)

where $I\left(*\right)$ is indicated function, OOB (Out-of-Bag) is a predictive error, $y_t = f\left(x_t\right)$ is predicted value before permuting and, $y_t = f\left(x_t^n\right)$ is predicted value after permuting.

III. Compute the Z-score as:

$$
Z - score = \frac{MDA}{SD}
$$

(2)

where, SD is the standard deviation of accuracy loss and, then maximum Z-score (Z_{max}) is determined among duplicated attributes.

IV. Following that, predictors are identified as "Unimportant" when $Z - score < Z_{max}$ and "Confirmed" as important when $Z - score > Z_{max}$ during the process.

C. TIME SEQUENTIAL PREDICTIVE METHOD: LONG SHORT-TERM MEMORY NETWORK

The LSTM is a special Recurrent Neural Network (RNN) [32] related to conventional artificial neural networks that are mainly used to identify patterns in sequences of data [33]. The LSTMs operates with special units, denoted as memory blocks that consist of input, output, and forget gates and these memory blocks continuously update and control the information flow [34]. The calculations are described in 4 steps as follows [35]:

I. The LSTM layer decides which information should be forgotten or remembered, based on the last hidden layer output h_{t-1} and the new input x_t by using "forget gate" f_t:

$$
f_t = \sigma \left(w_f \left[h_{t-1}, x_t\right] + b_f\right)
$$

(3)

where w_f is the weight matrix; b_f is the bias vector and $\sigma \left(\ldots\right)$ is the logistic sigmoid function.

II. The LSTM layer decides what information needs to be stored in the new cell state c_t that is represented by the new candidate cell state $\overline{C}_t$ after updating information by using "input gate" i_t:

$$
\overline{C}_t = tanh \left(w_C \left[h_{t-1}, x_t\right] + b_C\right)
$$

(4)

$$
i_t = \sigma \left(w_i \left[h_{t-1}, x_t\right] + b_i\right),
$$

(5)

where $tanh \left(\ldots\right)$ is the hyperbolic tangent function.

III. The old cell state C_{t-1} updates to C_t by the "forget gate" f_t to remove unnecessary information and the "input gate" i_t to get a new candidate cell state $\overline{C}_t$:

$$
C_t = f_t * C_{t-1} + i_t * \overline{C}_t
$$

(6)

IV. Finally, the output h_t is derived using "output gate" o_t and the cell state C_t:

$$
o_t = \sigma \left(w_o \left[h_{t-1}, x_t\right] + b_o\right)
$$

(7)

$$
h_t = o_t * tanh \left(C_t\right)
$$

(8)

III. MATERIALS AND METHOD

A. STUDY REGION AND DATASET

This study is centred in Queensland (QLD) Australia, where 84% of the total land resources are used for agricultural operations [36]. The Queensland government declared 67.4% of the land area of Queensland drought-affected in the year 2020 [37]. Therefore, developing a precise model to forecast water losses due to ET is useful for strategic planning in water resources management in the state.

The three examined sites located in arid and semi-arid areas in QLD, Australia selected for this study are Gatton $-152.34^\circ E, 27.54^\circ S$, Fordsdale $-152.12^\circ E, 27.72^\circ S$ and Cairns $-145.75^\circ E, 16.87^\circ S$ (see Figure 1). The land resources of these selected sites are mainly used for agricultural purposes.

To construct a target hybrid model, data for eight daily predictive climatic variables for the period 01 February 2003 to 19 April 2011 were extracted from the databases of NASA's Goddard Online Interactive Visualization and Analysis Infrastructure (GIOVANNI) - Atmospheric Infrared Sounder (AIRS) and GLDAS model satellite and Scientific Information for Land Owners (SILO). The GIOVANNI provides easy and user-friendly access to visualize and analyse the vast amount of Earth Science-related remote sensing data [38] that can be extracted easily without the requirement for advanced prior knowledge of complex remote sensing datasets. In addition, SILO data source provides ground-based data for predictor variables and it assists to further improve the model's performance. This database is operationally managed by the Queensland Government [39].
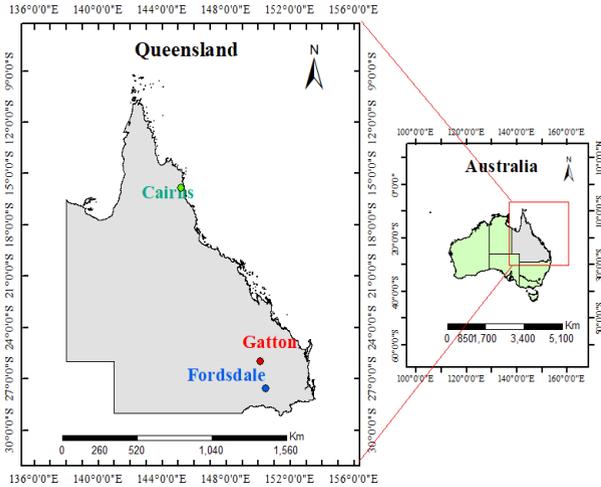
FIGURE 1. Study sites in Queensland, Australia where the proposed MEMD-Boruta-LSTM model was implemented.

TABLE 1. List of satellite-based Goddard Online Interactive Visualization and Analysis Infrastructure (GIOVANNI) and the ground-based Scientific Information for Land Owners (SILO) predictor variables used to forecast daily Reference Evapotranspiration (*ET*). Note: Atmospheric Infrared Sounder (AIRS) and GLDAS model are the two platforms in the GIOVANNI data source.

Data source		Name of input variable	Acronym	Unit
GIOVANNI-Satellite data	AIRS	Surface Temperature-Day	Tsd	°C
		Surface Temperature-Night	Tsn	°C
		Air Pressure	Pa	hPa
	GLDAS Model	Bare Soil Evaporation	Ebs	$kgm^{-2}s^{-1}$
		Transpiration	TR	$kgm^{-2}s^{-1}$
SILO-Ground based data		Maximum Temperature	Tmax	°C
		Minimum Temperature	Tmin	°C
		Radiation	Ra	MJm^{-2}

Missing data due to instrumental and equipment failures were filled with daily mean data of previous years [40]. Table 1 shows a summary of predictive variables and sources of data. For the target variable that is daily *ET*, point-based data is extracted from the SILO database.

B. DEVELOPMENT OF THE PROPOSED MULTI-STAGE DEEP HYBRID MEMD-BORUTA-LSTM MODEL

The proposed multi-stage MEMD-Boruta-LSTM model was developed using an Intel Core i7 @ 3.3 GHz and 16 GB memory computer; built using freely available DL libraries: Keras [46] and TensorFlow [47] in Python [48]. The MEMD data pre-processing method and Boruta feature selection method were implemented using MATLAB R2019b and R respectively, while "*matplotlib*" and "*seaborn*" tools in

Python were used for visualizations. The MEMD-Boruta-LSTM hybrid model was developed using historical time series inputs as follows:

Stage 1: In this study, before performing MEMD, firstly, all nine variables (eight predictors + target) (see Table 1) were partitioned into 50% for training (*i.e.*, 1500 data points) and other 50% for testing (*i.e.*, 1500 data points) for all study sites [49] to avoid having a different number of Intrinsic Mode Functions (IMFs). Deo, *et al.* [50] pointed out that, if the complete dataset (training, cross-validation, and testing) is decomposed together without partitioning as explained above, future data (that is testing and yet unseen data by the forecasting model at a particular time step) would unintentionally add bias into the forecast. Thus, it is an important requirement during the decomposition stage to avoid incorporating future datasets that are to be used in the testing phase with the calibration dataset i.e., training and cross-validation in this study.

The MEMD was performed in the decomposition process independently for each training, and testing data partitions for both predictor and target variables for all three sites. In this process, the recommended predefined parameters: ensemble number ($N = 500$) and amplitude of the added white noise ($\varepsilon = 0.2$) were applied [51]–[54]. All the first IMFs of predictor and target variables were pooled into one set. All the second IMFs of predictor and target variables were pooled into one set. This pooling was carried out until the i$^{\text{th}}$ IMFs including residuals.

Stage 2: Boruta-random forest is a feature selection technique available in R. Random Forest tree-based algorithm is embedded in this feature selection technique [21]. This feature selection algorithm is used to identify the significantly corelated predictor variables to the target variable in each IMFs and residuals using historical lagged data at (t-1).

Stage 3: After identifying the significantly corelated predictor variables for the model development, respective data of those variables were normalized to remove the variance of features [55] by converting them into $(0 - 1)$ range using equation (9):

$$X_n = \frac{X_{actual} - X_{min}}{X_{max} - X_{min}} \tag{9}$$

where X_{actual}, X_{max}, and X_{min} represent input data for actual, maximum, and minimum values respectively.

Stage 4: The LSTM model was employed to forecast daily *ET* in each IMF and residual using significantly corelated predictor variable data at (t-1) lag. To prepare the best model design, hyperparameters for the target model (MEMD-Boruta-LSTM) were identified using the *Hyperopt* library in Python [56], [57]. *Hyperopt* is one of the hyperparameter optimization algorithms that performed better than the *Grid search* and *Random search* algorithms as it ensures comparatively less time in the model training process while increasing the accuracy of the model [58]. Thereby optimal architecture of the hybrid MEMD-Boruta-LSTM model was used to predict daily *ET*. Finally forecasted *ET* in each
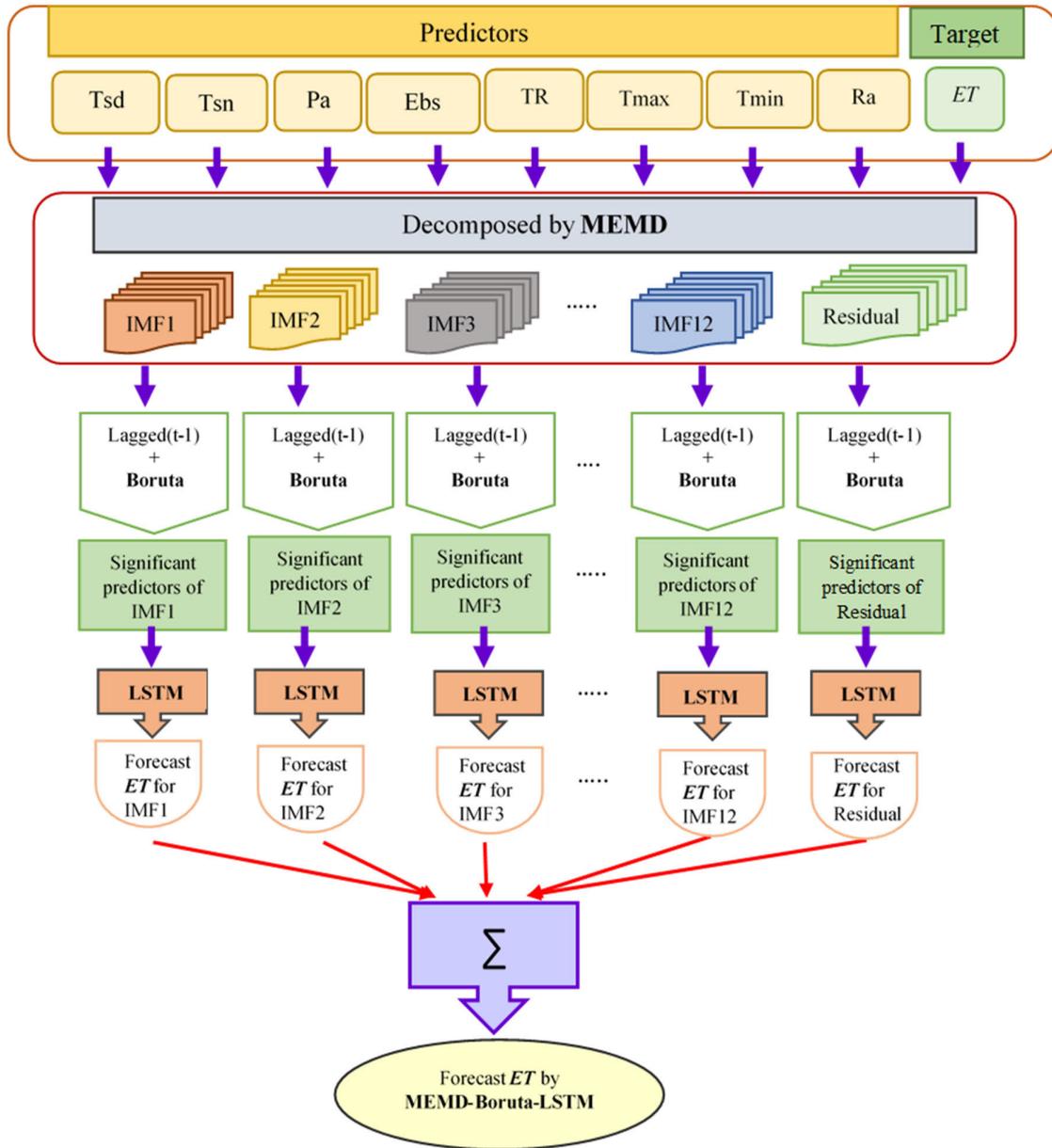
FIGURE 2. Workflow diagram detailing the necessary steps taken to design proposed deep hybrid MEMD-Boruta-LSTM model for daily evapotranspiration (*ET*) forecast. Note: *ET* = Evapotranspiration, MEMD = Multivariate Empirical Mode Decomposition, IMF = Intrinsic Mode Function, LSTM = Long Short-Term Memory. The details of predictors are given in TABLE 1.

IMF and a residual were cumulated to calculate forecasted daily *ET* for each study site. Figure 2 presents the workflow of the proposed multi-stage MEMD-Boruta-LSTM model. The same procedure is followed to develop hybridized DNN and DT with MEMD-Boruta (*i.e.*, MEMD-Boruta-DNN and MEMD-Boruta-DT models). Developed standalone LSTM, DNN, and DT models and hybrid MEMD-Boruta-DNN and MEMD-Boruta-DT were used as benchmark models for the model performance comparison.

C. MODEL PERFORMANCE EVALUATION
Model performances are evaluated using the statistical metrics [41]–[45] given below to confirm whether the target

predictive model is superior to other benchmark models and is sufficiently qualified for *ET* prediction in QLD,

I. Correlation Coefficient (*r*): The correlation coefficient measures the strength of the relationship between two variables and the values range between -1.0 and 1.0 [62]. The value given for perfect forecasting models is equal to $+1$ indicating strong positive relationship of forecasted values derived from the model with actual values, (10) as shown at the bottom of the next page.

II. Root Mean Square Error $\left(RMSE; mmday^{-1}\right)$: This measures the average model-performance error between predicted value (E_p^{FOR}) and observed value

(E_p^{OBS}) [63]. The *RMSE* value can range from 0 to ∞ and it becomes zero for the best predictive models.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(ET^{FOR,i} - ET^{OBS,i} \right)^2},$$
$$0 \leq RMSE < \infty \quad (11)$$

III. Mean Absolute Error $\left(MAE; \, mmday^{-1} \right)$: This error value provides an assessment of the actual forecasting errors in terms of the total number of observations [21]. *MAE* can range from 0 to ∞ and it becomes zero for best predictive models. The *MAE* gives a more precise measure of average model error than the *RMSE* since it is not influenced by extreme outliers [41].

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| ET^{FOR,i} - ET^{OBS,i} \right|,$$
$$0 \leq MAE < \infty \quad (12)$$

IV. Relative Root Mean Squared Error (*RRMSE*): The *RRMSE* is used to measure overall forecasting accuracy of the models and always gives positive values [21]. If the value for *RRMSE* is less than 10%, model performance is considered to be outstanding, while model performance is considered to be good if it is lying between 10% to 20%. If the value for *RRMSE* error lies between 20% to 30%, model performance is considered as fair. If the value for *RRMSE* error is higher than 30% model performance is considered to be poor [64].

$$RRMSE = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(ET^{FOR,i} - ET^{OBS,i} \right)^2}}{\frac{1}{N} \sum_{i=1}^{N} ET^{OBS,i}} \times 100$$
$$(13)$$

V. Relative Mean Absolute Percentage Error (*RMAE*): The relative mean absolute percentage error measures the size of the error in percentage terms.

$$RMAE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{ET^{FOR,i} - ET^{OBS,i}}{ET^{OBS,i}} \right| \times 100 \quad (14)$$

VI. Nash-Sutcliffe Index (*NS*): The *NS* [43] measures how well the plotted line between observed data and simulated data fits into 1:1. The *NS* is equal to 1, if the model forecasted data is perfectly matched to the observed data. NSE = 0 indicates that the model predictions are as accurate as the mean of the observed data while, Inf

TABLE 2. Summarized results of MEMD process.

Study site	Parameters used in MEMD		Number of initial predictor variables	Number of IMFs & residual	Total number of predictor variables after MEMD
	Ensemble number (N)	Amplitude of the added white noise (ε)			
Gatton	500	0.2	8	13	104 (13×8)
Fordsdale	500	0.2	8	13	104 (13×8)
Cairns	500	0.2	8	12	96 (12×8)

< NSE < 0 indicates that observed mean is a better predictor than the model [62].

$$NS = 1 - \left[\frac{\sum_{i=1}^{N} \left(ET^{OBS,i} - ET^{FOR,i} \right)^2}{\sum_{i=1}^{N} \left(ET^{OBS,i} - \overline{ET^{OBS}} \right)^2} \right],$$
$$-\infty < E_{NS} \leq 1 \quad (15)$$

VII. Willmott's Index (*WI*): Willmott index is a standardized measure of the degree of model prediction error and the value for *WI* ranges from 0 to 1, whereas this value equals 1 for best predictive models, (16) as shown at the bottom of the page.

VIII. Legate and McCabe Index *(LM)*: The *LM* is an advanced assessment index based on *WI* and *NS* values. This index can be used to assess the goodness-of-fit of a hydrologic or hydro climatic model and is more effective than correlation and correlation-based measures (e.g., the Coefficient of Determination (r^2), WI and NS) [41]. The value for *LM* ranges from $-\infty$ to 1, whereas this value equals one for best predictive models.

$$LM = 1 - \left[\frac{\sum_{i=1}^{N} \left| ET^{FOR,i} - ET^{OBS,i} \right|}{\sum_{i=1}^{N} \left| \left(ET^{OBS,i}, -\overline{ET^{OBS,i}} \right) \right|} \right],$$
$$-\infty < LM \leq 1 \quad (17)$$

IX. Absolute Percentage Bias (*APB%*): The *APB* gives the error of forecasted values as a percentage concerning the observed values. The optimal value for APB is zero and lower-magnitude values closer to zero reflect good accuracy of the model [65].

$$APB = \left[\frac{\sum_{i=1}^{N} \left(ET^{OBS,i} - ET^{FOR,i} \right) \times 100}{\sum_{i=1}^{N} ET^{OBS,i}} \right] \quad (18)$$

$$r = \frac{\sum_{i=1}^{N} \left(ET^{OBS,i} - \overline{ET^{OBS,i}} \right) \left(ET^{FOR,i} - \overline{ET^{FOR}} \right)}{\sqrt{\sum_{i=1}^{N} \left(ET^{OBS,i} - \overline{ET^{OBS,i}} \right)^2} \sqrt{\sum_{i=1}^{N} \left(ET^{FOR,i} - \overline{ET^{FOR,i}} \right)^2}}, \quad -1 \leq r \leq 1 \quad (10)$$

$$WI = 1 - \left[\frac{\sum_{i=1}^{N} \left(ET^{OBS,i} - ET^{FOR,i} \right)^2}{\sum_{i=1}^{N} \left(\left| \left(ET^{FOR,i} - \overline{ET^{OBS,i}} \right) \right| + \left| \left(ET^{OBS,i} - \overline{ET^{OBS,i}} \right) \right| \right)^2} \right], \quad 0 \leq WI \leq 1 \quad (16)$$

TABLE 3. Summarized results of Boruta feature selection process.

Study site	Number of initial predictor variables in each IMF and residual	Step 1: MEMD — Total number of predictor variables after MEMD	Step2: Boruta Feature Selection — Number of selected predictor variables in each IMF and residuals													Total number of predictor variables identified after feature selection
			IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10	IMF11	IMF12	Residual	
Gatton	8	104 (13×8)	6	6	7	8	8	8	8	8	8	8	8	8	8	99
Fordsdale	8	104 (13×8)	6	6	6	8	8	8	8	8	8	8	8	8	8	98
Cairns	8	96 (12×8)	6	6	6	6	8	8	8	8	8	8	8	-	8	88

TABLE 4. List of selected predictor variables identified in each IMF and residual used to develop hybrid MEMD-Boruta-LSTM, MEMD-Boruta-DNN and MEMD-Boruta-DT models in Cairns site.

		LSTM	DNN	DT
Standalone model		Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	Pa$_{t-1}$, Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	Pa$_{t-1}$, Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
Hybrid model	IMF1	**6 significant inputs** Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**6 significant inputs** Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**6 significant inputs** Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
	IMF2	**6 significant inputs** Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**6 significant inputs** Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**6 significant inputs** Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
	IMF3	**6 significant inputs** Tsd$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**6 significant inputs** Tsd$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**6 significant inputs** Tsd$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
	IMF4	**6 significant inputs** Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**6 significant inputs** Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**6 significant inputs** Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
	IMF5	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
MEMD-Boruta	IMF6	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
	IMF7	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
	IMF8	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
	IMF9	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
	IMF10	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
	IMF11	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$
	Residual	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$	**8 significant inputs** Pat-1,Tsd$_{t-1}$, Tsn$_{t-1}$, Ebs$_{t-1}$, TR$_{t-1}$, Tmax$_{t-1}$, Tmin$_{t-1}$, Ra$_{t-1}$

X. Kling-Gupta Efficiency (*KGE*): The *KGE* measures the goodness-of-fit of the model. This metric can be decomposed into the contribution of mean, variance, and correlation on the model performance [45]. Perfect models will give value one for the *KGE* index [65].

KGE

$$= 1 - \sqrt{(r-1)^2 + \left(\frac{CV_{FOR}}{CV_{OBS}}\right)^2 + \left(\frac{\overline{ET^{FOR,i}}}{\overline{ET^{OBS,i}}} - 1\right)^2} \tag{19}$$

where CV = Coefficient of Variation, where $ET^{OBS,i}$ and $ET^{FOR,i}$ are observed and forecasted ith value of

the evapotranspiration ET, $\overline{ET^{OBS,i}}$ and $\overline{ET^{FOR,i}}$ are the observed and forecasted average of ET and N is the total number of data points of the test dataset.

IV. RESULTS AND DISCUSSIONS

In the decomposition process, training and testing datasets for Gatton and Fordsdale sites were decomposed into 12 IMFs and a residual component (i.e. 104 (=**8 × 13**) predictors) whereas 11 IMFs and a residual component (i.e. 96 (=**8 × 12**) predictors) were generated by decomposing training and testing datasets for the Cairns site (see Table 2). In the Boruta feature selection process, 99 predictor variables were identified as significantly corelated to the target variable ET in all IMFs and the residual of Gatton, while 98 and 88 predictor variables were identified for Fordsdale, and Cairns sites respectively (see Table 3). Table 4 shows the selected final predictor variables in each IMF and the residual used to develop target hybrid MEMD-Boruta-LSTM model and benchmark models in Cairns site. Identified hyperparameters for the LSTM in target model through the hyperparameter optimization process are listed in Table 5.

The performance of the multi-stage deep MEMD-Boruta-LSTM model and other comparative models: MEMD-Boruta-DNN, MEMD-Boruta-DT, LSTM, DNN and, DT in the testing phase were assessed using statistical metrics calculated using equations (10) to (19), visual graphs, and error distributions between forecasted and observed ET.

Table 6 shows the results derived for statistical metrics: Correlation Coefficient (r), Root Mean Squared Error ($RMSE$; $mm\ day^{-1}$), Mean Absolute Error (MAE; $mm\ day^{-1}$), Willmott's Index (WI), Nash-Sutcliffe coefficient (NS), and Legates and McCabe's (LM). According to the results shown in table 6, the proposed multi-stage deep MEMD-Boruta-LSTM model has yielded the highest r, WI, NS, and LM and lowest $RMSE$ and MAE values over the other benchmark models at all study sites. For instance, values scored for r, WI, NS, and LM by this proposed model for the Gatton site where it showed the best performances among all study sites are 0.9668, 0.9723, 0.8960, and 0.6996 respectively and higher than the respective values scored by other benchmark models. Furthermore, for the same site, this proposed model scored 0.5307 and 0.4204 for $RMSE$ and MAE respectively, and these are the lowest recorded values. These results indicate that the proposed multi-stage deep hybrid MEMD-Boruta-LSTM model can be confidently employed for forecasting daily ET and for achieving higher forecasting accuracy compared to counterpart models (MEMD-Boruta-DNN and, MEMD-Boruta-DT) and standalone models (LSTM, DNN, and DT).

In terms of the Absolute Percentage Bias (APB%) error and Kling-Gupta Efficiency (KGE) calculated in the testing phase, Figure 3(a) and 3(b) show that the proposed deep multi-stage MEMD-Boruta-LSTM model generates better performance in terms of APB% error percentage and KGE respectively. Figure 3(a) illustrations that the proposed MEMD-Boruta-LSTM model has scored the lowest APB

TABLE 5. List of hyperparameters for the LSTM model. The optimal parameters used for all sites are boldfaced (in blue). Note: ReLU, Uniform, He_uniform, Glorot_uniform, and adam stand for the rectified linear units, uniform initializer, He uniform variance scaling initializer, Glorot uniform initializer, and adaptive moment estimation respectively.

Model hyperparameter Name	Search space for optimal hyperparameter
LSTM Layer 1	[50,**100**,150]
LSTM Layer 2	[50,**100**,150]
LSTM Layer 3	[50,100,**150**]
LSTM Layer 4	[10,20,30,40,**50**]
LSTM Layer 5	[10,20,**30**,40,50]
LSTM Layer 6	[10,20,30,40,**50**]
Epochs	[30, 500, 100,**2000**]
Activation Function	[relu, **tanh**, sigmoid]
Weight Initializer	[uniform , he_uniform, **glorot_uniform**]
Recurrent Activation Function	[**relu**, tanh, sigmoid]
Optimizer	[**adam**]
Dropout Ratio	[0.1, **0.2**, 0.3]
Batch Size	[10,20,**30**]

Architecture of the backpropagation algorithm	
Alpha, α	0.001
Epsilon, ε	0.0000001
Beta, β_1, β_2	0.9, 0.999

α = Learning rate

ε = Small number to prevent any division by zero

β_1, β_2= 1st, 2nd moment estimation exponential decay rate

error percentage (9.2-12.3%) while other all comparative models' APB error percentages are within (11.6-19.7%) range for all sites. According to Figure 3(b), the proposed MEMD-Boruta-LSTM model has yielded the highest KGE values (0.89-0.91) while, KGE values are less than 0.86 for other all benchmark models for all sites. These results also provide strong evidence to recognize the superior potentiality of the proposed multi-stage MEMD-Boruta-LSTM model in daily ET forecasting over the other benchmark models.

The radar plots in Figure 4 demonstrate the proposed deep multi-stage MEMD-Boruta-LSTM model yielded the lowest values for $RRMSE$,% and $RMAE$,% for all sites (12.59% and 10.89% at Gatton, 16.21% and 15.06% at Fordsdale, 12.47% and 11.29% at Cairns respectively). Further, all values scored for $RRMSE$ and $RMAE$ for all sites by this proposed deep

TABLE 6. Performance evaluation of the proposed hybrid MEMD-Boruta-LSTM model in the testing phase for the comparative counterpart models in terms of the Pearson's Correlation Coefficient (r), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Willmott's Index (WI), Nash Sutcliffe coefficient (NS) and Legates and McCabe's (LM). The best model is boldfaced (in blue).

Study Site	Predictive Model	Model Performance Metrics					
		r	RMSE (mm day⁻¹)	MAE (mm day⁻¹)	WI	NS	LM
Gatton	**MEMD-Boruta-LSTM**	**0.9668**	**0.5307**	**0.4204**	**0.9723**	**0.8960**	**0.6996**
	MEMD-Boruta-DNN	0.8978	0.7637	0.5855	0.9277	0.7841	0.5807
	MEMD-Boruta-DT	0.9195	0.6495	0.4904	0.9547	0.8439	0.6488
	LSTM	0.8614	0.8627	0.6671	0.9200	0.7254	0.5231
	DNN	0.8527	0.9638	0.7471	0.8827	0.6562	0.465
	DT	0.8282	0.9431	0.7128	0.8972	0.6708	0.4895
Fordsdale	**MEMD-Boruta-LSTM**	**0.9343**	**0.5773**	**0.4380**	**0.9609**	**0.8424**	**0.6404**
	MEMD-Boruta-DNN	0.8427	0.8791	0.6824	0.8906	0.6348	0.4403
	MEMD-Boruta-DT	0.8936	0.7171	0.5549	0.9329	0.7575	0.5449
	LSTM	0.8364	0.8853	0.6366	0.8959	0.6295	0.4773
	DNN	0.8134	0.8836	0.7006	0.8628	0.6311	0.4252
	DT	0.7990	0.8827	0.6692	0.8818	0.6318	0.4511
Cairns	**MEMD-Boruta-LSTM**	**0.9044**	**0.5218**	**0.4130**	**0.9307**	**0.7655**	**0.5369**
	MEMD-Boruta-DNN	0.8108	0.6382	0.4881	0.8731	0.6482	0.4519
	MEMD-Boruta-DT	0.8240	0.6131	0.4663	0.9014	0.6754	0.4763
	LSTM	0.7261	0.7521	0.5610	0.8157	0.5128	0.3709
	DNN	0.7043	0.7700	0.5980	0.7855	0.4879	0.3285
	DT	0.6890	0.7827	0.5939	0.8033	0.4709	0.3331
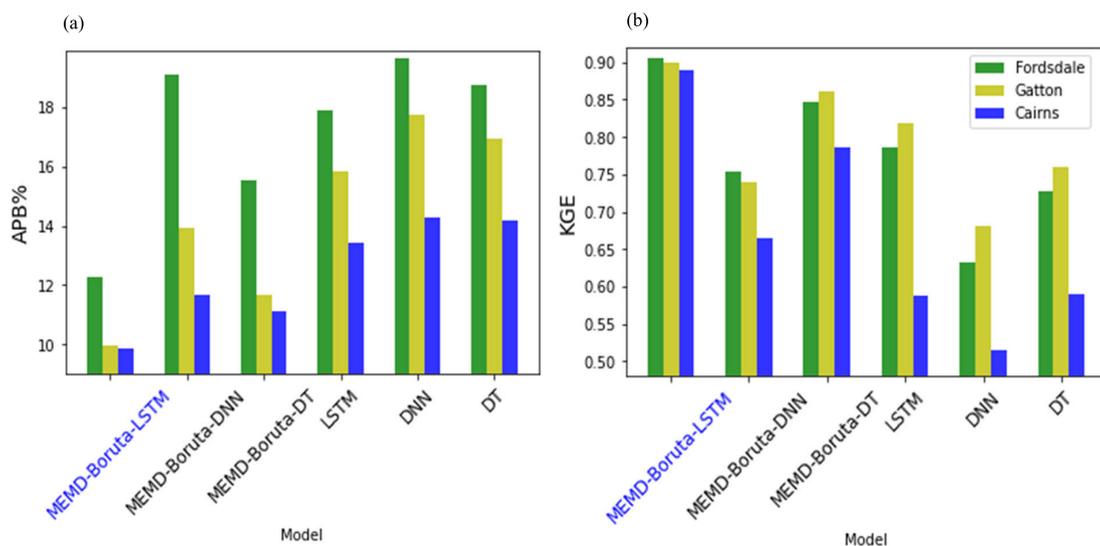
FIGURE 3. Bar graphs show the comprehensive assessment of the performance of the proposed MEMD-Boruta-LSTM model against the counterpart models, based on the (a) Absolute Percentage Bias (*APB*, %) error and (b) Kling-Gupta Efficiency (*KGE*) in the testing phase for the all study sites. The best model for all sites is boldfaced (in blue).
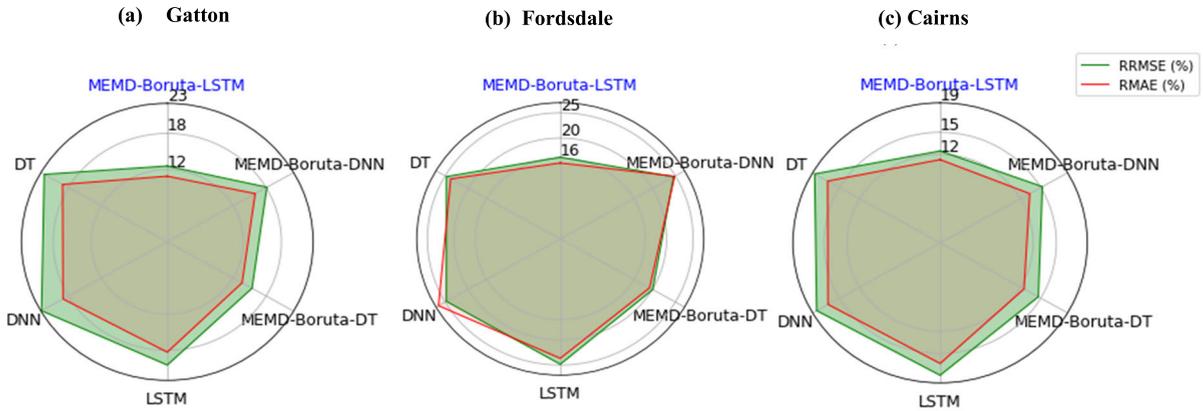
FIGURE 4. The radar plots showing the Relative Root Mean Squared Error (*RRMSE* %) and Relative Mean Absolute Error (*RMAE* %) of the MEMD-Boruta-LSTM hybrid model and comparative models constructed for 1-day evapotranspiration forecasting in the testing phase. The best model is boldfaced (in blue).

MEMD-Boruta-LSTM model is lying within the range of 10%-20%. Therefore, this proposed model can be categorized under the good model group having lower model errors of less than 20% [59], [60].

To further validate the proposed MEMD-Boruta-LSTM model the absolute Forecasting Errors (|FE| = |Observed *ET* − Forecasted *ET*|; mm) of this proposed model and all other benchmark models are compared. The box plots in Figure 5 depict the distribution of |FE| in the testing phase with their upper, median and, lower quartiles for all models and weather stations. The results of box plots shown in Figure 5 indicate that the proposed multi-stage MEMD-Boruta-LSTM model presented the smallest quartiles for |FE| for all sites followed by MEMD-Boruta-DNN, MEMD-Boruta-DT, LSTM, DNN, and DT. These results also clearly indicate that the proposed deep multi-stage MEMD-Boruta-LSTM model is superior to the other benchmark models.

The Empirical Cumulative Distribution Function (ECDF, Figure 6) is also used to illustrate the forecasting skills in terms of the absolute Forecasting Error, |FE| (mm) at each site. Forecasting errors of good models should be closer to zero. The all-hybrid MEMD-Boruta-LSTM, MEMD-Boruta-DNN, and MEMD-Boruta-DT models performed better than standalone LSTM, DNN, and DT models. Based on the forecasting error (0 to ±4 *mm*), Figure 6 visibly depicts that the proposed MEMD- Boruta-LSTM model is the most accurate compared to all other benchmark models.

In summary, the proposed multi-stage hybrid DL model (*i.e.*, MEMD-Boruta-LSTM) provided significant high performance with the lowest values of the absolute and relative errors i.e., *APB, RMSE, MAE, RRMSE*, and *RMAE*, including the highest *r; WI, NS, LM* and *KGE* in respect to the other benchmark models. Consequently, it is promising that the results confirm the deep multi-stage MEMD-Boruta-LSTM model has the potential to forecast daily *ET* and its perfor-
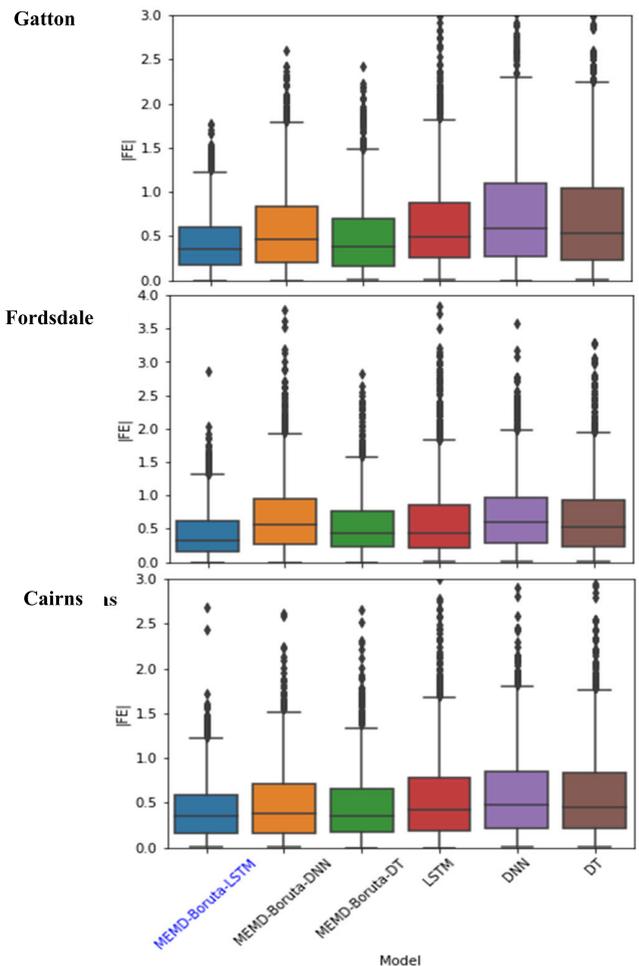
FIGURE 5. The box plots of the absolute value of the Forecasting Errors (|FE|) in the testing phase, generated by the hybrid MEMD-Boruta-LSTM model compared to that of the other predictive models implemented at all study sites. The best model is boldfaced (in blue).

mance exceeds that of all other comparative hybrid DL and standalone models for all the study sites in Queensland.
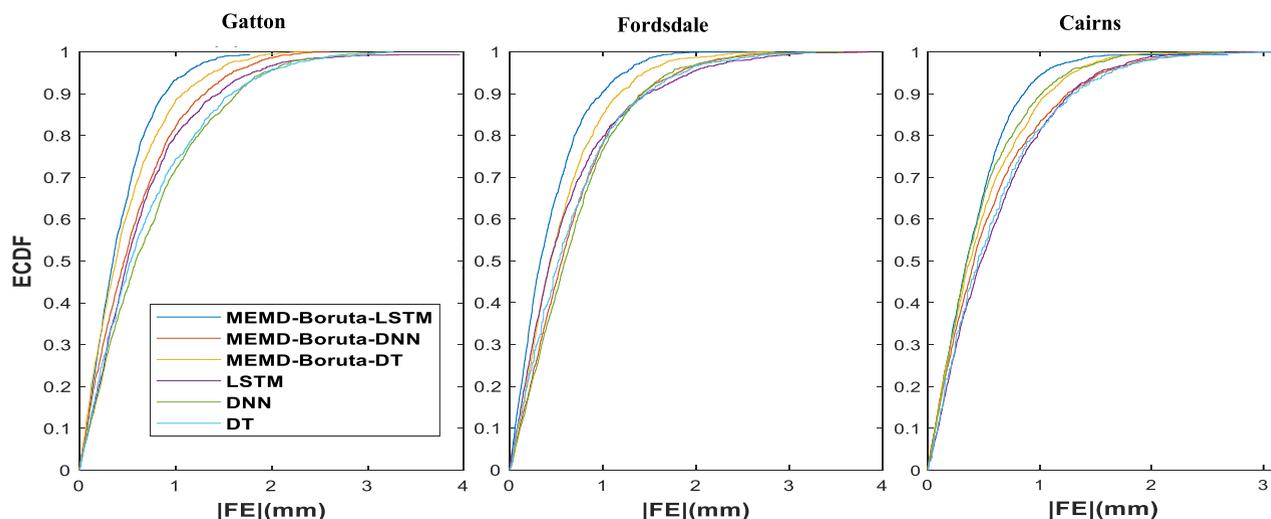
FIGURE 6. Empirical Cumulative Distribution Function (ECDF) of absolute forecasting error, |FE| (mm) of the testing data using MEMD-Boruta-LSTM *vs*. MEMD-Boruta-DNN, MEMD-Boruta-DT, and standalone LSTM, DNN, and DT models in forecasting *ET* for all study sites.

V. CONCLUSION

This study aims to design a novel deep learning multi-stage hybrid MEMD-Boruta-LSTM model as a practical tool to forecast daily *ET* using satellite and ground-based variables.

The Multivariate Empirical Mode Decomposition (MEMD) is incorporated with LSTM to decompose predictor variable data into IMFs and residuals and the Boruta-Random Forest (Boruta) feature selection method has been employed to screen the most correlated predictor variables to target variable *ET* in each IMFs and residuals. The daily predictor and target variable data (01 February 2003 to 19 April 2011) were extracted from GIOVANNI-AIRS, GLDAS model satellites, and the SILO ground database of the Queensland government. The test sites included Gatton, Fordsdale, and Cairns, which are located in drought-prone regions in Queensland, Australia. The integration of LSTM with MEMD and Boruta resulted in a novel multi-stage deep learning MEMD-Boruta-LSTM hybrid model whose performance was evaluated using statistical score metrics and compared with the other hybrid and standalone models namely, MEMD-Boruta-DNN, MEMD-Boruta-DT, LSTM, DNN, and DT based approaches.

The MEMD-Boruta-LSTM hybrid model yielded the highest values for normalized performance metrics: *r, NS, WI, LM* (see Table 6) and the lowest values for *RMSE, MAE, RRMSE,* and *APB* for all sites. Meanwhile, the results also revealed that all hybrid models (MEMD-Boruta-LSTM, MEMD-Boruta-DNN, MEMD-Boruta-DT) remarkably outperformed in comparison with the standalone models (LSTM, DNN, DT) in forecasting *ET* at all study sites (see Table 6). This comparison provides strong evidence to verify that MEMD decomposition and Boruta feature selection methods can be used effectively to improve the forecasting accuracy of any model. All findings of this study confirm that the proposed multi-stage deep hybrid MEMD-Boruta-LSTM model out-

performed the comparative hybrid and standalone models in forecasting *ET* at a daily forecasting horizon.

The novel proposed deep hybrid MEMD-Boruta-LSTM model can be practically employed for precise forecasting of *ET*. Evapotranspiration is the main causative natural phenomenon that contributes to the water losses from croplands. By multiplying forecasted *ET* with the relevant crop factor, which is a unique value for individual crops, the water loss due to evapotranspiration can be estimated in advance, which will be helpful in planning precise irrigation schedules for the future while avoiding the wastage of water resources in drought-prone areas. In addition, this multi-stage deep learning hybrid model used for forecasting *ET* is likely to lead to significant financial benefits to the farmers, in arid and semi-arid regions where agricultural practices are adversely affected by the scarcity of water resources.

VI. LIMITATION AND FUTURE RESEARCH

For this model development, data were extracted only for three sites within Queensland (as a case study) as it is impracticable to select more sites representing the whole drought-affected region in Australia or elsewhere. However, this pioneering study has produced a new modelling framework for *ET* forecasting and paves the way for future studies with a wider scope. For example, the geographic consistency of the MEMD-Boruta-LSTM hybrid model, together with its accuracy can be considered in future research. Moreover, the potential use of multi-stage MEMD-Boruta-LSTM for multi-step ahead daily *ET* (7 days, 15 days, 30 days) forecasting can be researched. Further, instead of the MEMD technique for data pre- Decomposition (VMD) technique [61] can be used with Boruta-LSTM to build up a new two-stage deep forecasting model to forecast *ET*.

REFERENCES

[1] L. S. Pereira, R. G. Allen, M. Smith, and D. Raes, "Crop evapotranspiration estimation with FAO56: Past and future," *Agricult. Water Manage.*, vol. 147, pp. 4–20, Jan. 2015.

[2] G. Huang, L. Wu, X. Ma, W. Zhang, J. Fan, X. Yu, W. Zeng, and H. Zhou, "Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions," *J. Hydrol.*, vol. 574, pp. 1029–1041, Jul. 2019, doi: 10.1016/j.jhydrol.2019.04.085.

[3] M. Wu, Q. Feng, X. Wen, R. C. Deo, Z. Yin, L. Yang, and D. Sheng, "Random forest predictive model development with uncertainty analysis capability for the estimation of evapotranspiration in an arid oasis region," *Hydrol. Res.*, vol. 51, no. 4, pp. 648–665, Aug. 2020, doi: 10.2166/nh.2020.012.

[4] V. Nourani, G. Elkiran, and J. Abdullahi, "Multi-step ahead modeling of reference evapotranspiration using a multi-model approach," *J. Hydrol.*, vol. 581, Feb. 2020, Art. no. 124434, doi: 10.1016/j.jhydrol.2019.124434.

[5] Y. Tikhamarine, A. Malik, A. Kumar, D. Souag-Gamane, and O. Kisi, "Estimation of monthly reference evapotranspiration using novel hybrid machine learning approaches," *Hydrol. Sci. J.*, vol. 64, no. 15, pp. 1824–1842, Nov. 2019, doi: 10.1080/02626667.2019.1678750.

[6] Z. Chen, S. Sun, Y. Wang, Q. Wang, and X. Zhang, "Temporal convolution-network-based models for modeling maize evapotranspiration under mulched drip irrigation," *Comput. Electron. Agricult.*, vol. 169, Feb. 2020, Art. no. 105206, doi: 10.1016/j.compag.2019.105206.

[7] P. de Oliveira e Lucas, M. A. Alves, P. C. de Lima e Silva, and F. G. Guimarães, "Reference evapotranspiration time series forecasting with ensemble of convolutional neural networks," *Comput. Electron. Agricult.*, vol. 177, Oct. 2020, Art. no. 105700, doi: 10.1016/j.compag.2020.105700.

[8] J. Zhang, Y. Zhu, X. Zhang, M. Ye, and J. Yan, "Developing a long short-term memory (LSTM) based model for predicting water table depth in agricultural areas," *J. Hydrol.*, vol. 561, pp. 918–929, Jun. 2018.

[9] S. Ghimire, R. C. Deo, N. Raj, and J. Mi, "Deep learning neural networks trained with MODIS satellite-derived predictors for long-term global solar radiation prediction," *Energies*, vol. 12, no. 12, p. 2407, Jun. 2019.

[10] M. Fu, T. Fan, Z. Ding, S. Q. Salih, N. Al-Ansari, and Z. M. Yaseen, "Deep learning data-intelligence model based on adjusted forecasting window scale: Application in daily streamflow simulation," *IEEE Access*, vol. 8, pp. 32632–32651, 2020.

[11] M. Gauch, F. Kratzert, D. Klotz, G. Nearing, J. Lin, and S. Hochreiter, "Rainfall-runoff prediction at multiple timescales with a single long short-term memory network," *Hydrol. Earth Syst. Sci.*, vol. 25, no. 4, pp. 2045–2062, 2021.

[12] J. Yin, Z. Deng, A. V. M. Ines, J. Wu, and E. Rasu, "Forecast of short-term daily reference evapotranspiration under limited meteorological variables using a hybrid bi-directional long short-term memory model (Bi-LSTM)," *Agricult. Water Manage.*, vol. 242, Dec. 2020, Art. no. 106386, doi: 10.1016/j.agwat.2020.106386.

[13] L. B. Ferreira and F. F. da Cunha, "Multi-step ahead forecasting of daily reference evapotranspiration using deep learning," *Comput. Electron. Agricult.*, vol. 178, Nov. 2020, Art. no. 105728, doi: 10.1016/j.compag.2020.105728.

[14] M. S. AL-Musaylh, R. C. Deo, Y. Li, and J. F. Adamowski, "Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting," *Appl. Energy*, vol. 217, pp. 422–439, May 2018, doi: 10.1016/j.apenergy.2018.02.140.

[15] R. Prasad, M. Ali, P. Kwan, and H. Khan, "Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation," *Appl. Energy*, vol. 236, pp. 778–792, Feb. 2019.

[16] N. ur Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proc. R. Soc. Lond. A, Math., Phys. Eng. Sci.*, vol. 466, no. 2117, pp. 1291–1302, 2009.

[17] X. Lang, Q. Zheng, Z. Zhang, S. Lu, L. Xie, A. Horch, and H. Su, "Fast multivariate empirical mode decomposition," *IEEE Access*, vol. 6, pp. 65521–65538, 2018.

[18] M. Ali, R. C. Deo, T. Maraseni, and N. J. Downs, "Improving SPI-derived drought forecasts incorporating synoptic-scale climate indices in multi-phase multivariate empirical mode decomposition model hybridized with simulated annealing and kernel ridge regression algorithms," *J. Hydrol.*, vol. 576, pp. 164–184, Sep. 2019.

[19] M. B. Kursa, A. Jankowski, and W. R. Rudnicki, "Boruta—A system for feature selection," *Fundamenta Informaticae*, vol. 101, no. 4, pp. 271–285, 2010.

[20] A. M. Ahmed, R. C. Deo, A. Ghahramani, N. Raj, Q. Feng, Z. Yin, and L. Yang, "LSTM integrated with Boruta-random forest optimiser for soil moisture estimation under RCP4. 5 and RCP8. 5 global warming scenarios," *Stochastic Environ. Res. Risk Assessment*, vol. 35, pp. 1–31, Jan. 2021.

[21] R. Prasad, R. C. Deo, Y. Li, and T. Maraseni, "Weekly soil moisture forecasting with multivariate sequential, ensemble empirical mode decomposition and boruta-random forest hybridizer algorithm approach," *Catena*, vol. 177, pp. 149–166, Jun. 2019.

[22] J. Qu, K. Ren, and X. Shi, "Binary grey wolf optimization-regularized extreme learning machine wrapper coupled with the Boruta algorithm for monthly streamflow forecasting," *Water Resour. Manage.*, vol. 35, no. 3, pp. 1029–1045, Feb. 2021.

[23] A. A. M. Ahmed, R. C. Deo, Q. Feng, A. Ghahramani, N. Raj, Z. Yin, and L. Yang, "Deep learning hybrid model with boruta-random forest optimiser algorithm for streamflow forecasting with climate mode indices, rainfall, and periodicity," *J. Hydrol.*, vol. 599, Aug. 2021, Art. no. 126350.

[24] B. Lyu, Y. Zhang, and Y. Hu, "Improving PM2.5 air quality model forecasts in China using a bias-correction framework," *Atmosphere*, vol. 8, no. 12, p. 147, Aug. 2017.

[25] I. Craig, A. Green, M. Scobie, and E. Schmidt, "Controlling evaporation loss from water storages," Nat. Centre Eng. Agricult., Univ. Southern Queensland, Toowoomba, QLD, Australia, Tech. Rep. 1000580/1, 2005.

[26] N. Ur Rehman and D. P. Mandic, "Filter bank property of multivariate empirical mode decomposition," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2421–2426, May 2011.

[27] N. E. Huang, M.-L. C. Wu, S. R. Long, S. S. P. Shen, W. Qu, P. Gloersen, and K. L. Fan, "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis," *Proc. R. Soc. Lond. A, Math. Phys. Eng. Sci.*, vol. 459, no. 2037, pp. 2317–2345, 2003.

[28] W. Hu and B. C. Si, "Soil water prediction based on its scale-specific control using multivariate empirical mode decomposition," *Geoderma*, vol.s. 193–194, pp. 180–188, Feb. 2013.

[29] M. B. Kursa and W. R. Rudnicki, "Feature selection with the Boruta package," *J. Stat. Softw.*, vol. 36, no. 11, pp. 1–13, 2010.

[30] J.-H. Hur, S.-Y. Ihm, and Y.-H. Park, "A variable impacts measurement in random forest for mobile cloud computing," *Wireless Commun. Mobile Comput.*, vol. 2017, pp. 1–13, Oct. 2017.

[31] C. Strobl, A.-L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis, "Conditional variable importance for random forests," *BMC Bioinform.*, vol. 9, no. 307, pp. 1–11, 2008.

[32] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, *arXiv:1406.1078*.

[33] N. K. Manaswi, *Deep Learning with Applications Using Python: Chatbots and Face, Object, and Speech Recognition with TensorFlow and Keras*, 1 ed. Berkeley, CA, USA: Apress L, 2018.

[34] J. Chen, G.-Q. Zeng, W. Zhou, W. Du, and K.-D. Lu, "Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization," *Energ. Convers. Manage*, vol. 165, pp. 681–695, Jun. 2018.

[35] X. Zhang, Q. Zhang, G. Zhang, Z. Nie, Z. Gui, and H. Que, "A novel hybrid data-driven model for daily land surface temperature forecasting using long short-term memory neural network based on ensemble empirical mode decomposition," *Int. J. Environ. Res. Public Health*, vol. 15, no. 5, p. 1032, May 2018, doi: 10.3390/ijerph15051032.

[36] *About my Region—Queensland*, DOAWE, Dept. Agricult., Water Environ, Canberra, ACT, Australia, 2021.

[37] Queensland. *Drought Declarations*. Accessed: Aug. 20, 2021. [Online]. Available: https://www.longpaddock.qld.gov.au

[38] W. Teng, H. Rui, B. Vollmer, R. de Jeu, F. Fang, G.-D. Lei, and R. Parinussa, "NASA Giovanni: A tool for visualizing, analyzing, and intercomparing soil moisture data," *Remote Sens. Terrestrial Water Cycle*, vol. 206, p. 331, Oct. 2014.

[39] A. Morshed, J. Aryal, and R. Dutta, "Environmental spatio-temporal ontology for the linked open data cloud," in *Proc. 12th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun.*, Jul. 2013, pp. 1907–1912.

[40] S. Ghimire, R. C. Deo, N. Raj, and J. Mi, "Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms," *Appl. Energy*, vol. 253, Nov. 2019, Art. no. 113541, doi: 10.1016/j.apenergy.2019.113541.

[41] D. R. Legates and G. J. McCabe, "Evaluating the use of 'goodness-of-fi' measures in hydrologic and hydroclimatic model validation," *Water Resour. Res.*, vol. 35, no. 1, pp. 233–241, Jan. 1999, doi: 10.1029/1998wr900018.

[42] D. R. Legates and G. J. McCabe, "A refined index of model performance: A rejoinder," *Int. J. Climatol.*, vol. 33, no. 4, pp. 1053–1056, Mar. 2013, doi: 10.1002/joc.3487.

[43] J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models part I—A discussion of principles," *J. Hydrol.*, vol. 10, no. 3, pp. 282–290, Apr. 1970, doi: 10.1016/0022-1694(70)90255-6.

[44] C. J. Willmott, S. M. Robeson, and K. Matsuura, "A refined index of model performance," *Int. J. Climatol.*, vol. 32, no. 13, pp. 2088–2094, 2012.

[45] H. Kling and H. Gupta, "On the development of regionalization relationships for lumped watershed models: The impact of ignoring sub-basin scale variability," *J. Hydrol.*, vol. 373, nos. 3–4, pp. 337–351, Jul. 2009.

[46] N. Ketkar, "Introduction to keras," in *Deep Learning With Python*. CA, USA: Springer, 2017, pp. 97–111.

[47] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, and M. Kudlur, "Tensorflow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Operating Syst. Design Implement. (OSDI)*, 2016, pp. 265–283.

[48] M. F. Sanner, "Python: A programming language for software integration and development," *J. Mol. Graph Model.*, vol. 17, no. 1, pp. 57–61, 1999.

[49] J. Quilty and J. Adamowski, "Addressing the incorrect usage of wavelet-based hydrological and water resources forecasting models for real-world applications with best practices and a new forecasting framework," *J. Hydrol.*, vol. 563, pp. 336–353, Aug. 2018.

[50] R. C. Deo, M. K. Tiwari, J. F. Adamowski, and J. M. Quilty, "Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model," *Stochastic Environ. Res. Risk Assessment*, vol. 31, no. 5, pp. 1211–1240, Jul. 2017.

[51] Q. Ouyang, W. Lu, X. Xin, Y. Zhang, W. Cheng, and T. Yu, "Monthly rainfall forecasting using EEMD-SVR based on phase-space reconstruction," *Water Resour. Manage.*, vol. 30, no. 7, pp. 2311–2325, May 2016.

[52] Y. Ren, P. N. Suganthan, and N. Srikanth, "A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods," *IEEE Trans. Sustain. Energy*, vol. 6, no. 1, pp. 236–244, Dec. 2014.

[53] W. C. Wang, D. M. Xu, K. W. Chau, and S. Chen, "Improved annual rainfall-runoff forecasting using PSO-SVM model based on EEMD," *J. Hydroinform.*, vol. 15, no. 4, pp. 1377–1390, 2013.

[54] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, 2008.

[55] S. García, S. Ramírez-Gallego, J. Luengo, J. M. Benítez, and F. Herrera, "Big data preprocessing: Methods and prospects," *Big Data Analytics*, vol. 1, no. 1, pp. 1–22, Dec. 2016, doi: 10.1186/s41044-016-0014-0.

[56] B. Komer, J. Bergstra, and C. Eliasmith, "Hyperopt-sklearn," in *Automated Machine Learning*. Cham, Switzerland: Springer, 2019, pp. 97–111.

[57] J. Bergstra, B. Komer, C. Eliasmith, D. Yamins, and D. D. Cox, "Hyperopt: A Python library for model selection and hyperparameter optimization," *Comput. Sci. Discovery*, vol. 8, no. 1, Jul. 2015, Art. no. 014008.

[58] S. Putatunda and K. Rama, "A comparative analysis of hyperopt as against other approaches for hyper-parameter optimization of XGBoost," in *Proc. Int. Conf. Signal Process. Mach. Learn. (SPML)*, 2018, pp. 6–10.

[59] K. Mohammadi, S. Shamshirband, M. H. Anisi, K. A. Alam, and D. Petković, "Support vector regression based prediction of global solar radiation on a horizontal surface," *Energy Convers. Manage.*, vol. 91, pp. 433–441, Feb. 2015.

[60] K. Mohammadi, S. Shamshirband, C. W. Tong, M. Arif, D. Petković, and S. Ch, "A new hybrid support vector machine–wavelet transform approach for estimation of horizontal global solar radiation," *Energy Convers. Manage.*, vol. 92, pp. 162–171, Mar. 2015.

[61] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.

[62] *AgriMetSoft (2019). Online Calculators*, AgriMetSoft, South Korea, 2020.

[63] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Res.*, vol. 30, no. 1, pp. 79–82, 2005.

[64] C. Ertekin and O. Yaldiz, "Comparison of some existing models for estimating global solar radiation for Antalya (Turkey)," *Energy Convers. Manage.*, vol. 41, no. 4, pp. 311–330, 2000, doi: 10.1016/s0196-8904(99)00127-2.

[65] S. Ghimire, R. C. Deo, N. Raj, and J. Mi, "Deep learning neural networks trained with MODIS satellite-derived predictors for long-term global solar radiation prediction," *Energies*, vol. 12, no. 12, p. 2407, Jun. 2019.

[66] T. Yang, L. Zhang, T. Kim, Y. Hong, D. Zhang, and Q. Peng, "A large-scale comparison of artificial intelligence and data mining (AI&DM) techniques in simulating reservoir releases over the upper Colorado region," *J. Hydrol.*, vol. 602, Nov. 2021, Art. no. 126723.

[67] H. Apaydin, H. Feizi, M. T. Sattari, M. S. Colak, S. Shamshirband, and K.-W. Chau, "Comparative analysis of recurrent neural network architectures for reservoir inflow forecasting," *Water*, vol. 12, no. 5, p. 1500, 2020.

[68] H. Fan, M. Jiang, L. Xu, H. Zhu, J. Cheng, and J. Jiang, "Comparison of long short term memory networks and the hydrological model in runoff simulation," *Water*, vol. 12, no. 1, p. 175, Jan. 2020.

[69] I.-F. Kao, Y. Zhou, L.-C. Chang, and F.-J. Chang, "Exploring a long short-term memory based encoder-decoder framework for multi-step-ahead flood forecasting," *J. Hydrol.*, vol. 583, Apr. 2020, Art. no. 124631.

[70] B. B. Sahoo, R. Jha, A. Singh, and D. Kumar, "Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting," *Acta Geophysica*, vol. 67, no. 5, pp. 1471–1481, 2019.

[71] Y. Wu, Y. Ding, Y. Zhu, J. Feng, and S. Wang, "Complexity to forecast flood: Problem definition and spatiotemporal attention LSTM solution," *Complexity*, vol. 2020, pp. 1–13, Mar. 2020.

[72] T. Yang, L. Zhang, T. Kim, Y. Hong, D. Zhang, and Q. Peng, "A large-scale comparison of artificial intelligence and data mining (AI&DM) techniques in simulating reservoir releases over the upper Colorado region," *J. Hydrol.*, vol. 602, Nov. 2021, Art. no. 126723.

[73] A. R. S. Kumar, M. K. Goyal, C. S. P. Ojha, R. D. Singh, P. K. Swamee, and R. K. Nema, "Application of ANN, fuzzy logic and decision tree algorithms for the development of reservoir operating rules," *Water Resour. Manag.*, vol. 27, no. 3, pp. 911–925, Feb. 2013.

[74] T. M. Shafapour, B. Pradhan, and M. N. Jebur, "Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS," *J. Hydrol.*, vol. 504, no. 8, pp. 69–79, Nov. 2013.

[75] Y. Cui, L. Song, and W. Fan, "Generation of spatio-temporally continuous evapotranspiration and its components by coupling a two-source energy balance model and a deep neural network over the Heihe river basin," *J. Hydrol.*, vol. 597, Jun. 2021, Art. no. 126176.

[76] S. Lee, D. Kaown, E. H. Koh, K. S. Ko, and K. K. Lee, "Delineation of groundwater quality locations suitable for target end-use purposes through deep neural network models," *Wiley Online Library*, vol. 50, pp. 0047–2425, Mar. 2021.

[77] X. H. Le, H. V. Ho, G. Lee, and S. Jung, "Application of long short-term memory (LSTM) neural network for flood forecasting," *Water*, vol. 11, no. 7, p. 1387, Jul. 2019.

W. J. M. LAKMINI PRARTHANA JAYASINGHE (Member, IEEE) received the B.Sc. degree (Hons.), in 2008, and the M.Phil. degree in mathematical sciences from the University of Kelaniya, Sri Lanka, in 2014. She is currently pursuing the Ph.D. degree in artificial intelligence with the University of Southern Queensland, Australia. She has been a Senior Lecturer with the Department of Mathematical Sciences, Wayamba University of Sri Lanka, since 2010. Her current research work is co-funded by the University of Southern Queensland, and the Wayamba University of Sri Lanka. Her research interests include deep learning, remote sensing, machine leaning, and mathematical modeling. She is a member of ADSN, NSF, and SLAAS. She has won Physical Sciences Award, in 2008, provided by SLAAS, Sri Lanka.

RAVINESH C. DEO (Senior Member, IEEE) is currently a Professor with the University of Southern Queensland, Australia. His research interests include data science, knowledge & data engineering, deep learning, artificial intelligence (AI), and machine learning (ML) for decision systems. Due to his research leadership, he was awarded internationally competitive fellowships, such as Queensland Smithsonian, Australia–China, JSPS, Chinese Academy of Science Presidential Fellowship, Australia–India Strategic Fellowship and Endeavour Fellowship. He is a 2021 Clarivate Highly Cited Researcher. He has supervised over 25 research degrees, received Employee Excellence Award, such as an Excellence in Research, an Excellence in Postgraduate Supervision, Publication Excellence and Teaching Commendation. He is on the Editorial Board of *Stochastic Environmental Research & Risk Assessment*, IEEE ACCESS, *Remote Sensing, Energies* and *Journal of Hydrologic Engineering*. He has published over 230 articles, *incl*. 170 journals, seven books in Elsevier, Springer and IGI and 23 book chapters with cumulative citations exceeding eight,100 with an H-index of 48 and SciVal Field-Weighted Citation Index over 3.0.

AFSHIN GHAHRAMANI received the master's and Ph.D. degrees from the Tokyo University of Agriculture and Technology. He is currently a Senior Researcher with the University of Southern Queensland, Australia. Previously, he worked as a Postdoctoral Fellow and a Research Scientist at CSIRO and the University of Tsukuba, Japan. He has published more than 40 research papers and have led more than ten research projects. His research interests include sustainable and efficient water and soil management by undertaking multidisciplinary research in the areas of water/soil/agricultural systems & modeling.

SUJAN GHIMIRE received the degree in mechanical engineering from Kathmandu University, Nepal, in 2002, the M.S. degree in renewable energy from the Institute of Engineering, Nepal, in 2007, and the Ph.D. degree in global solar radiation modeling and prediction using artificial intelligence from the University of Southern Queensland, Australia. He is currently an Adjunct Lecturer with University of Southern Queensland and is working as the Territory Customer Support Manager with John Deere Ltd., Australia. His current research interests include optimization algorithms, nature-inspired metaheuristics, machine learning, and feature selection problem for real world problems.

NAWIN RAJ received the B.Sc., B.Ed., PGDMA, and M.Sc. degrees from the University of the South Pacific in the area of computational fluid dynamics, and the Ph.D. degree from the School of Sciences, University of Southern Queensland (USQ), QLD, Australia, in 2015. From 2007 to 2010, he was a Lecturer with Fiji National University. From 2015 to 2016, he worked as a Learning Advisor at USQ, where he is currently a Lecturer. He is a member of Australian Mathematical Society and the Queensland College of Teachers. His research interests include artificial intelligence, deep learning, non-linear oscillation, computational fluid dynamics, and oceanography.

• • •