Deep multi-stage reference evapotranspiration forecasting model: Multivariate empirical mode decomposition integrated with Boruta-random forest algorithm

WJM Lakmini Prarthana Jayasinghe1, (Member, IEEE), Ravinesh C Deo2, (Senior Member, IEEE), Afshin Ghahramani3, Sujan Ghimire4 and Nawin Raj2

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

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

This work was supported by the School of Science, University of Southern Queensland, Australia and Wayamba University of Sri Lanka

ABSTRACT Evapotranspiration, as a combination of evaporation and transpiration of water vapour, is a primary component of the global hydrological cycle. It accounts for significant losses of soil moisture from the earth to the atmosphere. Thus, reliable methods to monitor and forecast evapotranspiration are required for decision-making in many sectors. 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 (≤ 17%), Absolute Percentage Bias (≤ 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, 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, Nouri, 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. Tikhrame 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 [72-75]. 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 [76], which is a common characteristic related with many types of hydrological time series [71]. 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 [77]. 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 [79]. The DNN model has been employed in water resource management e.g. development of spatial-temporally continuous evapotranspiration model [80], development of model for mapping suitable groundwater extraction location [81] and shown better performances compared to benchmark models. LSTM is also extensively used for flood forecasting [82], and predicting water table [83], 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 preprocessing [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 OVERVIWES

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 multivariate 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 formula 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 $x_t \in \mathbb{R}^n$ be the group of predictors for the set of $T$ and $y_r \in \mathbb{R}$ be the target for $n$ number of inputs, where $t = 1, 2, ..., T$.

I. Create a randomly ordered duplicated (shadow) variable, $x'_t$ for $x_t$, and then predict the target $y_r$.
II. Calculate Mean Decrease Accuracy (MDA) for every $x_t$ and $x'_t$ over all trees, $m_{\text{true}} = 500$ in this study [30, 31]:

$$MDA = \frac{1}{m_{\text{true}}} \sum_{m=1}^{m_{\text{true}}} \frac{\sum_{t=1}^{T} (I(y_r - f(x_t)) - \sum_{t=1}^{T} OOB(y_r - f(x'_t)))}{\sum_{t=1}^{T} OOB(y_r - f(x'_t)))}$$

(1)

where $I(*)$ is indicated function, $OOB(t)$ (Out-of-Bag) is a predictive error, $y_r = f(x_t)$ is predicted value before permuting and, $y'_r = f(x'_t)$ is predicted value after permuting.

III. Compute the $Z$-score as:

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

(2)

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

IV. Following that, predictors are identified as “Unimportant” when $Z - \text{score} < Z_{\text{max}}$ and "Confirmed" as important when $Z - \text{score} > Z_{\text{max}}$ 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 operate 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” $f_t$:

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

(3)

where $w_f$ is the weight matrix; $b_f$ is the bias vector and $\sigma(\cdot)$ 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 $\tilde{c}_t$, after updating information by using “input gate” $i_t$:

$$\tilde{c}_t = \text{tanh}(w_c[h_{t-1}, x_t] + b_c)$$

(4)

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

(5)

where $\text{tanh}(\cdot)$ 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 $\tilde{c}_t$:

$$c_t = f_t \ast c_{t-1} + i_t \ast \tilde{c}_t$$

(6)

IV. Finally, the output $h_t$ is derived using “output gate” $o_t$ and the cell state $c_t$:

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

(7)

$$h_t = o_t \ast \text{tanh}(c_t)$$

(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°E, 27.54°S, Fordsdale − 152.12°E, 27.72°S and Cairns − 145.75°E, 16.87°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

VOLUME XX, 2017
(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]. 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 (ε = 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-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 [62]. This feature selection algorithm is used to identify the significantly correlated predictor variables to the target variable in each IMFs and residuals using historical lagged data at (t-1).

Stage 3: After identifying the significantly correlated 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):

![FIGURE 1. Study sites in Queensland, Australia where the proposed MEMD-Boruta-LSTM model was implemented.](image1)

| Data source          | Name of input variable | Acronym | Unit          |
|----------------------|------------------------|---------|---------------|
| AIRS                 | Surface Temperature-Day | Tsd     | °C            |
| GLDAS Model          | Bare Soil Evaporation  | Ebs     | kg m⁻² s⁻¹    |
| GLDAS Model          | Transpiration          | TR      | kg m⁻² s⁻¹    |
| SILO-Ground based data | Maximum Temperature    | Tmax    | °C            |
| SILO-Ground based data | Minimum Temperature    | Tmin    | °C            |
| SILO-Ground based data | Radiation              | Ra      | MJ m⁻²        |
\[ X_n = \frac{X_{\text{actual}} - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \]  

(9)

where \( X_{\text{actual}} \), \( X_{\text{max}} \), and \( X_{\text{min}} \) 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 correlated 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 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 [63]. 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.

\[
r = \frac{\sum_{i=1}^{N}(ET_{OBS,i} - \bar{ET}_{OBS})(ET_{FOR,i} - \bar{ET}_{FOR})}{\sqrt{\sum_{i=1}^{N}(ET_{OBS,i} - \bar{ET}_{OBS})^2 \sum_{i=1}^{N}(ET_{FOR,i} - \bar{ET}_{FOR})^2}}, -1 \leq r \leq 1 \quad (10)
\]

II. Root Mean Square Error (RMSE; mm/day\(^{-1}\)): This measures the average model-performance error between predicted value (\(\bar{ET}_{FOR}\)) and observed value (\(ET_{OBS}\)) [64]. The RMSE value can range from 0 to \(\infty\) and it becomes zero for the best predictive models.

\[
RMSE = \frac{1}{N} \sum_{i=1}^{N} (ET_{FOR,i} - ET_{OBS,i})^2, 0 \leq RMSE < \infty \quad (11)
\]

III. Mean Absolute Error (MAE; mm/day\(^{-1}\)): This error value provides an assessment of the actual forecasting errors in terms of the total number of observations [62]. MAE can range from 0 to \(\infty\) 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 [65].

\[
MAE = \frac{1}{N} \sum_{i=1}^{N} |ET_{FOR,i} - ET_{OBS,i}|, 0 \leq MAE < \infty \quad (12)
\]

IV. Relative Root Mean Squared Error (RRMSE): The RRMSE is used to measure overall forecasting accuracy of the models and always gives positive values [62]. 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 [66].

\[
RRMSE = \frac{\frac{1}{N} \sum_{i=1}^{N} (ET_{FOR,i} - ET_{OBS,i})^2}{\frac{1}{N} \sum_{i=1}^{N} ET_{OBS,i}} \times 100 \quad (13)
\]

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 [67] 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 < NSE < 0 indicates that observed mean is a better predictor than the model [63].

\[
NS = 1 - \frac{\sum_{i=1}^{N} (ET_{OBS,i} - \bar{ET}_{OBS})^2}{\sum_{i=1}^{N} (ET_{FOR,i} - \bar{ET}_{FOR})^2}, -\infty < E_{NS} \leq 1 \quad (15)
\]

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.

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

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) [68]. The value for LM ranges from \(-\infty\) to 1, whereas this value equals one for best predictive models.

\[
LM = 1 - \frac{\sum_{i=1}^{N} \left( ET_{OBS,i} - \bar{ET}_{OBS} - ET_{FOR,i} + \bar{ET}_{FOR} \right)^2}{\sum_{i=1}^{N} \left( ET_{OBS,i} - \bar{ET}_{OBS} \right)^2}, -\infty < LM \leq 1 \quad (17)
\]

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 [69].

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

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 [70]. Perfect models will give value one for the KGE index [69].
where $CV_{OF_{i}}$ and $ET_{OF_{i}}$ are observed and forecasted $i$th value of the evapotranspiration $ET$, $CV_{OB_{i}}$ and $ET_{OB_{i}}$ 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. $104 (=8 \times 13)$ predictors) whereas 11 IMFs and a residual component (i.e. $96 (=8 \times 12)$ 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 correlated 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.

| Study site | Parameters used in MEMD | Number of initial predictor variables in each IMF and residual | Step 1: MEMD | Step2: Boruta Feature Selection | Total number of predictor variables identified after feature selection |
|-----------|--------------------------|-------------------------------------------------------------|-------------|-----------------------------|------------------------------------------------------------------|
| Gatton    | 500 Ensemble number (N)  | 8                                                           | 13          | 104 (13x8)                 |                                                                  |
| Fordsdale | 500 Amplitude of the added white noise (ε) | 8                                                           | 13          | 104 (13x8)                 |                                                                  |
| Cairns    | 500 Number of initial predictor variables | 8                                                           | 12          | 96 (12x8)                  |                                                                  |

**TABLE 2. Summarized results of MEMD process**

**TABLE 3. Summarized results of Boruta feature selection process**

(a) (b)

**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).
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) illustrates that the proposed MEMD-Boruta-LSTM model has scored the lowest APB 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-

| 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 |
|----------------------------------|-----------------|-----------------|-----------------|
| IMF1  Pat-1, Tsd-1, Tsn-1, Ebs-1, TR-1, Tmax-1, Tmin-1, Ra-1 | 6 significant inputs | 6 significant inputs | 6 significant inputs |
| IMF2  Pat-1, Tsd-1, Tsn-1, Ebs-1, TR-1, Tmax-1, Tmin-1, Ra-1 | 6 significant inputs | 6 significant inputs | 6 significant inputs |
| IMF3  Pat-1, Tsd-1, Tsn-1, Ebs-1, TR-1, Tmax-1, Tmin-1, Ra-1 | 6 significant inputs | 6 significant inputs | 6 significant inputs |
| IMF4  Pat-1, Tsd-1, Tsn-1, Ebs-1, TR-1, Tmax-1, Tmin-1, Ra-1 | 6 significant inputs | 6 significant inputs | 6 significant inputs |
| IMF5  Pat-1, Tsd-1, Tsn-1, Ebs-1, TR-1, Tmax-1, Tmin-1, Ra-1 | 6 significant inputs | 6 significant inputs | 6 significant inputs |
| IMF6  Pat-1, Tsd-1, Tsn-1, Ebs-1, TR-1, Tmax-1, Tmin-1, Ra-1 | 6 significant inputs | 6 significant inputs | 6 significant inputs |
| IMF7  Pat-1, Tsd-1, Tsn-1, Ebs-1, TR-1, Tmax-1, Tmin-1, Ra-1 | 6 significant inputs | 6 significant inputs | 6 significant inputs |
| IMF8  Pat-1, Tsd-1, Tsn-1, Ebs-1, TR-1, Tmax-1, Tmin-1, Ra-1 | 6 significant inputs | 6 significant inputs | 6 significant inputs |
| IMF9  Pat-1, Tsd-1, Tsn-1, Ebs-1, TR-1, Tmax-1, Tmin-1, Ra-1 | 6 significant inputs | 6 significant inputs | 6 significant inputs |
| IMF10 Pat-1, Tsd-1, Tsn-1, Ebs-1, TR-1, Tmax-1, Tmin-1, Ra-1 | 6 significant inputs | 6 significant inputs | 6 significant inputs |
| IMF11 Pat-1, Tsd-1, Tsn-1, Ebs-1, TR-1, Tmax-1, Tmin-1, Ra-1 | 6 significant inputs | 6 significant inputs | 6 significant inputs |

VOLUME XX, 2017
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 10.89% at Gatton, 16.21% and 15.06% at Fordsdale, 12.47% and 11.29% at Cairns respectively). 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 performance exceeds that of all other comparative hybrid DL and standalone models for all the study sites in Queensland.

**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
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 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).

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 outperformed 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.

ACKNOWLEDGMENT
The authors are grateful to NASA’s Goddard Online Interactive Visualization and Analysis Infrastructure (GIOVANNI) satellite and Scientific Information for Land Owners (SILO) for the meteorological data used and University of Southern Queensland (USQ), Australia and Wayamba university of Sri Lanka for funding this research. The authors are also thankful to the editor and reviewers for their valuable suggestions in improving the quality of the paper.

REFERENCES
[1] L. S. Pereira, R. G. Allen, M. Smith, and D. Raes, “Crop evapotranspiration estimation with FAO56: Past and future,” Agricultural Water Management, vol. 147, pp. 4-20, 2015.
[2] G. Huang et al., “Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions,” Journal of hydrology (Amsterdam), vol. 574, pp. 1029-1041, 2019, doi: 10.1016/j.jhydrol.2019.04.085.
[3] W. Min et al., “Random forest predictive model development with uncertainty analysis capability for the estimation of evapotranspiration in an arid oasis region,” Hydrology research, vol. 51, no. 4, pp. 648-665, 2020, doi: 10.2166/hh.2020.012.
[4] V. Nourani, G. Elkiran, and J. Abdullahi, “Multi-step ahead modeling of reference evapotranspiration using a multi-model approach,” Journal of hydrology (Amsterdam), vol. 581, p. 124434, 2020, 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,” Hydrological sciences journal, vol. 64, no. 15, pp. 1824-1842, 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,” Computers and electronics in agriculture, vol. 169, p. 105206, 2020, 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,” Computers and electronics in agriculture, vol. 177, p. 105700, 2020, doi: 10.1016/j.compag.2020.105700.
[8] J. Zhang, Y. Zhu, X. Zhang, M. Ye, and J. Yang, “Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas,” Journal of hydrology, vol. 561, pp. 918-929, 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, 2019.
[10] M. Fu, T. Fan, Z. a. 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,” Hydrology and Earth System Sciences, 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),” Agricultural water management, vol. 242, p. 106386, 2020, doi: 10.1016/agwat.2020.106386.
[13] L. B. Ferreira and F. F. da Cunha, “Multi-step ahead forecasting of daily reference evapotranspiration using deep learning,” Computers and electronics in agriculture, vol. 178, p. 105728, 2020, doi: 10.1016/j.compag.2020.105728.
[14] M. S. Al-Musaylyh, R. C. Deo, Y. Li, and J. F. Adamowski, “Two-stage particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting,” Applied energy, vol. 217, pp. 422-439, 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,” Applied energy, vol. 236, pp. 778-792, 2019.
[16] N. Rehman and D. P. Mandic, “Multivariate empirical mode decomposition,” Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 466, no. 2117, pp. 1291-1302, 2020.
[17] X. Lang et al., “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,” Journal of hydrology, vol. 576, pp. 164-184, 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 et al., “LSTM integrated with Boruta-random forest optimiser for soil moisture estimation under RCP4. 5 and RCP8. 5 global warming scenarios,” Stochastic Environmental Research and Risk Assessment, pp. 1-31, 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, 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 Resources Management, vol. 35, no. 3, pp. 1029-1045, 2021.
[23] A. M. Ahmed et al., “Deep learning hybrid model with Boruta-Random forest optimiser algorithm for streamflow forecasting with climate mode indices, rainfall, and periodicity,” Journal of Hydrology, vol. 599, p. 126350, 2021.
[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. 8, p. 147, 2017.
[25] I. Craig, A. Green, M. Scobie, and E. Schmidt, “Controlling evaporation loss from water storages,” 2005.
[26] N. Ur Rehman and D. P. Mandic, “Filter bank property of multivariate empirical mode decomposition,” IEEE transactions on signal processing, vol. 59, no. 5, pp. 2421-2426, 2011.
[27] N. E. Huang et al., “A confidence limit for the empirical mode decomposition and Hilbert spectral analysis,” Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, vol. 459, no. 2037, pp. 2319-2358, 2003.
[28] W. Hu and B. C. Si, “Soil water prediction based on its scale-empirical mode decomposition using deep learning,” Geoderma, vol. 193, pp. 180-188, 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 communications and mobile computing, vol. 2017, 2017.
[31] C. Strobl, A.-L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis, “Conditional variable importance for random forests,” BMC bioinformatics, vol. 9, no. 1, pp. 1-11, 2008.
K. Cho et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078, 2014.

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

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," Energy conversion and management, vol. 165, pp. 681-695, 2018, doi: 10.1016/j.enconman.2018.03.098.

X. Zhang, Q. Zhang, G. Zhang, Z. Nie, Z. Gui, and H. Que, "A Novel Feature Data-Driven Model for Daily Land Surface Temperature Forecasting Using Long Short-Term Memory Neural Network Based on Ensemble Empirical Mode Decomposition," International journal of environmental research and public health, vol. 15, no. 5, 1032, 2018, doi: 10.3390/ijerph15051032.

DOAWE. "About my region – Queensland." Department of Agriculture, Water and the Environment (accessed).

Queensland. 'Drought Declaratons' https://longpaddock.qld.gov.au (accessed).

W. Teng et al., "NASA Giovanni: A Tool for Visualizing, Analyzing, and Intercomparing Soil Moisture Data," Remote Sensing of the Terrestrial Water Cycle, vol. 206, p. 331, 2014.

A. Morshed, J. Aryal, and R. Dutta, "Environmental spatio-temporal ontology for the Linked open data cloud," in 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, 2013, IEEE, pp. 1907-1912.

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," Applied energy, vol. 253, p. 113541, 2019, doi: 10.1016/j.apenergy.2019.113541.

D. R. Legates and G. J. McCabe, "Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation," Water resources research, vol. 35, no. 1, pp. 233-241, 1999, doi: 10.1029/1998wr900255.

D. R. Legates and G. J. McCabe, "A refined index of model performance: a rejoinder," International journal of climatology, vol. 33, no. 4, pp. 1053-1056, 2013, doi: 10.1002/joc.3487.

J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models part I — A discussion of principles," Journal of Hydrology, vol. 10, no. 3, pp. 282-290, 1970, doi: https://doi.org/10.1016/0022-1694(70)90055-6.

G. Willmott, R. M. McRoberts, and K. Matsurra, "A refined index of model performance," International Journal of Climatology, vol. 32, no. 13, pp. 2088-2094, 2012.

H. Kling and H. Gupta, "On the development of regionalization relationships for lumped watershed models: The impact of ignoring sub-basin scale variability," Journal of Hydrology, vol. 373, no. 3-4, pp. 337-351, 2009.

N. Kekar, "Introduction to keras," in Deep learning with Python: Springer, 2017, pp. 97-111.

M. Abadi et al., "Tensorflow: A system for large-scale machine learning," in 12th [USENIX] symposium on Operating Systems Design and Implementation ([OSDI]16), 2016, pp. 265-283.

M. F. Sanner, "Python: a programming language for software integration and development," J Mol Graph Model, vol. 17, no. 1, pp. 57-61, 1999.

J. Vinyt 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," Journal of hydrology, vol. 563, pp. 336-353, 2018.

R. C. Deo, M. K. Tiwari, J. F. Adamowski, and J. M. Qi, "Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model," Stochastic environmental research and risk assessment, vol. 31, no. 5, pp. 1211-1240, 2017.

Q. Quyang, W. Lu, X. Xin, Y. Zhang, W. Cheng, and T. Yu, "Monthly rainfall forecasting using EEMD-SVR based on phase-space reconstruction," Water resources management, vol. 30, no. 7, pp. 2311-2325, 2016.

Y. Ren, P. Sugarthan, and N. Srikanth, "A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods," IEEE Transactions on Sustainable Energy, vol. 6, no. 1, pp. 236-244, 2014.

W.-c. Wang, D.-m. Xu, K.-w. Chau, and S. Chen, "Improved annual rainfall-runoff forecasting using PSO–SVM model based on EEMD," Journal of Hydroinformatics, vol. 15, no. 4, pp. 1377-1390, 2013.

Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-aided data analysis method," Advances in adaptive data analysis, vol. 1, no. 1, pp. 1-41, 2009.

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, p. 1, 2016, doi: 10.1186/s41044-016-0014-0.

B. Komers, J. Bergstra, and C. Elissamth, "Hyperopt-sklearn," in Automated Machine Learning: Springer, Cham, 2019, pp. 97-111.

J. Bergstra, B. Komers, C. Elissamth, D. Yamins, and D. D. Cox, "Hyperopt: a python library for model selection and hyperparameter optimization," Computational Science & Discovery, vol. 8, no. 1, p. 014008, 2015.

S. Putatunda and K. Rama, "A comparative analysis of hyperopt as against other approaches for hyper-parameter optimization of XGBoost," in Proceedings of the 2018 International Conference on Signal Processing and Machine Learning, 2018, pp. 6-10.

K. Mohammadi, S. Shamshirband, M. H. Anisi, K. A. Alam, and D. Petkovic, "Support vector regression based prediction of global solar radiation on a horizontal surface," Energy Conversion and Management, vol. 91, pp. 433-441, 2015.

K. Mohammadi, S. Shamshirband, C. W. Tong, M. Arif, D. Petkovic, and S. Ch, "A new hybrid support vector machine–wavelet transform approach for estimation of horizontal global solar radiation," Energy Conversion and Management, vol. 92, pp. 162-171, 2015.

K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," IEEE transactions on signal processing, vol. 62, no. 3, pp. 531-544, 2013.

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, 2019.

AgriMetSoft. "AgriMetSoft (2019). Online Calculators." AgriMetSoft (accessed).

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 research, vol. 30, no. 1, pp. 79-82, 2005.

D. R. Legates and G. J. McCabe Jr, "Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation," Water resources research, vol. 31, no. 1, pp. 235-244, 1995.

C. Ertekin and O. Yalızd, "Comparison of some existing models for estimating global solar radiation for Antalya (Turkey)," Energy conversion and management, vol. 41, no. 4, pp. 311-330, 2000, doi: 10.1016/j.enconman.2009.09.0127-2.

J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models part I—A discussion of principles," Journal of hydrology, vol. 10, no. 3, pp. 282-290, 1970.

D. R. Legates and G. J. McCabe, "Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation," Water resources research, vol. 35, no. 1, pp. 233-241, 1999.

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, 2019.

H. Kling and H. Gupta, "On the development of regionalization relationships for lumped watershed models: The impact of ignoring sub-basin scale variability," Journal of Hydrology, vol. 373, no. 3-4, pp. 337-351, 2009.

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 River," Journal of Hydrology, vol. 602, p. 126723, 2021.

H. Apaydın, 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, pp. 1500, 2020.
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, 2020.

L.-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," *Journal of Hydrology*, vol. 583, p. 124631, 2020.

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.

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, 2020.

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," *Journal of Hydrology*, vol. 602, p. 126723, 2021.

M. K. Goyal, C. Ojha, R. Singh, P. Swamee, and R. Nema, "Application of ANN, fuzzy logic and decision tree algorithms for the development of reservoir operating rules," *Water resources management*, vol. 27, no. 3, pp. 911-925, 2013.

M. S. Tehrany, 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," *Journal of Hydrology*, vol. 504, pp. 69-79, 2013.

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 Hethe River Basin," *Journal of Hydrology*, vol. 597, p. 126176, 2021.

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*, 0047-2425, 2021.

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, 2019.

J. Zhang, Y. Zhu, X. Zhang, M. Ye, and J. Yang, "Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas," *Journal of hydrology*, vol. 561, pp. 918-929, 2018.

W.J.M. LAKMINI PRARTHANA JAYASINGHE (IEEE Member) received the B.Sc. (Hons) (2008) and M.Phil. (2014) degrees in Mathematical Sciences from University of Kelaniya, Sri Lanka. She won Physical Science Award in 2008 given by SLAAS, Sri Lanka. She is currently pursuing the Ph.D. degree in artificial intelligence with the University of Southern Queensland, Australia. She has been a senior lecturer attached to the department of Mathematical Sciences at Wayamba University of Sri Lanka from 2010. Lakmini’s current research work is funded by University of Southern Queensland, Australia. Her research interests include deep learning, remote sensing, machine learning and mathematical modeling. She is a member of IEEE, ADSN, NSF and SLAAS.

RAVINESH C. DEO (Senior IEEE Member) is a Professor at University of Southern Queensland, Australia. His interests are data science, knowledge & data engineering, deep learning, artificial intelligence (AI) and machine learning (ML) for decision support systems. He was awarded internationally competitive fellowships: Queensland Smithsonian, Australia–China, JSPS, Chinese Academy of Science Presidential Fellowship, Australia–India Strategic Fellowship and Australian Endeavour Fellowship. He supervised over eighteen Ph.D/MSc degrees, has received Employee Excellence Award [Ronel Erwee Memorial Award for Excellence in Postgraduate Supervision], Faculty Research Excellence, IEEE Highly Cited Paper, Publication Excellence and Teaching Excellence Commendations. He was industry mentor through many funded grants worth $2.6M, a visiting professor internationally and is professionally affiliated under scientific bodies. He is on Editorial Boards of Stochastic Environmental Research & Risk Assessment, IEEE Access, Remote Sensing, Energies and Journal of Hydrologic Engineering. He published over 230 articles, incl. 170 journals (85% being Quartile 1), six books in Elsevier, Springer and IGI and 23 Book Chapters. His work has cumulative citation exceeding 7,300 with an h-Index 46 and SciVal Field-Weighted Citation Index 3.1.

AFSHIN GAHRAHMANI received the Master and PhD from Tokyo University of Agriculture and Technology. He is a Senior researcher at the University of Southern Queensland, Australia. Previously, he worked as a postdoctoral fellow and research scientist in CSIRO and University of Tsukuba, Japan. Dr Ghahramani as published more than 40 research papers and have led more than 10 research projects. His research supports sustainable and efficient water and soil management by undertaking multidisciplinary research in the areas of water/soil/agricultural systems & modelling.

SUJAN GHIMIRE received the degree in mechanical engineering from Kathmandu University, Nepal in 2002, the M.S. degree in renewable energy from institute of engineering, Nepal in 2007, and Ph.D. degrees in global solar radiation modelling and prediction using Artificial intelligence from The University of Southern Queensland (USQ), Australia. He is currently working as a territory customer support manager with the John Deere limited. His current research interests include optimization algorithms, nature-inspired metaheuristics, machine learning, and feature selection problem for real world problems.

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

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/