Journal of Hydrology 617 (2023) 129105

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Development of a TVF-EMD-based multi-decomposition technique integrated with Encoder-Decoder-Bidirectional-LSTM for monthly rainfall forecasting

Mehdi Jamei [a,*], Mumtaz Ali [b], Anurag Malik [c], Masoud Karbasi [d], Priya Rai [e], Zaher Mundher Yaseen [f,g]

[a] Faculty of Engineering, Shohadaye Hoveizeh Campus of Technology, Shahid Chamran University of Ahvaz, Dashte Azadegan, Iran
[b] UniSQ College, University of Southern Queensland 4350, Australia
[c] Punjab Agricultural University, Regional Research Station, Bathinda-151001, Punjab, India
[d] Water Engineering Department, Faculty of Agriculture, University of Zanjan, Zanjan, Iran
[e] Department of Soil and Water Conservation Engineering, College of Technology, G.B. Pant University of Agriculture and Technology, Pantnagar-263145, Uttarakhand, India
[f] Civil and Environmental Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia
[g] Interdisciplinary Research Center for Membranes and Water Security, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia

ARTICLE INFO

Keywords:
Rainfall forecasting, EDBi-LSTM
Empirical mode decomposition
Singular valued decomposition

ABSTRACT

Accurate forecasting of rainfall is extremely important due to its complex nature and enormous impacts on hydrology, floods, droughts, agriculture, and monitoring of pollutant concentration levels. In this study, a new multi-decomposition deep learning-based technique was proposed to forecast monthly rainfall in Himalayan region of India (i.e., Haridwar and Nainital). In the first stage, the original rainfall signals as the individual accessible datasets were decomposed into intrinsic mode decomposition functions (IMFs) through the time-varying filter-based empirical mode decomposition (TVF-EMD) technique, and then the significant lagged values were computed from the decomposed sub-sequences (i.e., IMFs) using the partial autocorrelation function (PACF). In the second stage, the PACF-based decomposed IMFs signals were again decomposed by the Singular Valued Decomposition (SVD) approach to reduce the dimensionality and enhance the forecasting accuracy. The machine learning approaches including the bidirectional long-short term memory reinforced with the Encoder-Decoder Bidirectional (EDBi-LSTM), Adaptive Boosting Regression (Adaboost), Generalized Regression Neural Network (GRNN), and Random Forest (RF) were used to construct the hybrid forecasting models. Also, several statistical metrics i.e., correlation coefficient (R), root mean square error (RMSE), and Nash-Sutcliffe efficiency (NSE) and graphical interpretation tools were employed to evaluate the hybrid ($T_{VF\text{-}EMD}$-S_{VD}-R_F, $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$, $T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$, and $T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}) and standalone counterpart (EDBi-LSTM, Adaboost, RF, and GRNN) models. The outcomes of monthly rainfall forecasting ascertain that the $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ in the Haridwar (R = 0.5870, RMSE = 118.4782 mm, and NSE = 0.3116) and Nainital (R = 0.9698, RMSE = 44.3963 mm, NSE = 0.9388) outperformed the benchmarking models.

1. Introduction

Increasing human population and industrialization have significant detrimental consequences on natural resources (Ahmed et al., 2021). Furthermore, natural calamities such as droughts, floods, storms, and tsunamis have become more common in recent years (Malik et al., 2020; Yaseen and Shahid, 2020). Hence, the modeling of climate extremes require precise time series analysis (Ashwini et al., 2021; Halder et al., 2021). Rainfall is truly the only input element for the hydrological cycle but its scarcity or excess on Earth, on the other hand, has an impact on massive flooding and severe droughts that occur at both long and short periods (Patra et al., 2012; Zeleke and Raes, 1999). Both long-and short-term rainfall forecasting is critical for the proper management of water resources and watershed sustainability (Alamgir et al., 2020;

* Corresponding author.
E-mail address: M.jamei@shhut.ac.ir (M. Jamei).

https://doi.org/10.1016/j.jhydrol.2023.129105
Received 6 May 2022; Received in revised form 22 November 2022; Accepted 8 January 2023
Available online 11 January 2023

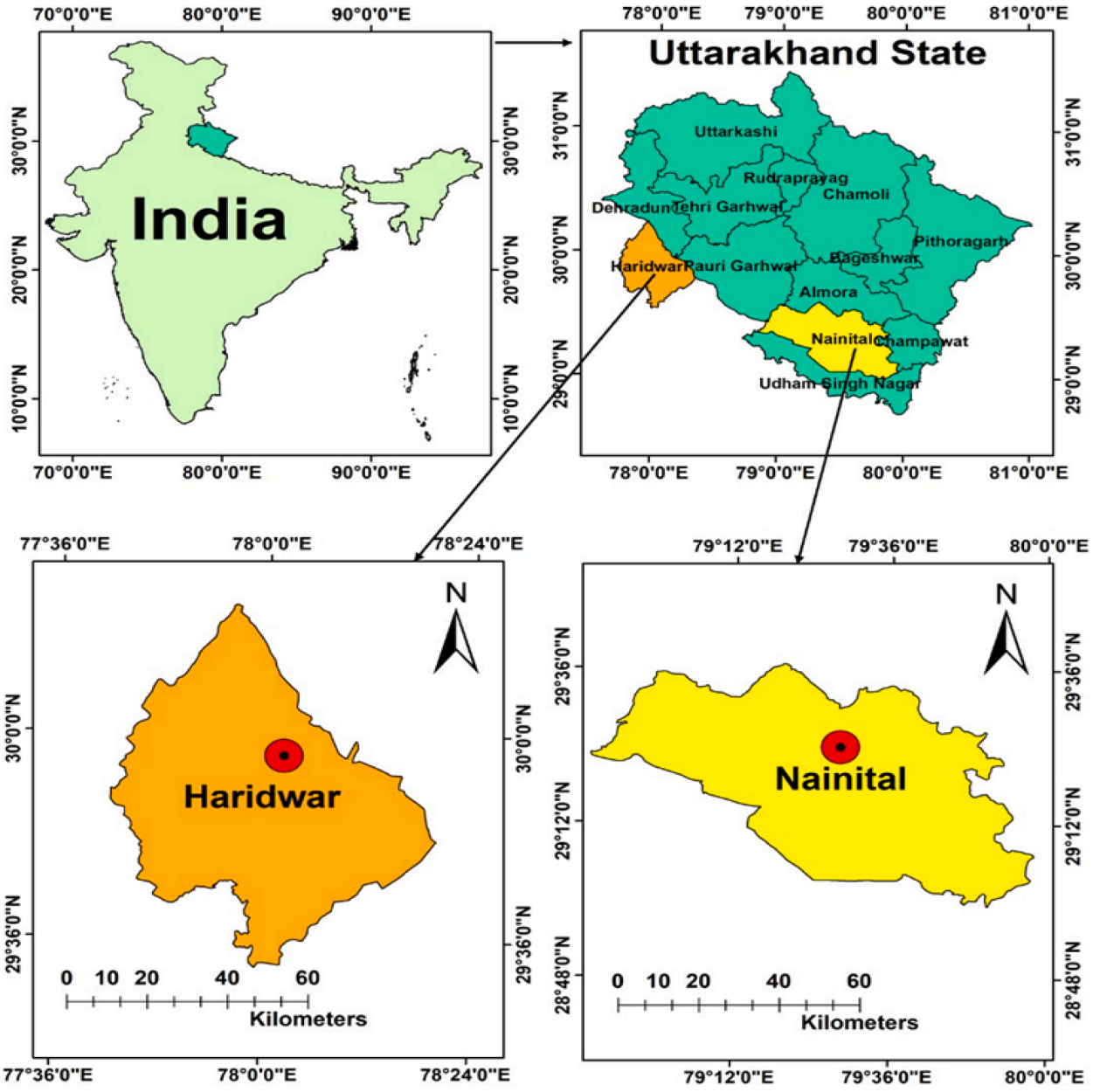


Fig. 1. The location map of two synaptic stations in Uttarakhand State of India.

Table 1
Summary of geographical properties of the study stations.

Site	Latitude	Longitude	Elevation	Data duration
Haridwar	29°55′26′′ N	78°03′04′′ E	276 *m*	1961–2015
Nainital	29°23′20′′ N	79°27′20′′ E	1945 *m*	1961–2015

Mohsenipour et al., 2020; Yaseen et al., 2019). For instance, the extent of rainfall determines the status of groundwater, which can then deliver water at any time. Furthermore, rainfall has a considerable impact on natural occurrences such as farming techniques.

Rainfall prediction can help people to avoid numerous natural disasters and save lives (Prasetya and Djamal, 2019). Rainfall forecasting accuracy can also aid in the creation of efficient structural and non-structural solutions for disaster relief. Rainfall time series prediction accuracy is determined by the uncertainty mitigation strategies (e.g., stochastic or deterministic) (Teegavarapu and Chandramouli, 2005). Deterministic dynamical prediction models are based on physical rules

Table 2
Descriptive statistics of rainfall datasets during1961 to 2015 at study sites.

Site	Haridwar		Nainital	
Phase	Training	Testing	Training	Testing
Number of values	495	165	495	165
Minimum	0.0	0.0	0.0	0.0
Q25%	4.1	4.75	9.1	9.1
Median	26.2	29	41.3	49.3
Q75%	122.2	157.9	178.6	186.4
Maximum	788.7	660	805.3	799.5
Range	788.7	660	805.3	799.5
Mean	95.29	98.99	116.5	131.9
Standard deviation	147.3	142.9	159.3	179.5
CV	154.6 %	144.4 %	136.8 %	136.1 %
Skewness	2.109	1.827	1.728	1.684
Kurtosis	4.289	3.055	2.403	2.222

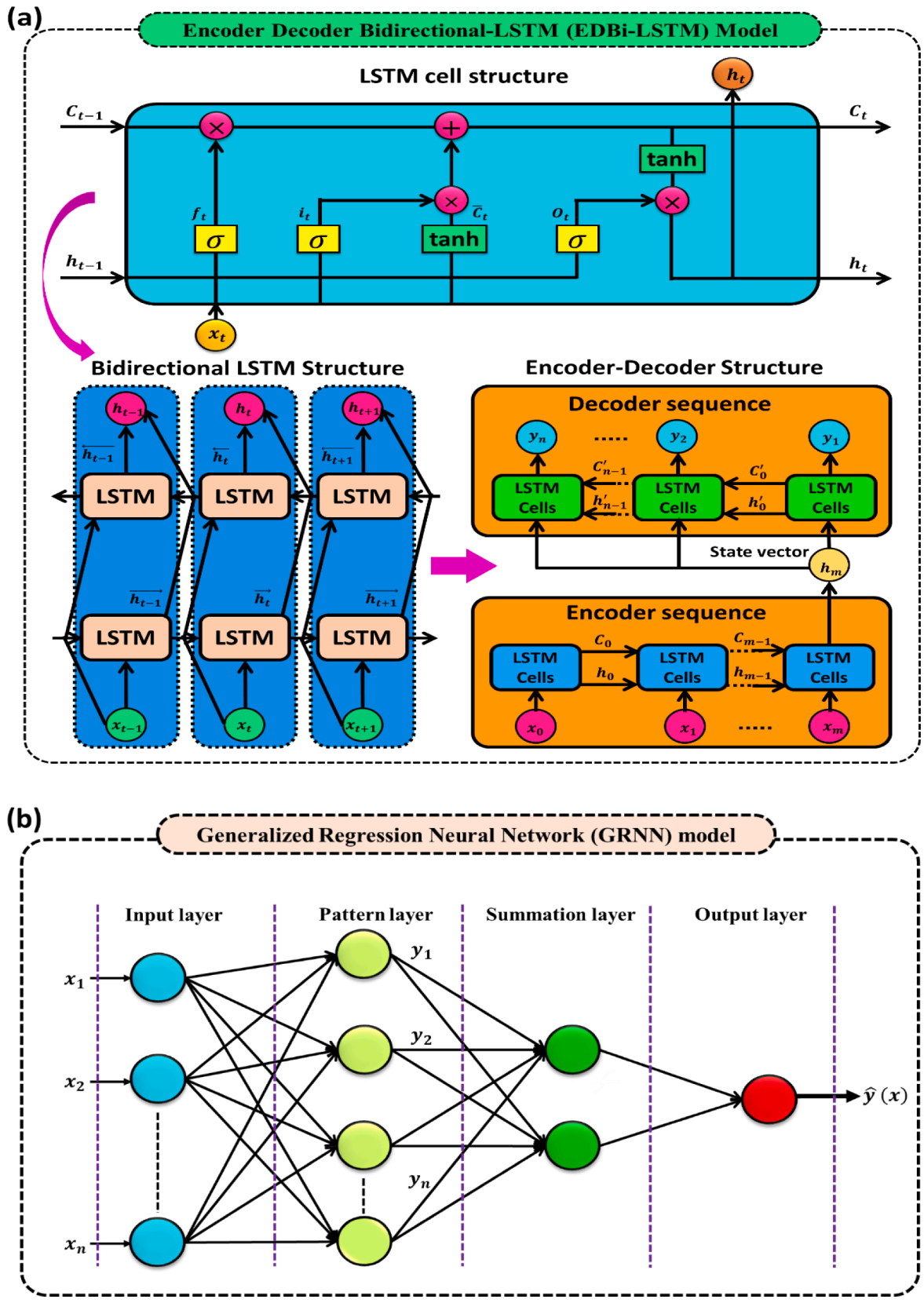


Fig. 2. The architecture of the EDBi-LSTM model (**a**) and GRNN model (**b**).

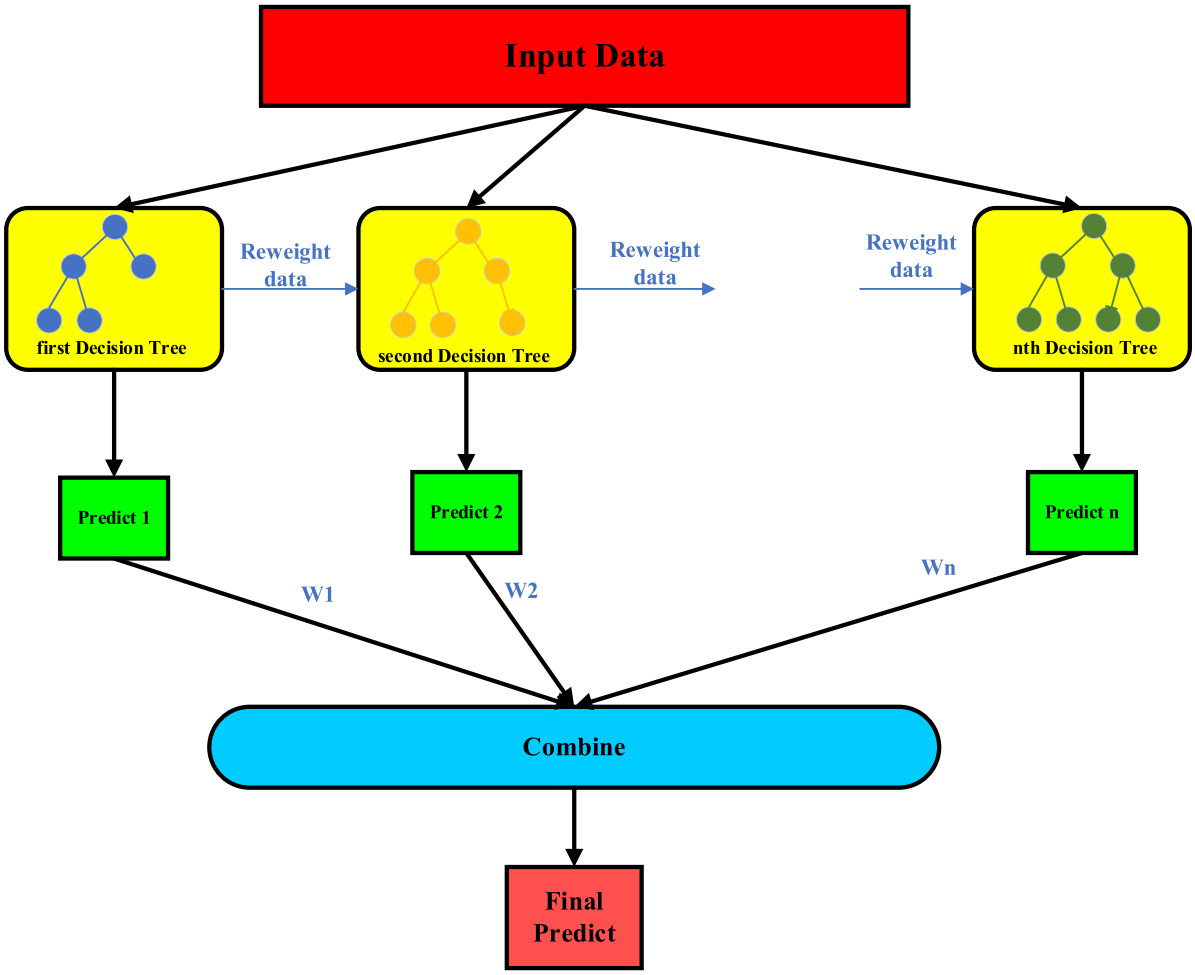


Fig. 3. Flowchart of AdaBoost algorithm.

governing the interaction between land, ocean, and atmosphere; they can anticipate changes in the pattern of rainfall as a result of changes in the Earth's atmosphere (Freeze, 1982). Rainfall prediction using the dynamical models is frequently subject to substantial errors at the local scale. On the other hand, statistical models are easy to set up and use, and they are more effective for forecasting smooth variations in rainfall on a small scale (Zaw and Naing, 2008). As a result, most scholars rely on statistical models for local-scale rainfall prediction. Time series models have also received adequate attention not just for hydrological tasks, but also for several other science and engineering tasks. Hence, this study is aimed at building reliable and resilient hybrid intelligence models that can simulate the current non-linear pattern in rainfall by analysing the historical data and comprehending the time series data's fundamental mechanisms.

Rainfall forecasting has received much attention to be simulated using advanced computer aid models including machine learning models (ML) (Adaryani et al., 2022; Diez-Sierra and del Jesus, 2020). The ML models are a well-established tool that has been applied successfully in different engineering disciplines and particularly in climate and hydrology applications (Danandeh Mehr et al., 2018; Fahimi et al., 2017; Omeje et al., 2021; Thamilselvan et al., 2022). The main merit of those models are able to comprehend the input features with the targeted parameters without essential knowledge of the main feature characteristics (Herath et al., 2020). Over the years, ML models with different versions have been applied for rainfall prediction and forecasting. For example, artificial neural networks (Ramirez et al., 2005), recurrent artificial neural networks (Hong, 2008), random forest (Yu et al., 2017), support vector machine (Lu and Wang, 2011), fuzzy logic (Patel and Parekh, 2014), extreme learning machine (Dash et al., 2018), logistic regression method (Moon et al., 2019), genetic programming (Cramer et al., 2017), and deep learning-based models took a serious advancement on this perspective (Basha et al., 2020; Hernández et al., 2016; Salehin et al., 2020). Although, multiple models have been explored for rainfall prediction. Yet, shortage in those standalone models is still associated, for example, internal parameters tuning, type of data used for simulating the rainfall, degree of the nonlinearity, and stochasticity. Researchers have started to explore new trends of research development to overcome the forgoing obstacles by using some advanced hybrid ML models (Diop et al., 2020; Pham et al., 2020; Yaseen et al., 2019), data time series decomposition (Ali et al., 2020), cross station simulation (Oleiwi et al., 2018), application of radar spatial data (Prudden et al., 2020).

Although the use of ML models and their many variants has been reported severally for rainfall prediction, many studies on rainfall prediction using integrative models have also been reported based on the use of feature selection and hybrid predictive models. The combination of advanced tuning methods and the most recently researched version of a deep learning model resulted in a significant improvement in modeling various engineering problems. Till now, rainfall process prediction using a hybrid encoder-decoder-bidirectional-LSTM model coupled with TVF-EMD-based multi-decomposition technique is not yet to be presented..

The primary purpose of this study is to see if a hybrid artificial intelligence (AI) model combined with a data processing method can be used efficiently for forecasting rainfall patterns on a monthly basis. The suggested model combines a new multi-composition technique comprised of the TVF-EMD and SVD combination coupled with the

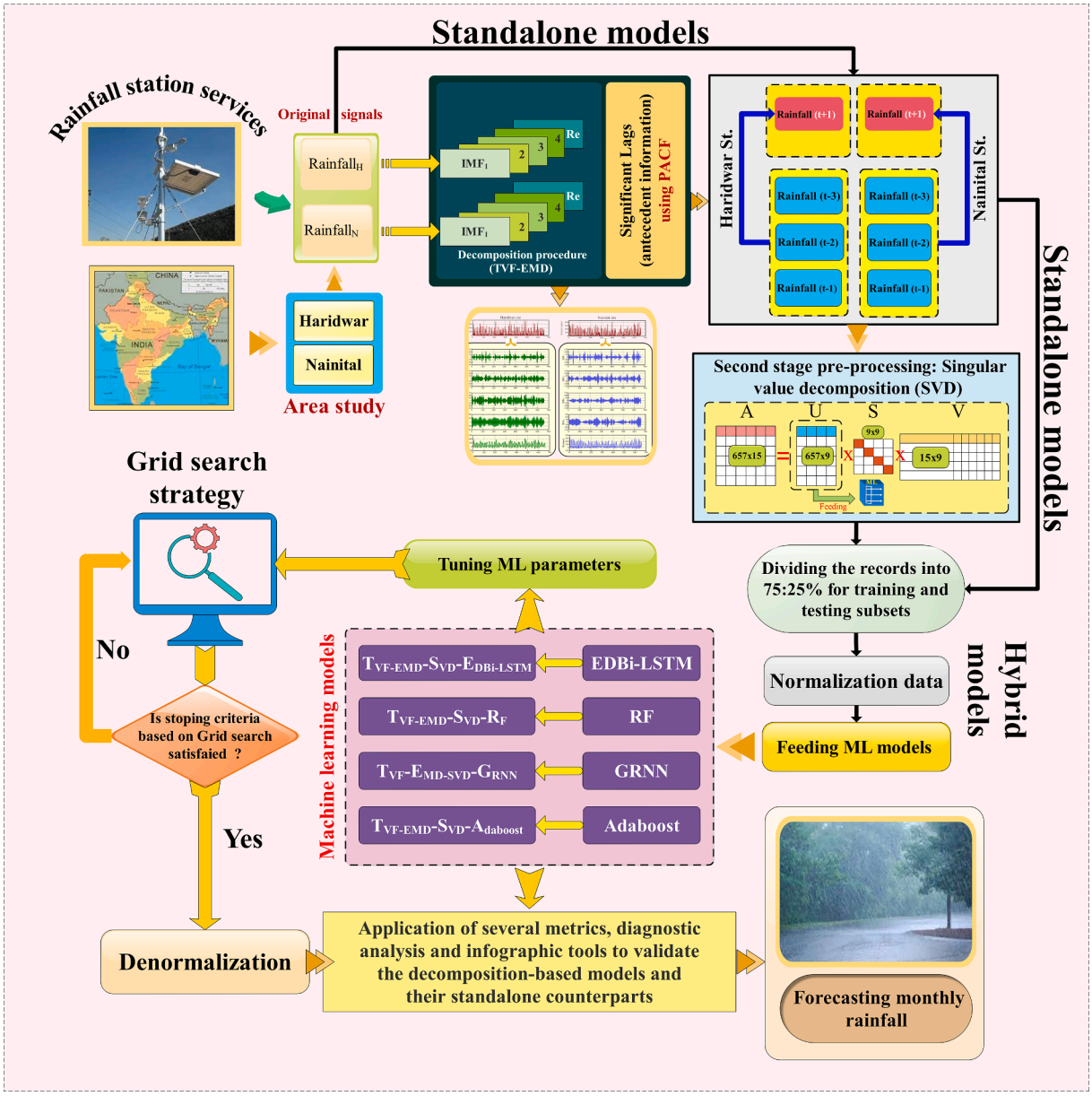


Fig. 4. A schematic view of the workflow to forecast monthly rainfall in Himalayan region of India using the multi-decomposition-based ML techniques.

EDBi-LSTM approach, making it a unique integrative computer aid model based on advanced learning and climate dataset analysis for improved forecasting accuracy.; It is also believed that the proposed hybrid AI model can capture the deepening of monthly rainfall patterns at Haridwar and Nainital locationsinUttarakhand State of India. Besides, the other advanced ML approaches i.e., Adaboost, RF, and GRNN were examined to evaluate the robustness of the EDBi-LSTM in both hybrid and standalone counterpart states. The developed model, in the end, can be considered an alternative strategy for simulating the climatological processes based on the potential of hybridized AI models optimized with nature-inspired algorithms.

2. Theoretical overviews

2.1. Introducing study area and datasets description

Fig. 1 demonstrates the location map of the study area along with the two selected stations namely Haridwar and Nainital for monthly rainfall forecasting using advanced machine learning paradigms. These stations are positioned in the Uttarakhand State of India which has an area of 53812.47 km^2 with a varying altitude of 145 m to 7796 m above the mean sea level (Malik and Kumar, 2021, 2020). The high altitude (>4572) areas of Uttarakhand are cold over the year and not reachable because of heavy rainfall (Nandargi et al., 2016). The average annual rainfall varies between 206 and 3955 mm and is mostly received (60–85 %) during the rainy season (June to September) (Malik et al., 2021a). The temperature is distributed from sub-zero to 43 °C in the summer season (April to June) and 0–15 °C in the winter season (October to February) of the Uttarakhand State (Nandargi et al., 2016). Table 1 provides information about the latitude, longitude, elevation, and data duration of the Haridwar and Nainital stations.

The monthly rainfall time series data for 55-years (1961–2015) was obtained from IMD (India Meteorological Department), Pune. It was separated into training (495 datasets) and testing (165 datasets) phases. Table 2 outlines the statistical properties of training and testing datasets. In training, the CV (coefficient of variation) was 154.6 % and 136.8 %, and in testing the CV was 144.4 % and 136.1 % at Haridwar and Nainital stations. It indicates the robustness of the data and is suitable for

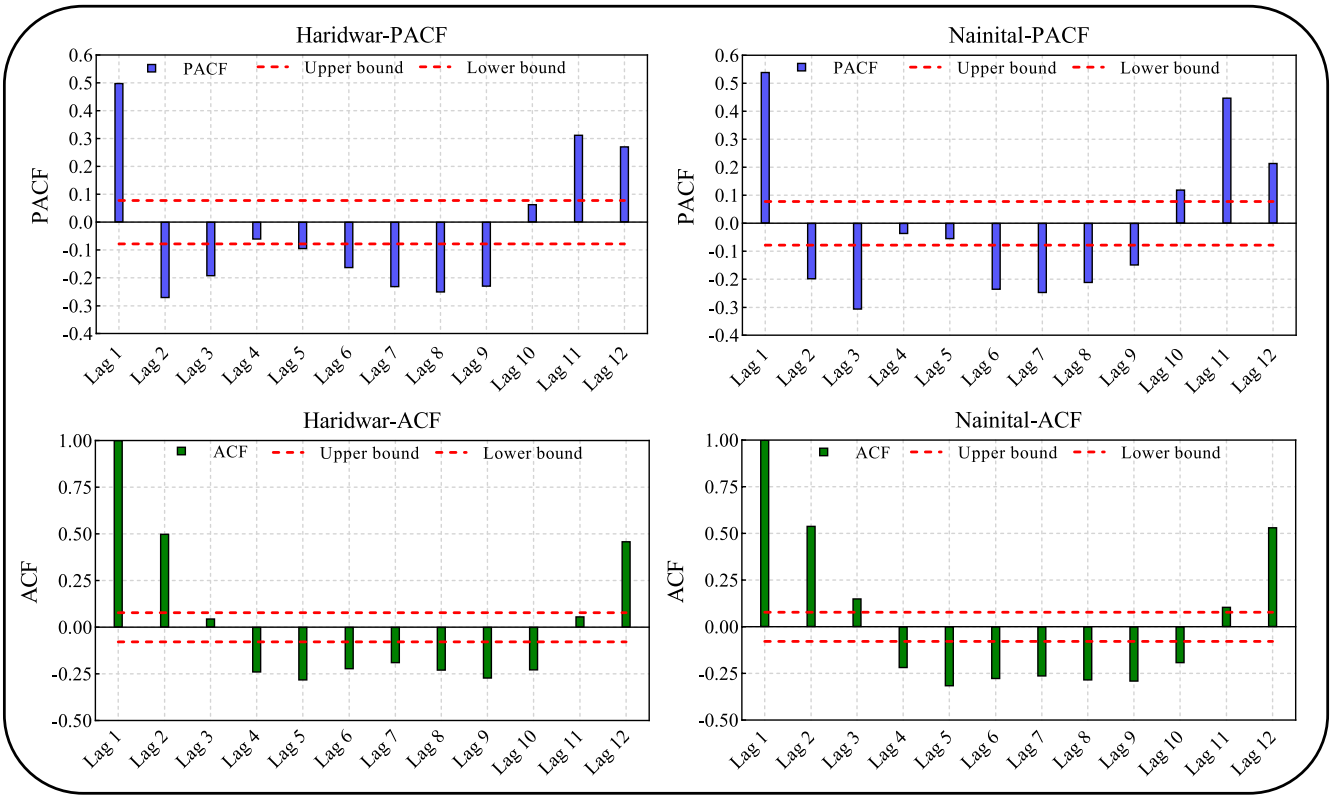


Fig. 5. Determination of significant lagged-time inputs to construct the hybrid and single AI models.

analysis.

2.2. Time varying filter based empirical mode decomposition (TVF-EMD)

The idea of the time-varying filter-based empirical mode decomposition (TVF-EMD) approach was given by Li et al. (2017) for handling the end effect and mode mixing problem within empirical mode decomposition (EMD), and achieving the shifting process through a time-varying filter. The bandwidth threshold and B-spline order are the main constraints that need to be selected properly for the application of the TVF-EMD technique (Li et al., 2017). If these constraints are inappropriate then the TVF-EMD model fails to optimize the mode mixing problem. Likewise, for performing time-varying filtering need to find out the local cut-off frequency (Zhang et al., 2021c). Moreover, the TVF-EMD approach yields more practicable results than the decomposition techniques (Wang et al., 2020; Zhang et al., 2021c). A time-varying filter is used to complete the shifting process of TVF-EMD and its implementation involves the following steps: (Li et al., 2017; Song et al., 2021b):

Step-1: Employ a B-spline approximation to compute the local cut-off frequency, which can be described as:

$$g_m^n(t) = \sum_{k=-\infty}^{\infty} c(k)\beta^n(t/m - k) \tag{1}$$

Here, $\beta^n(t)$ = B-spline function, and $c(k)$ = B-spline coefficient. The B-spline function is enlarged by a factor of m. The approximation result is determined by n, m, and $c(k)$. Thus, for given B-spline order n and knots m, B-spline approximation is to determine the $c(k)$ that reduces the approximation error $\left(\varepsilon_m^2\right)$ and computed as:

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

In which, $b_m^n(t) = \beta^n(t/m)$, $[\bullet]_{\uparrow m}$ = up-sampling operation by m, and $*$ = convolution operator, and the solution of $c(k)$ is

$$c(k) = \left[p_m^n * x\right]_{\downarrow m}(k) \tag{3}$$

where, $[\bullet]_{\downarrow m}$ = down-sampling operation by m, and p_m^n = pre-filter. Thus, we can write Eq. (1) as:

$$g_m^n(t) = \left[p_m^n * x\right]_{\downarrow m} * b_m^n(t) \tag{4}$$

In the above equation, the used B-spline approximation is a special form of low-pass filtering. Thus, the local cut-off frequency is projected from the input signal and employed to create the TVF. This process is carried out to obtain the local cut-off frequency, $\varphi_{bis}'(t) = \varphi_1'(t) + \varphi_2'(t)/2$, here, $\varphi_1'(t)$, and $\varphi_2'(t)$ are slow varying components. Realign the $\varphi_{bis}'(t)$ to solve the issue of intermittence i.e., noise, and obtain the final local cut-off frequency by interpolating among the peaks (or remainders).

Step-2: Filtering of the input signal using a time-varying filter (i.e., B-spline approximation filter) to achieve the local mean.

Step-3: Check whether the residual signal encounters the stopping conditions (or improves the stopping criteria) as follows (Wang et al., 2020):

$$\theta(t) = \frac{B_{Loughlin}(t)}{\varphi_{avg}(t)} \tag{5}$$

In which, $B_{Loughlin}(t)$ and $\varphi_{avg}(t)$ represents the Loughlin instantaneous bandwidth and weighted average instantaneous frequency of separate components. The present study utilized TVF-EMD for monthly rainfall forecasting on Haridwar and Nainital sites. Comprehensive information on the TVF-EMD model, readers can obtain from Li et al. (2017), and Zhang et al. (2018). For the EMD model, readers can obtain from Li et al. (2017), and Zhang et al. (2018).

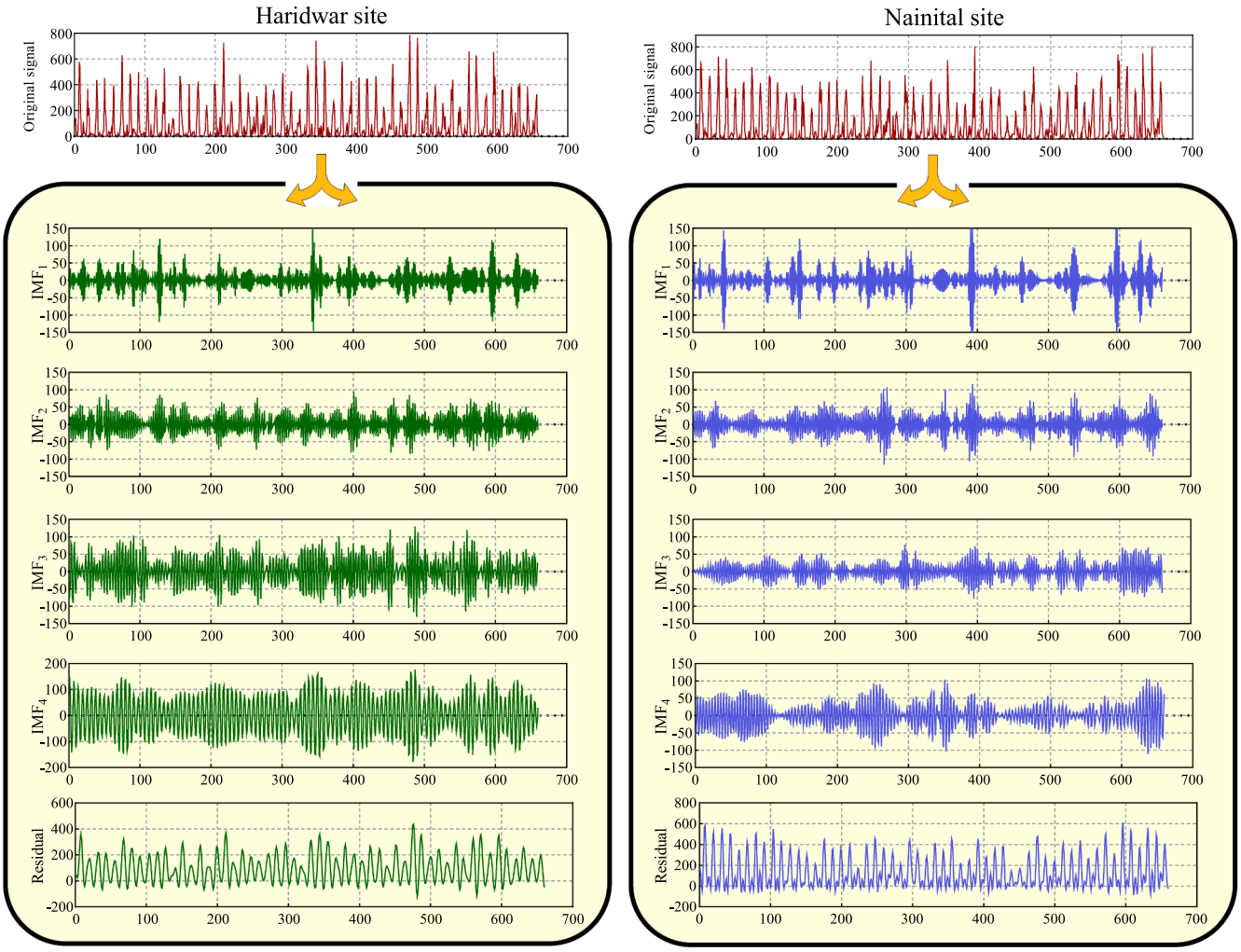


Fig. 6. Outcomes of decomposition procedure using the TVF-EMD including IMFs and residual for both original rainfall signals at Haridwar and Nainital stations.

Table 3
Multistage decomposition (TVF-EMD and SVD) adjustment for both study stations.

Site of study	Total number of IMFs	SVD decomposition setting	Values	Setting TVF-EMD	values	No. of data set in each site
Haridwar	15	Number of singular vectors	9	No. of decomposed IMFs (k)	5	After imposing the lags:
		Number of iterations	3	B-spline order	26	Training:493
		Block size	9	End_flag	0	Testing:164
		Setting mean center	False	Stopping criterion	0.1	
Nainital	15	Number of singular vectors	9	No. of decomposed IMFs (k)	5	After imposing the lags:
		Number of iterations	3	B-spline order	26	Training:493
		Block size	9	End_flag	0	Testing:164
		Setting mean center	False	Stopping criterion	0.1	

2.3. Singular valued decomposition (SVD)

The SVD is a matrix-based data decomposition technique and is usually employed in statistical and machine learning fields (Bretherton et al., 1992). It has been used for extracting the faint signal and proved to be robust in estimating the missing values from real-world data (Cong et al., 2013). The objective of the SVD algorithm is to find U, V, and W matrices. For any $m \times n$ particular input data in the form of matrix, A, is divided into $A = U\Omega V^T$, where $U = m \times m$ and orthonormal matrix, $\Omega = m \times n$ and diagonal matrix, and $V = n \times n$ and orthonormal matrix (Prasad et al., 2020). The W is the singular values of A. The matrix Ω contains singular values, σ_i, where $(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0)$. The truncated singular value decomposition for $1 \leq r \leq n$ is written as (Mardani et al., 2020):

$$A_k = Udiag\left(\sigma_1 \cdots \sigma_k, \frac{0, \cdots, 0}{n-k}\right) V^T \tag{6}$$

In the SVD, the A_k of rank, k can be found as:

$$A_k = \underbrace{\min}_{X:rank(X)=k} A - X_F \tag{7}$$

where F is Frobenius norm and defined as $\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}$.

2.4. Partial autocorrelation function

The selection of the optimal input features is one of the most important step in the construction of robust and reliable forecasting

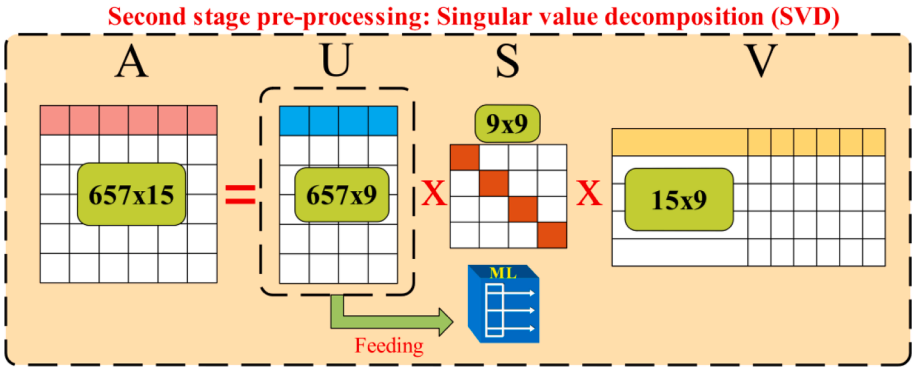


Fig. 7. Second-stage decomposition using the SVD technique.

Table 4
Hyper-parameters of all the AI approaches attained through a grid search strategy at study locations.

Model	Tuning parameters for ML techniques	
	Haridwar	Nainital
$T_{VF\text{-}EMD}$-S_{VD}-R_F	Tree number = 1500, Maximum Number of features = 20	Tree number = 1000, Maximum Number of features = 20
RF	Tree number = 1000, Maximum Number of features = 15	Tree number = 1000, Maximum Number of features = 15
$T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$	Number of Layers: 3, Learning Rate: 0.004, Neurons number: 300, Epochs: 300, Training Algorithm: Adam, activation='relu', Batch Size: 128	Number of Layers: 2, Learning Rate: 0.0008, Neurons number: 200, Epochs: 100, Training Algorithm: Adam, activation='relu', Batch Size: 128
EDBi-LSTM	Number of Layers: 4, Learning Rate: 0.06, Neurons number: 150, Epochs: 275, Training Algorithm: Adam, activation='relu', Batch Size: 128	Number of Layers: 2, Learning Rate: 0.0008, Neurons number: 200, Epochs: 100, Training Algorithm: Adam, activation='relu', Batch Size: 128
$T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$	Learning rate = 075, N_estimators = 100, loss function="Linear", subsample = 0.204	Learning rate = 0.90, N_estimators = 1500, loss function="Linear", subsample = 0.204
Adaboost	Learning rate = 0.03, N_estimators = 10, loss function="Linear", subsample = 0.916	Learning rate = 0.8, N_estimators = 5, loss function="Linear", subsample = 0.702
$T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}	Spread:0.15, Input neurons number:9	Spread:0.15, Input neurons number:9
GRNN	Spread:0.2, Input neurons number:3	Spread:0.15, Input neurons number:3

models. This research utilized a partial autocorrelation function (PACF) as a feature selection tool for nominating the significant input (or lags) from monthly rainfall time-series at 5 % confidence interval (upper and lower bounds). Mathematical PACF is expressed as (Hadi et al., 2019; Malik et al., 2021d):

$$PACF_{k,k} = \frac{ACF - \sum_{j=1}^{k-1} PACF_{k-1,j} ACF_{k-1}}{1 - \sum_{j=1}^{k-1} PACF_{k-1,j} ACF_{k-1}} \quad (8)$$

In which, ACF represents the autocorrelation function (ACF), and is computed as$ACF_k = \frac{\sum_{t=1}^{n-k}(x_t-\bar{x})(x_{t+k}-\bar{x})}{\sum_{t=1}^{n}(x_t-\bar{x})}$, here, x_t defines the value of the variable at time t, x_{t+k} describes the value of the variable at time $t+k$. Furthermore, k denotes the lag number with series x_t, $\bar{x}$indicates the mean of the whole rainfall data series and n outlines the number of observations.

2.5. Encoder-decoder bidirectional-LSTM (EDBi-LSTM)

Cho et al. (2014) proposed the concept of encoder-decoder bidirectional-long short-term memory (EDBi-LSTM) to diagnose models on different input and output time stages. Fig. 2 illustrates the typical structure of the EDBi-LSTM model which include the principles of the simple LSTM, Bidirectional-LSTM, and encoder-decoder sequence. The simple LSTM is an advanced version of a recurrent neural network (RNN) and learns the system based on chronological data (Hochreiter and Schmidhuber, 1997). The LSTM model gained popularity in different fields across the world (Feng et al., 2020; Ferreira and da Cunha, 2020; Kim and Kim, 2020; Li et al., 2022; Livieris et al., 2020; Seng et al., 2021; Xiang et al., 2020; Yin et al., 2020). The basic network of the LSTM model includes the forget gate, input gate, and output gate (**see** Fig. 2**a**). The flow of information inside the memory block, the

Table 5
Performance evaluation for hybrid and standalone AI approaches based on goodness-of-fit metrics at Haridwar.

Model	Metrics	R	RMSE	MAE	NSE	KGE	I_A
$T_{VF\text{-}EMD}$-S_{VD}-R_F	Training	0.9942	18.6262	13.4111	0.9840	0.9278	0.9957
	Testing	0.9475	48.1584	34.0708	0.8863	0.8314	0.9657
$T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$	Training	0.9838	26.5828	19.1662	0.9675	0.9605	0.9915
	Testing	**0.9688**	**35.4671**	**25.1577**	**0.9383**	**0.9422**	**0.9836**
$T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$	Training	0.9719	42.5195	37.3212	0.9168	0.7664	0.9755
	Testing	0.9425	53.6252	42.5407	0.8590	0.7528	0.9553
$T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}	Training	0.9918	20.0592	14.2333	0.9815	0.9444	0.9951
	Testing	0.9399	51.5689	34.8929	0.8696	0.8089	0.9603
RF	Training	0.9360	59.5204	38.7224	0.8370	0.7306	0.9448
	Testing	0.5495	120.1979	77.9971	0.2915	0.4288	0.7024
LSTM	Training	0.6811	108.5396	64.8198	0.4579	0.5288	0.7824
	Testing	**0.5870**	**118.4782**	**72.4846**	**0.3116**	**0.3927**	**0.7000**
Adaboost	Training	0.7109	104.4843	67.9237	0.4977	0.5987	0.8195
	Testing	0.5321	126.4811	79.4903	0.2155	0.4831	0.7205
GRNN	Training	0.5697	123.2012	80.2072	0.3016	0.2769	0.5958
	Testing	0.5674	119.7970	81.8863	0.2962	0.2691	0.5994

Table 6
Performance evaluation for hybrid and standalone AI approaches based on goodness-of-fit metrics at Nainital.

Model	Metrics	R	RMSE	MAE	NSE	KGE	I_A
$T_{VF\text{-}EMD}$-S_{VD}-R_F	Training	0.9906	25.9330	17.8388	0.9735	0.9019	0.9927
	Testing	0.9577	63.5489	46.9782	0.8746	0.7473	0.9584
$T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$	Training	0.9853	27.6399	20.1840	0.9699	0.9747	0.9925
	Testing	**0.9698**	**44.3963**	**28.5545**	**0.9388**	**0.9614**	**0.9846**
$T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$	Training	0.9673	51.5557	46.2244	0.8953	0.7334	0.9687
	Testing	0.9486	67.0786	54.7749	0.8602	0.7245	0.9566
$T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}	Training	0.9857	29.5257	19.6848	0.9657	0.9058	0.9905
	Testing	0.9505	68.0454	45.7621	0.8562	0.7220	0.9521
RF	Training	0.9471	57.3639	38.1771	0.8704	0.7778	0.9584
	Testing	0.6496	136.9900	82.9661	0.4171	0.4826	0.7565
LSTM	Training	0.7869	99.5844	63.1908	0.6095	0.6262	0.8546
	Testing	**0.6802**	**132.3965**	**81.4487**	**0.4556**	**0.5302**	**0.7856**
Adaboost	Training	0.7384	107.5139	70.6869	0.5449	0.6385	0.8387
	Testing	0.6086	143.4351	83.7947	0.3610	0.4974	0.7500
GRNN	Training	0.6834	118.4066	79.6756	0.4480	0.4475	0.7423
	Testing	0.6606	138.1628	85.5346	0.4071	0.3979	0.7204

information stored in a cell, and the output streams of the cell into the rest of the networks are controlled through these gates (Zhang et al., 2021c). The output from these gates is expressed as (Livieris et al., 2020; Yin et al., 2020):

$$f_t = \sigma\left(w_f \cdot [h_{t-1} x_t] + b_f\right) \tag{9}$$

$$i_t = \sigma(w_i \bullet [h_{t-1} x_t] + b_i) \tag{10}$$

$$\overline{C}_t = tanh(w_c \bullet [h_{t-1} x_t] + b_c) \tag{11}$$

$$C_t = f_t \times C_{t-1} + i_t \times \overline{C}_t \tag{12}$$

$$O_t = softsign(w_o \bullet [h_{t-1} x_t] + b_0) \tag{13}$$

$$h_t = O_t \times RLU(C_t) \tag{14}$$

In these Eqs., f_t, i_t, $\overline{C}_t$, C_t, O_t, and h_t represents the forget gate, input gate, new cell state candidate vectors, cell state, output gate, and final output. x_t defines the input at time t, σ indicates the sigmoid function of f_t and i_t. Furthermore, w_f, w_i, w_c, w_o, and b_f, b_i, b_c, and b_o describes the weights matrix and bias vectors of f_t, i_t, $\overline{C}_t$, and O_t, respectively. C_{t-1} represents the old memory cell unit at $t-1$ time, and h_{t-1} outlines the output of the hidden state at $t-1$ time. The *tanh* (hyperbolic: −1, 1), *softsign*, and RLU (Rectified Linear Unit) state the activation function of $\overline{C}_t$, O_t, and h_t units.

After that bidirectional-LSTM model was constructed to handle the gradient explosion and information morphing drawbacks during the backpropagation of LSTM (Graves and Schmidhuber, 2005; Hu and Zhang, 2018). Again, the predictive efficacy of bidirectional-LSTM was improved through encoder-decoder sequences. The EDBi-LSTM model is capable to process the natural language, and can also read & create a sequence of arbitrary length (Kurata et al., 2016). Additionally, the EDBi-LSTM model works in a symbiotic manner and the input sequential data is trained by the RNN encoder-decoder networks to generate the mapped output sequence. Recently, this model gained popularity in different domains because of its high predictive accuracy (Fan et al., 2019; Karbasi et al., 2022a, 2022b; Park et al., 2018; Wang and Zhang, 2018). Therefore, the present study explored the viability of the EDBi-LSTM model in monthly rainfall forecasting at Haridwar and Nainital locations. The EDBi-LSTM model includes two different parts (i) one for the interpretation of the input information of the sequence, and encoding the fixed-length vector, and (ii) a second for decoding the vector and outputting the forecasted sequence. The cell state stores the output of the encoded LSTM on m time steps, and supply the same as input to the decoder LSTM on n time steps (Fu et al., 2021). Accordingly, it can be stated that in each update, the decoder feeds the model so that the output from the preceding update is considered as the input of the recent update because this sequence-to-sequence network solves the time-step issue (Bian et al., 2020; Xiang et al., 2020; Zhang et al., 2021a).

2.6. Generalized regression neural network (GRNN)

The GRNN is a type of radial basis function (RBF) used to classify and control the nonlinear systems (Specht, 1991). The basic structure of the GRNN model comprises with input, pattern, summation, and output layers (**see** Fig. 2**b**). The output of the model is calculated based on the maximum probability principle (Cai et al., 2021). The optimal parameters of the GRNN model were obtained during the forward propagation step. The input layer is connected to the pattern layer and defines a training pattern. The output of the pattern layer is supplied to the summation layer which involves only two neurons namely S-summation (first), and D-summation (second). The final output of the GRNN model is obtained as (Ding et al., 2019; Specht, 1991):

$$\widehat{y}(x) = \frac{\sum_{i=1}^{n} y_i \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^{n} \exp\left(-\frac{D_i^2}{2\sigma^2}\right)} \tag{15}$$

where, $\widehat{y}$ is the estimated output, y_i denotes the output of the input sample, $\exp\left(-\frac{D_i^2}{2\sigma^2}\right)$ represents the activation function. D_i^2 defines the Euclidean distance from x, and computed as $D_i^2 = (x - x_i)^T (x - x_i)$, where, x is the input sample, x_i is training sample, and T state matrix transpose. Here, the GRNN model is discussed in brief, for more information refers to Specht (1991).

2.7. Random forest (RF)

The RF technique is a classification and regression technique that entails the construction of a series of tree predictors, each of which is created using a random vector independent from the input vector. In regression, the tree predictor uses numerical values rather than the class labels from the random forest classifier (Breiman, 1999). Random forest is a collection of CART model trees that have been proposed to improve its performance. These trees are constructed from the set of input vectors using the Bootstrap sampling method. Thus, the selected input of all of them follows the same distribution. Despite their unique and independent structure, the tree sets have a high correlation with each other, which, together with their strength and a large number, causes the convergence of the model error. Also, the mentioned features, along with estimating the internal error of the model, increase its ability to check more features and select the most effective parameters. In this tree network, the pruning process is not performed, so the model error is limited to avoid over-fitting the model.

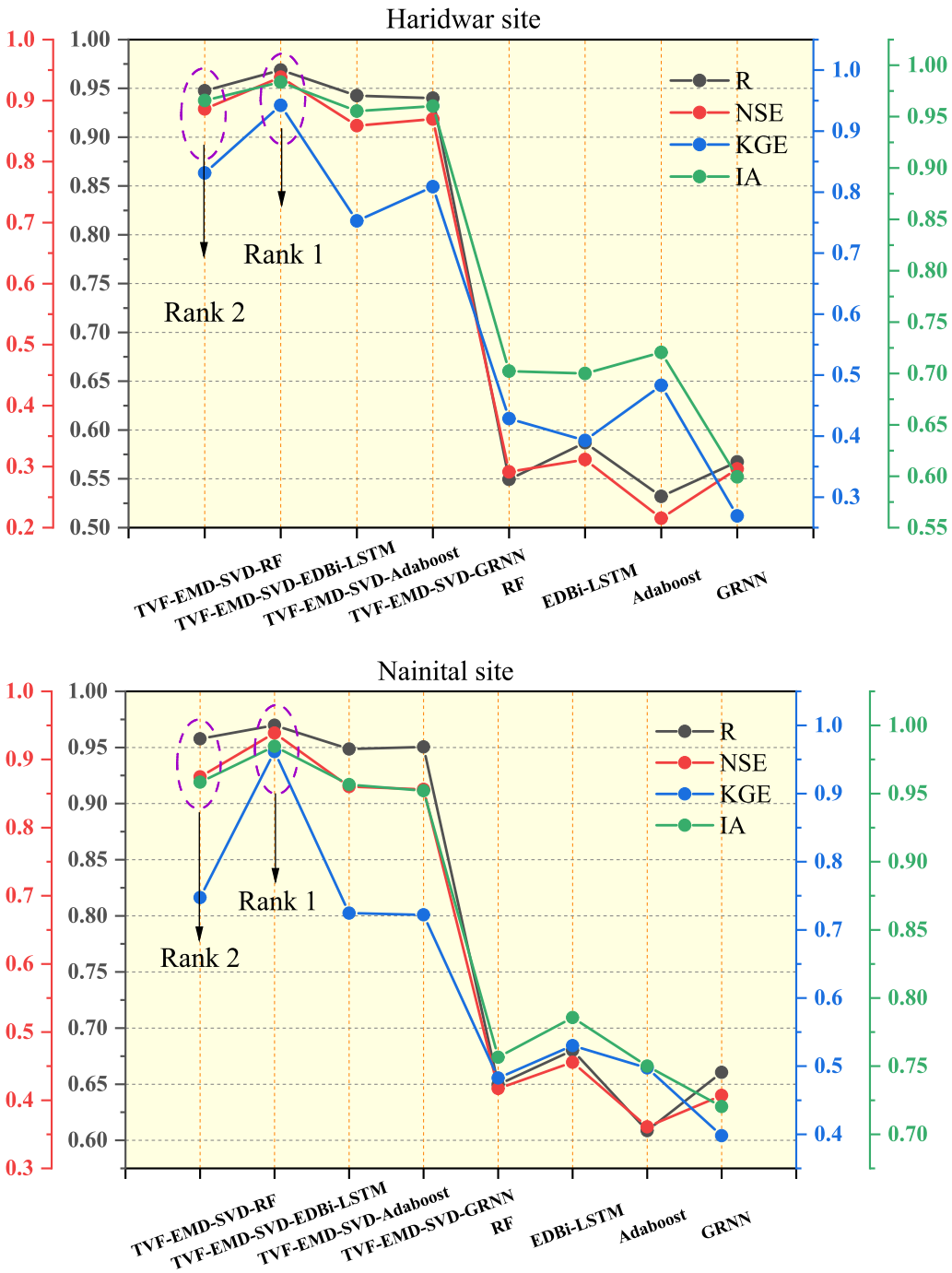


Fig. 8. Assessing the forecasting performance of all the provided models based on the multi-criterion metrics in the testing phase at Haridwar and Nainital sites.

The following are the steps in the RF method (Breiman, 2001):

i. From the original dataset X, different bootstrap samples X_i(i = bootstrap iteration) are randomly selected. Two-thirds of the samples are bootstrapped, while the remaining one-third are out-of-bag samples. Each tree is designed in such a way that it corresponds to a subset of the bootstrap.

ii. At each node in the tree, a new split is chosen at random from all indices, and the regression tree's splitting criterion is the input variable with the lowest mean square error (MSE).

iii. The data splitting procedure is repeated in each internal node until all randomized trees are created, and a stop condition is achieved.

iv. The following are the final regression results, where B represents the total number of trees in the forest and T_b signifies each tree:

$$\hat{y}(x_i) = \frac{1}{B}\sum_{b=1}^{B} T_b(x_i) \tag{16}$$

2.8. Adaptive boosting regression (Adaboost)

AdaBoost, an abbreviation for Adaptive Boosting, is a type of boosting algorithm proposed in 1997 by Freund and Schapire (1997). This approach first establishes an initial distribution on a training set, then repeats the process using adaptive weights to attain a stopping condition. Over the training set, one of the algorithm's key concepts is to preserve a weight distribution. Training samples are weighted based on a classifiers after it has been developed. The reweighted training samples

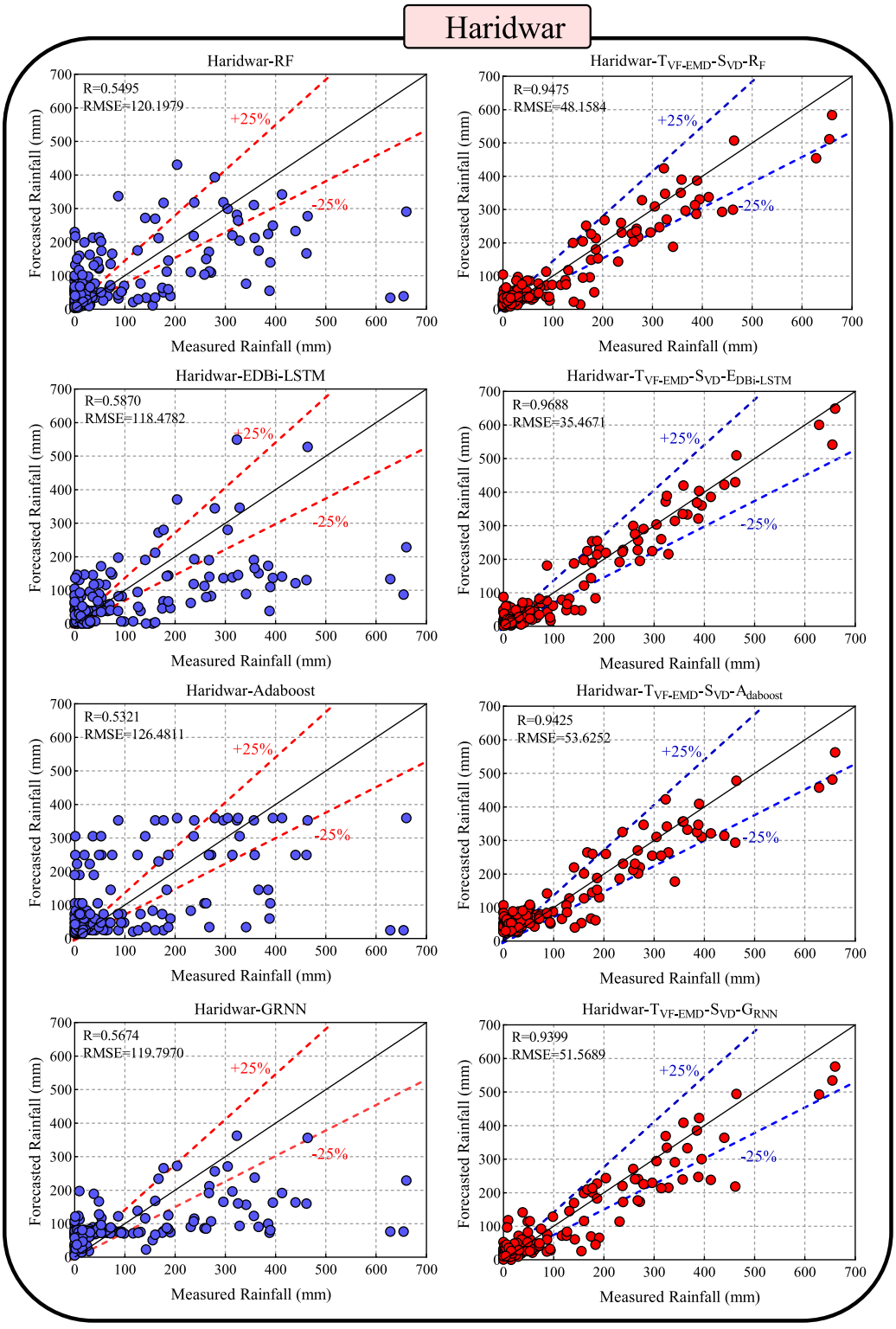


Fig. 9. Scatter plots of measured versus forecasted monthly rainfall values by hybrid and standalone AI models at Haridwar and Nainital stations during the testing phase.

are then used to build the next classifier. To sum it up, when training is complete, the individual classifiers are merged to form a final, highly accurate classifier. A weight is maintained for each individual occurrence, and the more weight an individual instance has, the greater weight that individual instance has on the classifier that is trained (Pedregosa et al., 2011). The AdaBoost approach uses the following algorithm to get its results (El Bilali et al., 2021):

i. Entering training data $D = \{(x_1, y_1), \cdots, (x_m, y_m)\}$, where x_m represents a feature vector, y_m represents a target vector, and m is the quantity of training data.
ii. Use the following equation to initialize the weight vector:

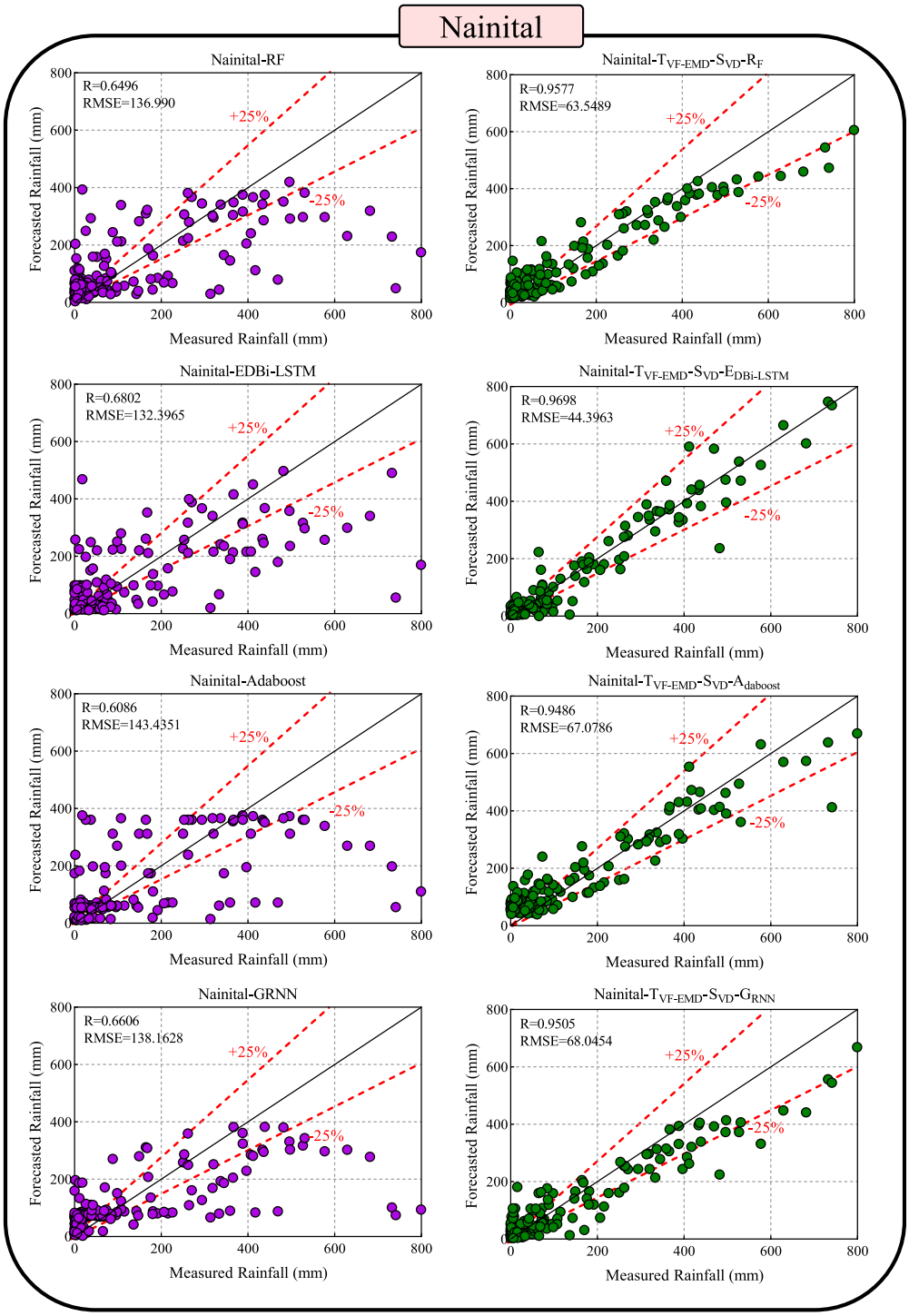


Fig. 9. (*continued*).

$$w_1(x_i) = 1/m \tag{17}$$

iii. T weak learner identification $h_t(t = 1, 2, \cdots, T)$ and initialization at t = 0

A) Determine $P_t(x_i)$

$$P_t(x_i) = \frac{w_t(x_i)}{\sum_{i=1}^{m} w_t(x_i)} \tag{18}$$

B) Determine the error rate:

$$\varepsilon_t = \sum P_t(x_i)[h_t(x_i) - y_i] \tag{19}$$

If $\varepsilon_t > 0.5$ and T = t-1 exit the loop.

C) Weight confidence α_t should be calculated.

$$\alpha_t = log\left(\frac{\varepsilon_t}{1 - \varepsilon_t}\right) \tag{20}$$

D) From i to m, adjust the weighting of all training data.

$$w_{t+1}(x_i) = w_t(x_i) \times e^{-y_i h_t(x_i)\alpha_t} \tag{21}$$

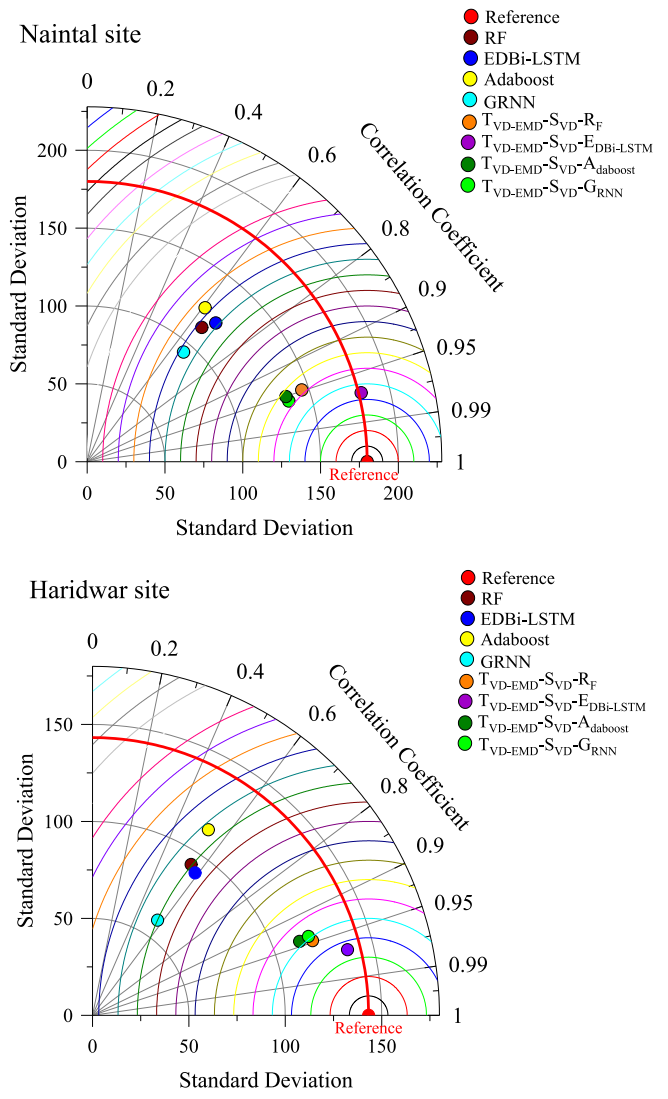


Fig. 10. Taylor diagram for evaluating the robustness of hybrid and standalone counterpart AI approaches at study sites in the testing phase.

E) If $\varepsilon_t > 0.001$ and $t < T$ return to A.

iv. The result is a strong learner created by combining weak learners:

$$H(x) = sign\left[\sum_{j}^{T} \alpha_t [h_t(x) = y]\right] \quad (22)$$

The CART decision tree is employed as a basic learner in this work. Fig. 3 depicts the Adaboost approach algorithm.

2.9. Goodness-of-fit metrics

The monthly rainfall forecasted accuracy of the designed AI models including the RF, EDBi-LSTM, Adaboost, GRNN, T_{VF-EMD}-S_{VD}-R_F, T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$, T_{VF-EMD}-S_{VD}-$A_{daboost}$, and T_{VF-EMD}-S_{VD}-G_{RNN} was evaluated by using six goodness-of-fit metrics i.e., root mean square error (RMSE), mean absolute error (MAE), Nash-Sutcliffe efficiency (NSE), Kling-Gupta efficiency (KGE), coefficient of correlation (R), and agreement index (I_A). The mathematical expression of RMSE (Ebtehaj et al., 2021; Karbasi et al., 2022c), MAE (Malik et al., 2021d), NSE (Nash and Sutcliffe, 1970), KGE (Gupta et al., 2009), R (Malik et al., 2021c; Moriasi et al., 2015), and I_A (Malik et al., 2021b; Willmott, 1981) are written as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Rain_{o,i} - Rain_{f,i}\right)^2} \ (0 < \text{RMSE} < \infty) \ (23).$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|Rain_{f,i} - Rain_{o,i}\right| \ (0 < \text{MAE} < \infty) \ (24).$$

$$NSE = 1 - \left[\frac{\sum_{i=1}^{N}\left(Rain_{o,i} - Rain_{f,i}\right)^2}{\sum_{i=1}^{N}\left(Rain_{o,i} - \overline{Rain_o}\right)^2}\right] (\infty < \text{NSE} < 1) \ (25).$$

$$KGE = 1 - \sqrt{(R-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \ (\infty < \text{KGE} < 1) \ (26).$$

$$R = \frac{\sum_{i=1}^{N}\left(Rain_{o,i} - \overline{Rain_o}\right)\left(Rain_{f,i} - \overline{Rain_f}\right)}{\sqrt{\sum_{i=1}^{N}\left(Rain_{o,i} - \overline{Rain_o}\right)^2 \sum_{i=1}^{N}\left(Rain_{f,i} - \overline{Rain_f}\right)^2}} \ (-1 < \text{R} < 1) \ (27).$$

$$I_A = 1 - \left[\frac{\sum_{i=1}^{N}\left(Rain_{f,i} - Rain_{o,i}\right)^2}{\sum_{i=1}^{N}\left(\left|Rain_{f,i} - \overline{Rain_o}\right| + \left|Rain_{o,i} - \overline{Rain_o}\right|\right)^2}\right] (0 < \text{I}_\text{A} \le 1) \ (28).$$

In Eqs. (23–28), $Rain_{o,i}$, and $Rain_{f,i}$ = observed and forecasted monthly rainfall values for i^{th} observations, $\overline{Rain_o}$, and $\overline{Rain_f}$ = mean of observed and forecasted monthly rainfall values, α = relative variability in the forecasted and observed monthly rainfall values, β = ratio between the mean forecasted and mean observed monthly rainfall values, and N = total number of observations.

3. Model development and strategies configuration

Deep learning methods have been developed with the aim of solving the engineering problems related to the forecasting of time series, which, according to their potential, can lead to more accurate and better results than other classical methods of ML (Karbasi et al., 2022b, 2022a). In simulating highly non-linear problems such as rainfall, even deep learning methods such as the EDBi-LSTM in individual form are not able to achieve proper accuracy. In this research, to overcome those drawbacks, the EDBi-LSTM scheme is integrated with powerful signal decomposition methods and other efficient pre-processing techniques. The TVF-EMD method, by eliminating the shortcomings of the EMD scheme in hybridization with a EDBi-LSTM algorithm, has a significant ability to solve highly non-linear problems. Also, the results can show that applying a dimensionality reduction method such as the SVD method can be effective in reducing the cost of calculations and increasing accuracy. Combining the benefits of each part of the hybrid model presented in this study has made it possible to get promising results to forecast the monthly rainfall in India's different climate zones.

The proposed hybrid AI models (i.e., T_{VF-EMD}-S_{VD}-R_F, T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$, T_{VF-EMD}-S_{VD}-$A_{daboost}$, and T_{VF-EMD}-S_{VD}-G_{RNN}) have been developed on a PC with an Intel Core i7 Intell (R) CPU @ 3.4 GHz and 8 GB of memory. The open-source Python libraries, Keras (Arnold, 2017) and TensorFlow (Abadi, 2016) were used to construct the EDBi-LSTM and Scikit-learn library was employed for the Adaboost development. Also, the RF, GRNN, TVF-EMD, SVD, and PACF approaches were adopted in the MATLAB 2018a environment. The schematic flowchart of the monthly forecasting rainfall using the developed hybrid AI models is illustrated in Fig. 4. The comprehensive details of constructing the hybrid AI models are described in the following steps:

4. Step 1: Significant time-lagged inputs extraction

One of the most critical steps in constructing the forecasting models is to determine the antecedent information and the lagged-time of the considered time-series. There are various techniques for lags determination such as PACF and types of feature selection methods which PACF is taken into account as the most common tool in the literature. Fig. 5 illustrates the graphical representation of the ACF and PACF of both rainfall signals of Haridwar and Nainital stations to determine the significant antecedent information (lags) to apply on decomposed components. As can be seen from Fig. 5, in both stations the rainfall (t-1), rainfall (t-2), and rainfall (t-3) are candidate for applying in models' development. Although the lagged times of 6, 7, 8, and 9 could be influential in the modeling, by furthermore assessments, the authors ascertained that their effect is negative on the accuracy of the models on both stations. So, it is ignored and confined to the first three lags.

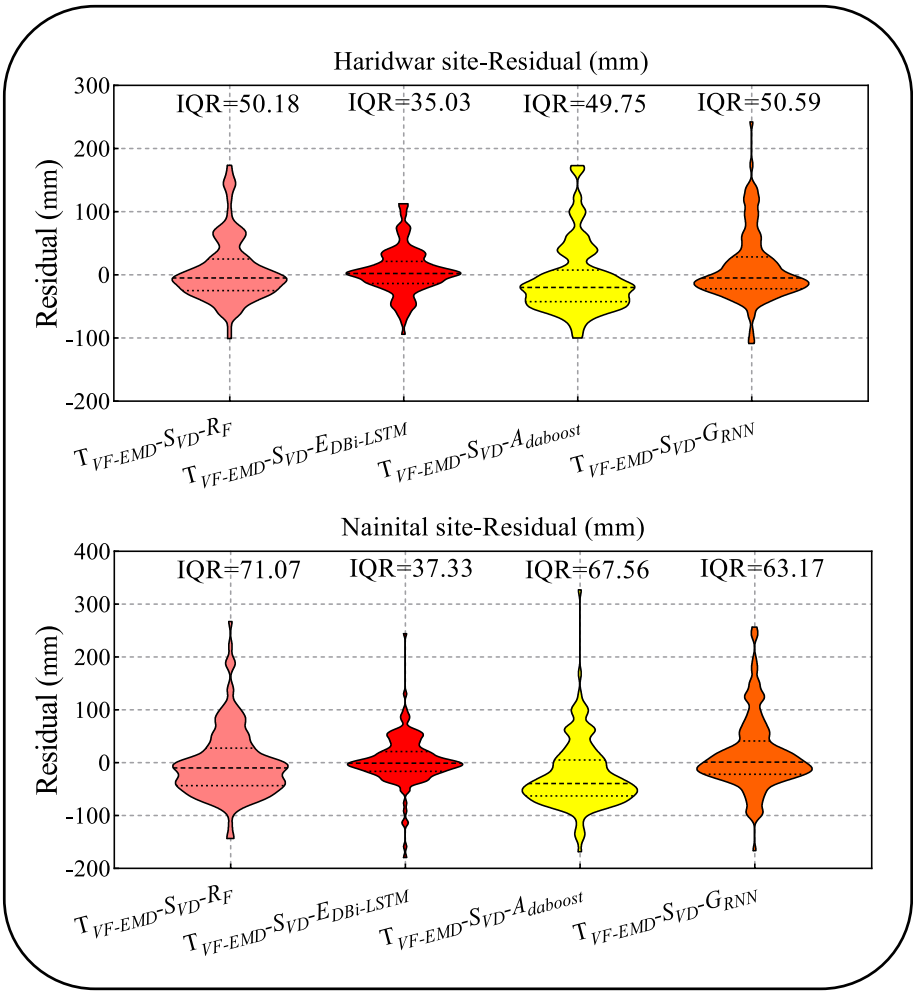


Fig. 11. Violin plots showing the residual generated between the forecasted and measured rainfall values by the hybrid AI models in the testing phase.

5. Step 2: First phase of signals decomposition (TVF-EMD)

The main aim of using the pre-processing TVF-EMD technique was its ability on mitigate mode mixing of the nonstationary rainfall signals and seek the amplitude modulation of the noisy modes which are yielded by the random distribution of extrema (Song, 2021; Song et al., 2021a). The TVF-EMD instead of the classical EMD scheme can handle the impact of noisy signals during the envelope construction step of decomposition techniques. For this purpose, various intrinsic mode functions (IMFs) were examined by trial-and-error procedure, and 5 sub-sequences (4 IMFs and one residual) were picked as the optimal IMFs.

Fig. 6 demonstrates the decomposed sub-sequences in each site of the study. As the default values, the B-spline order and stopping criterion were considered equal to 26 and 0.1, respectively for both sites. In this regard, for each site, 3 lags (antecedent candidate input) extracted from the PACF were typically decomposed and 15 sub-sequences (5 IMFs multiplied by 3 lags) were obtained from the first preprocessing stage. The other setting parameters of the TVF-EMD pre-processing technique such as B-spline order, End_flag, and stopping criterion are specified in Table 3.

6. Step 3: Second phase of decomposition (SVD)

The most important goal of the second decomposition stage is dimensionality reduction and increases the accuracy of rainfall forecasting. Here, the sub-sequences (i.e., IMFs and the residual) created by the TVF-EMD decomposition technique are used to further decompose by the SVD model considering a predefined 40 % dimensionality reduction. Fig. 7 describes the reduction of the candidate input matrix by using the SVD decomposition scheme with (number of singular vectors = 9). The predefined adjustment parameters of the SVD model such as the number of singular vectors, number of iterations, block size, and setting mean center criterion are given in Table 3. Besides, the numbers of training and testing datasets after imposing the lags are mentioned in Table 3.

7. Step 4: Models feeding and configuration

As mentioned before, the duration of modeling was from 1961 to 2015. The most common data-allocating approach for the training and testing datasets in time-series based engineering problems is ordinary direct dividing datasets into sub-sets (Prasad et al., 2020; Quilty and Adamowski, 2018). For this aim, 75 % of whole dataset (41 years) was used for training the models, and the rest 25 % dataset (14 years) was employed for testing of the models. Besides, prior splitting the datasets, all the data sets for the hybrid and standalone AI models were normalized between 0 and 1 to enhance the stability and convergence as the following formulation:

$$r'F = \frac{rF - rF_{\min}}{rF_{\max} - rF_{\min}} \tag{29}$$

where, $r'F$ is the normalized rainfall value and rF is the real-time rainfall value. Moreover, $rF_{\min}$ and $rF_{\max}$ are the minimum and maximum values of rainfall, respectively.

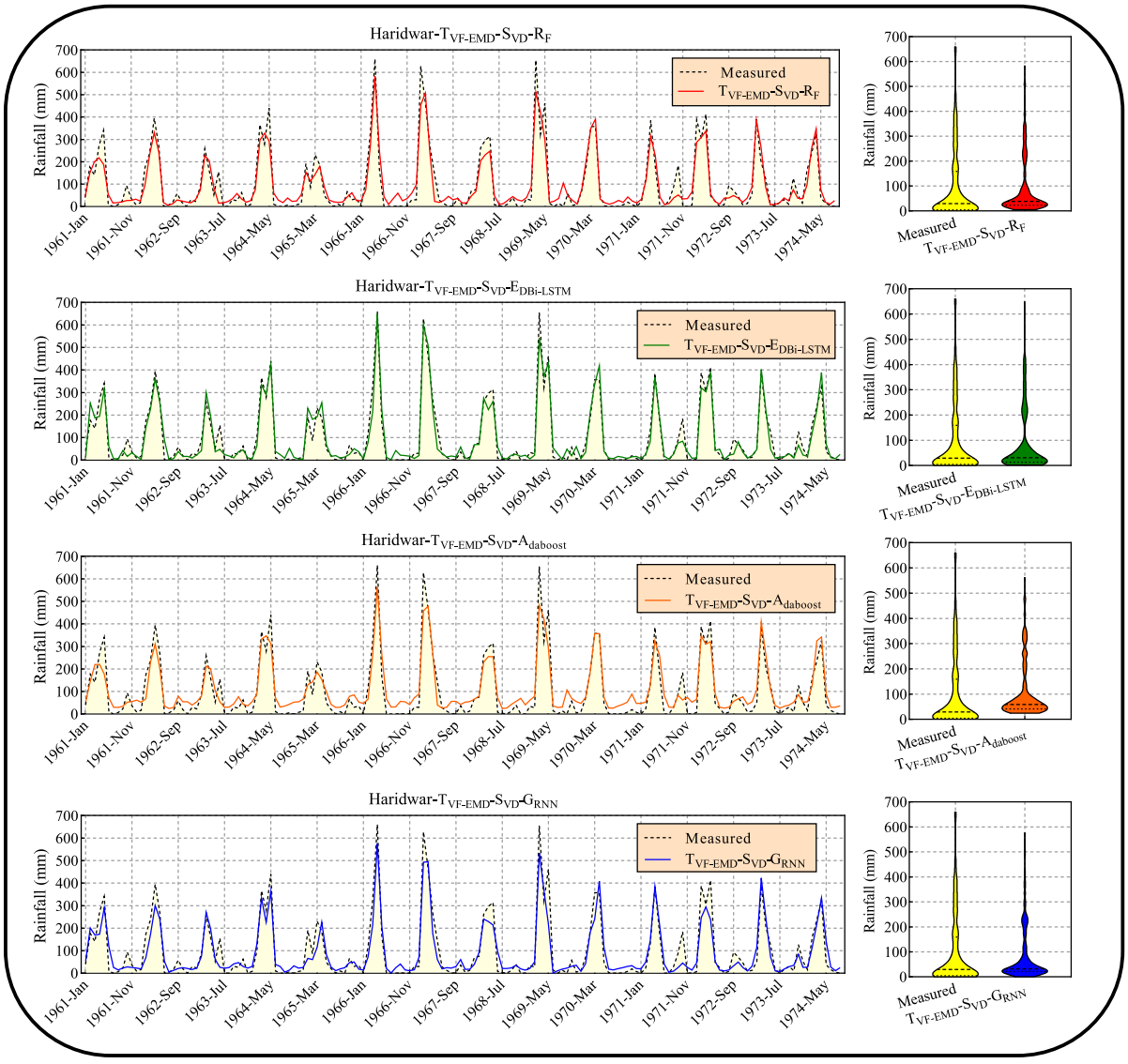


Fig. 12. Comparing the forecasted versus measured rainfall values in form of the trend plot and distribution density function-violin plot of hybrid AI models at Haridwar and Nainital for the testing phase.

7.1. Setting of machine learning approaches

Basically, the choice of regulatory parameters for the AI model has a significant impact on improving the forecasting results. To achieve this goal, *meta*-heuristic algorithms, random search, and grid search strategies are often used. In this research, the grid search method was used to optimize the parameters of the models. Table 4 lists all the optimized hyper-parameters of utilized AI methods in rainfall simulation. The important hyper-parameters ofthe EDBi-LSTM model are the number of layers, neuron number, learning rate, epoch value, training algorithm, and activation operator (Zhang et al., 2021b) whereas the significant setting parameters of the Adaboost approach (Jamei et al., 2021) are learning rate, n_estimators, and loss function. It is noteworthy that the RF and GRNN hyper-parameters are (tree number & maximum number of features) and spread value, respectively.

8. Application results and analysis

In this study, the proposed hybrid (i.e., $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$, $T_{VF\text{-}EMD}$-S_{VD}-R_F, $T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$, $T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}), and the standalone (i.e., LSTM, RF, Adaboost and GRNN) AI models were assessed based on R, RMSE, MAE, NSE, KGE and I_A indicators to forecast monthly rainfall at Haridwar and Nainital stations.

Table 5 highlighted the performance of the proposed hybrid $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model against the other benchmark comparison models on Haridwar station during training and testing periods. Based on R, RMSE, MAE, NSE, KGE, and I_A values, the hybrid $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model achieved the highest values of R = 0.9688, NSE = 0.9383, KGE = 0.9422, I_A = 0.9836, and lowest error magnitudes for RMSE = 35.4671 mm, MAE = 25.1577 mm, followed by the hybrid $T_{VF\text{-}EMD}$-S_{VD}-R_F, $T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$ and $T_{VF\text{-}EMD}$-S_{VD}-G_{RNN} models in the testing period. On the hand, the accuracy of standalone models was relatively poor in relation to the hybrid models, also it was found that the LSTM model in the testing period with R = 0.5870, RMSE = 118.4782 mm, MAE = 72.4846 mm, NSE = 0.3116, KGE = 0.3927, and I_A = 0.7000 appears to be better than the RF, Adaboost, and GRNN models. Overall, the hybrid $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model outperformed in both the training and testing scenarios against other comparing AI models to forecast the monthly rainfall at Haridwar.

Table 6 showing the accuracy of newly developed AI models (i.e.,

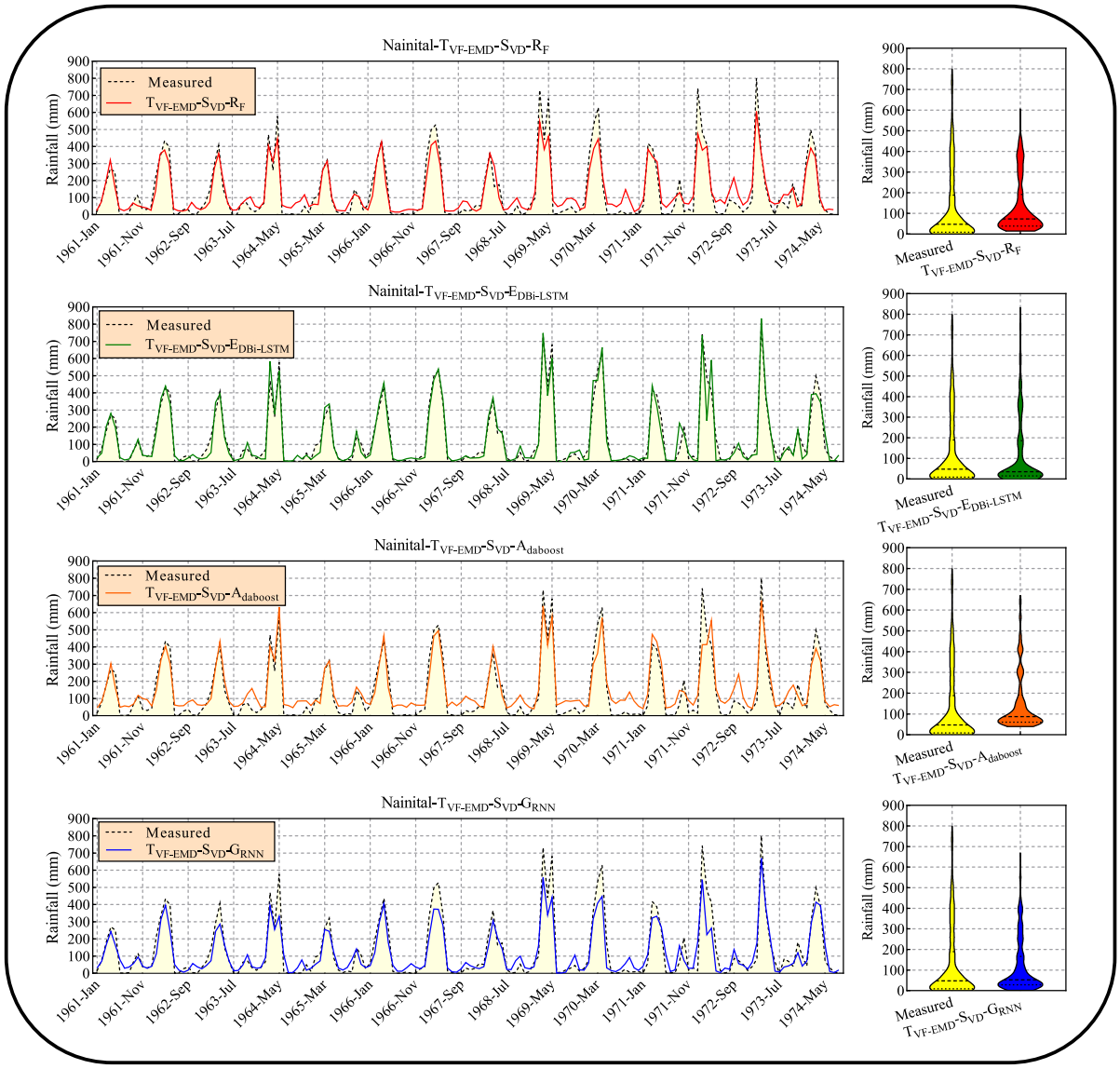


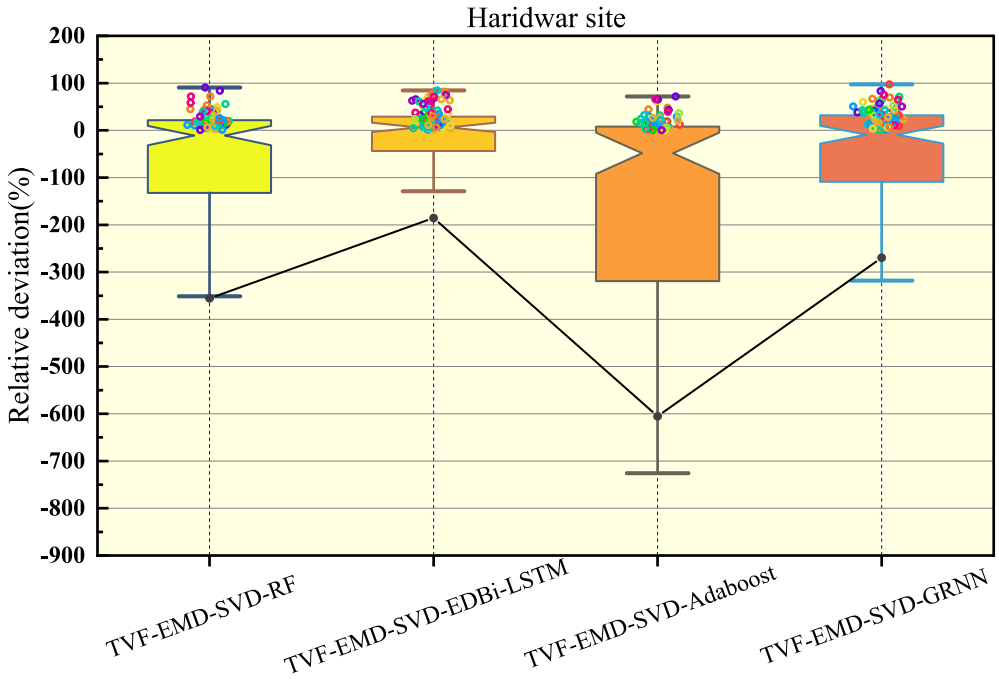
Fig. 12. (*continued*).

hybrid version and standalone) at Nainital station to forecast monthly rainfall. The hybrid T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$ model again appears to be the most accurate to forecast monthly rainfall based on the adopted assessment metrics. For example, these metrics in the testing period are R = 0.9698, RMSE = 44.3963 mm, MAE = 28.5545 mm, NSE = 0.9388, KGE = 0.9614, I_A = 0.9846 best for T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$ model as compared to T_{VF-EMD}-S_{VD}-R_F, T_{VF-EMD}-S_{VD}-$A_{daboost}$, T_{VF-EMD}-S_{VD}-G_{RNN}, LSTM, RF, Adaboost and GRNN models to forecast monthly rainfall on Nainital.

Fig. 8 ranks the proposed hybrid T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$, T_{VF-EMD}-S_{VD}-R_F, T_{VF-EMD}-S_{VD}-$A_{daboost}$, T_{VF-EMD}-S_{VD}-G_{RNN}, and the standalone LSTM, RF, Adaboost and GRNN models based on R, NSE, KGE, and I_A. The analysis undoubtedly proved that the hybrid models are ranked higher than the standalone counterpart models by achieving higher values of the R, NSE, KGE, and I_A for the Haridwar and Nainital sites. But the T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$ ranked at the top (ranked 1) exhibited higher magnitudes of the R, NSE, KGE, and I_A metrics in both the candidate sites (Haridwar and Nainital). The T_{VF-EMD}-S_{VD}-R_F model ranked on 2nd position followed by T_{VF-EMD}-S_{VD}-$A_{daboost}$, T_{VF-EMD}-S_{VD}-G_{RNN}, RF, EDBi-LSTM, Adaboost, and GRNN models for both locations. The T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$ model showpromising accuracy based on the R, NSE, KGE and I_A values for both sites against the other comparing AI models (Fig. 8).

Fig. 9 exposes the relationship between the monthly forecasted and measured rainfall (mm) of the T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$, T_{VF-EMD}-S_{VD}-R_F, T_{VF-EMD}-S_{VD}-$A_{daboost}$, T_{VF-EMD}-S_{VD}-G_{RNN}, LSTM, RF, Adaboost and GRNN models in terms of scatter plots along with R and RMSE for Haridwar and Nainital sites to assess the comparison of the models. For Haridwar site, the T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$ model had R = 0.9688, & RMSE = 35.4671 mm while T_{VF-EMD}-S_{VD}-R_F, T_{VF-EMD}-S_{VD}-$A_{daboost}$, and T_{VF-EMD}-S_{VD}-G_{RNN} models had R = 0.9475, 0.9425, 0.9399, and RMSE = 48.1584 mm, 53.6252 mm, & 51.5689 mm. This comparison indicates the better performance of the T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$ model than the other hybrid models in forecasting the monthly rainfall. In addition, he standalone LSTM, RF, Adaboost, and GRNN models perform poorly to forecast monthly rainfall at Haridwar. Similarly, the T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$ model was excellent in forecasting monthly rainfall at Nainital site followed by T_{VF-EMD}-S_{VD}-R_F, T_{VF-EMD}-S_{VD}-G_{RNN}, and T_{VF-EMD}-S_{VD}-$A_{daboost}$ while again the standalone counterpart models were found to be very poor. Thus, Fig. 9 established that the T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$ model was better in forecasting monthly rainfall for both sites.

Fig. 10 shows the Taylor diagram of referenced/measured and forecasted rainfall for Haridwar and Nainital sites using the T_{VF-EMD}-S_{VD}-$E_{DBi-LSTM}$, T_{VF-EMD}-S_{VD}-R_F, T_{VF-EMD}-S_{VD}-$A_{daboost}$, T_{VF-EMD}-S_{VD}-G_{RNN},



Metric	$T_{VF\text{-}EMD}$-S_{VD}-R_F	$T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$	$T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$	$T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}
Q25%	-279.6	-45.5	-454.8	-127.9
IQR	293.06	80.35	458.411	156.11
Q75%	13.46	34.85	3.611	28.21
Mean	-830.8	-295.3	-1411	-628.5

Fig. 13. Relative deviation percentage distribution attained by the hybrid AI models at Haridwar and Nainital stations during the testing phase.

LSTM, RF, Adaboost and GRNN models. Taylor diagram (Taylor, 2001) is basically based on standard deviation, RMSE, and coefficient of correlation in a polar coordinate system to portrays how far lies the forecasting models from the referenced/measured rainfall. A model with a minimum value of standard deviation, RMSE, and maximum value of R is believed to be the best. An alternative approach to determine the model precision is if the forecasted rainfall is closer to the measured rainfall is indicated as the best model. Fig. 10 clearly shows that the $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model was close to the measured/referenced rainfall for both sites, followed by $T_{VF\text{-}EMD}$-S_{VD}-R_F, $T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$, and $T_{VF\text{-}EMD}$-S_{VD}-G_{RNN} models. The standalone model's i.e., the LSTM, RF, Adaboost, and GRRN lies far away from the measured rainfall depicting poor accuracy.

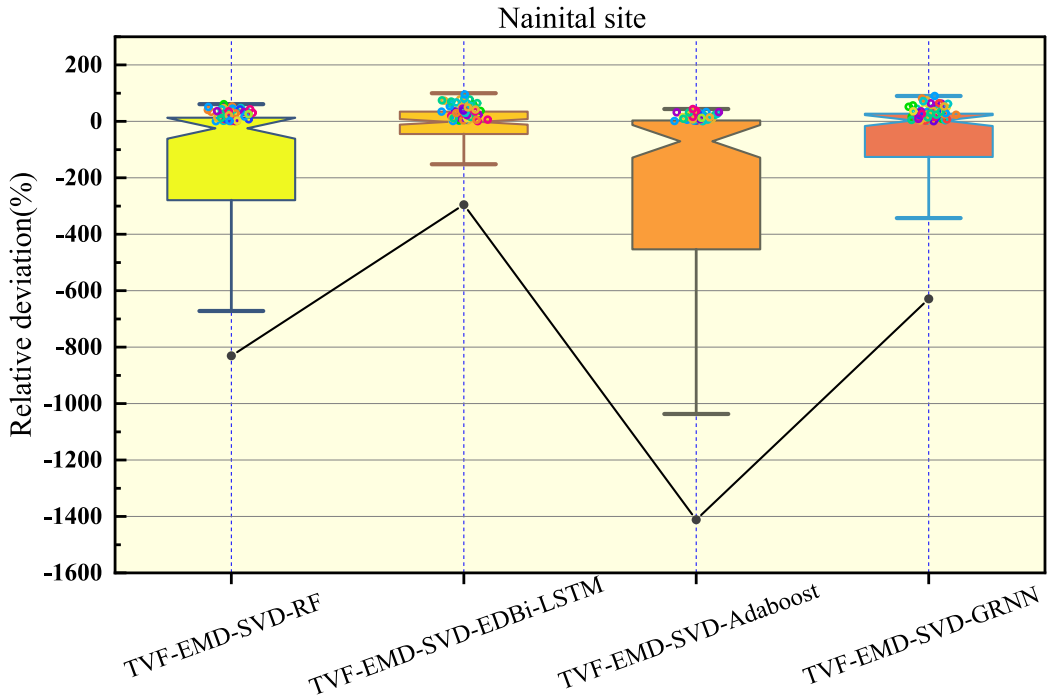
Fig. 11 reveals the residuals between the forecasted and measured rainfall in violin plots of the $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$, $T_{VF\text{-}EMD}$-S_{VD}-R_F, $T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$, and $T_{VF\text{-}EMD}$-S_{VD}-G_{RNN} models at Haridwar and Nainital sites. The $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model generates the lowest residual in violin plot distributions confirming that the errors were very small between the forecasted and measured rainfall with IQR = 35.03 for Haridwar and IQR = 37.33 for Nainital, respectively.

Fig. 12 demonstrates the comparison between the measured rainfall against the forecasted rainfall generated by the hybrid version of the $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$, $T_{VF\text{-}EMD}$-S_{VD}-R_F, $T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}, and $T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$ (Left) models along with their violin plots distribution (Right). By analyzing, the $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model seemed to be the most precise in forecasting the rainfall which exactly follows the pattern and stability corresponding to the measured rainfall with their same violin distribution generation of the monthly rainfall for Haridwar and Nainital sites. This can be easily seen from Fig. 12 where the comparing benchmark models i.e., the $T_{VF\text{-}EMD}$-S_{VD}-R_F, $T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}, and $T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$ yield lower accuracy in terms of time-series plots for forecasted versus measured rainfall with distinct violin distribution sizes which confirms respectively large fluctuations and spikes. Overall, the $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model achieves better accuracy against comparing AI hybrid models for Haridwar and Nainital sites.

A clear distinction of comparison of the hybrid $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model was established using the boxplots (Fig. 13) in combination with the relative deviation metrics (i.e., Q25%, inter quartile range: IQR, Q75%, and mean) demonstrates that the distribution of forecasted rainfall generated from benchmarking hybrid AI models (i.e., $T_{VF\text{-}EMD}$-S_{VD}-R_F, $T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}, and $T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$) for Haridwar and Nainital sites were fairly scattered registering some outliers. Yet, the box-plot distribution of the hybrid $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model has a very accurate interpretation for both stations based on Q25%, IQR, Q75%, and mean values in Fig. 13. Thus, the boxplots for both sites ascertain the excellent forecasting accuracy of the hybridized $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model.

The empirical cumulative distribution function (ECDF) of the forecasted monthly rainfall by the $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ versus $T_{VF\text{-}EMD}$-S_{VD}-R_F, $T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}, and $T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$ models are plotted in Fig. 14.. For Haridwar and Nainital, the ECDF of the $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model displayed a very close profile against the benchmark hybrid comparing AI models which further establishes better precision of the $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model in forecasting monthly rainfall.

The scope of this work can be widen using the satellite-derived input predictors to improve the forecasting capability of the $T_{VF\text{-}EMD}$-S_{VD}-E_{DBi}-



Metric	$T_{VF\text{-}EMD}$-S_{VD}-R_F	$T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$	$T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$	$T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}
Q25%	-135.7	-44.28	-320.6	-114
IQR	157.49	73.68	328.904	145.91
Q75%	21.79	29.4	8.304	31.91
Mean	-355.5	-185.4	-605.2	-269.7

Fig. 13. (*continued*).

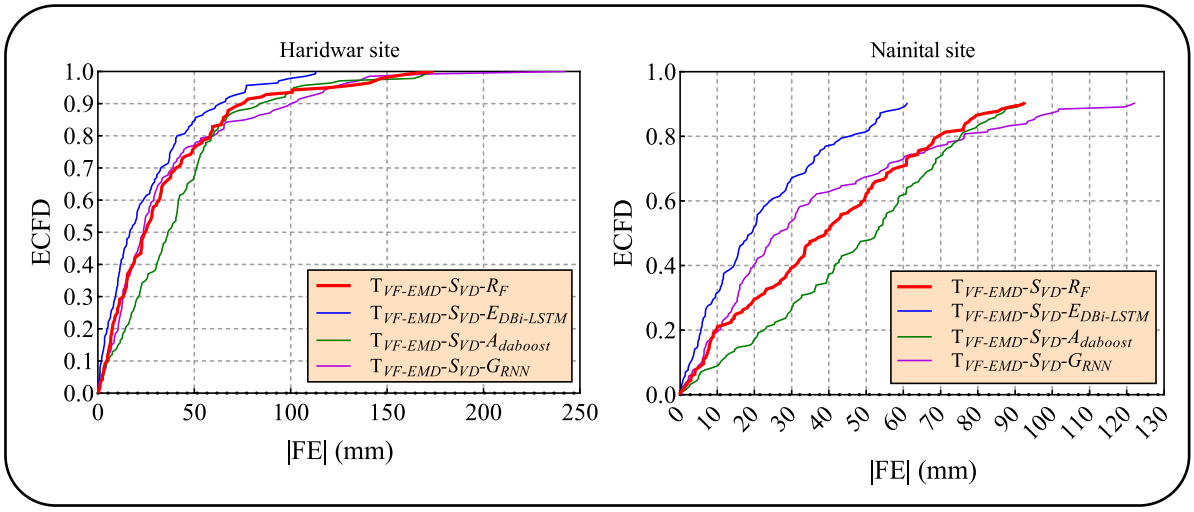


Fig. 14. Empirical cumulative distribution function (ECDF) of the forecasting residual (FE) generated by the hybrid AI models on Haridwar and Nainital in the testing stage.

LSTM model. In addition, the synoptic scale climate mode indices are extremely influencing rainfall patterns, and the employment of these indices can suggest a more detail study in parallel with the integration with federated learning (Yaseen, 2022). The hybridization of deep learning with numerical weather prediction (NWP) models can offer a new emerging area in this field of the research. Further, the scope can be enhanced by implementing the ensemble strategies with uncertainty assessment. The multi-decomposition $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ modelling framework can potentially be employed in other sectors of interest such as air quality monitoring systems, agriculture crops, droughts, hydrology, and renewable and sustainable energy areas under climate change scenarios. In a nutshell, it is established that the multi-decomposition

$T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ modelling framework can be helpful for water resources management, extreme rainfall events, droughts, and hydrology.

9. Conclusion and remarks

This research assesses how to provide a robust and efficient model to accurately forecast the monthly rainfall with the nonstationary signal using a double-decomposition scheme and advanced machine learning models. To achieve this goal, a novel TVF-EMD-based multi-decomposition modelling framework was designed to forecast monthly rainfall. The proposed multi-decomposition modelling framework was based on the TVF-EMD which is integrated with SVD and encoder-decoder-bidirectional-LSTM (i.e., $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$) for monthly rainfall forecasting on Haridwar and Nainital sites, India. Furthermore, for comparison purposes, three advanced data-driven models including the RF, Adaboost, and GRRN were coupled to construct the hybrid AI models i.e., $T_{VF\text{-}EMD}$-S_{VD}-R_F, $T_{VF\text{-}EMD}$-S_{VD}-G_{RNN}, and $T_{VF\text{-}EMD}$-S_{VD}-$A_{daboost}$.

The original signals of monthly rainfall were decomposed into several IMFs using the TVF-EMD technique to overcome the non-stationarity and non-linearity issues. Further, the SVD algorithm was introduced to reduce the dimensionality of IMFs signals which can convert the signals (IMFs) into a more stable and linear form. Finally, the SVD-based signals were then supplied as inputs into the encoder-decoder-bidirectional-LSTM to design the $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model for monthly rainfall forecasting. Likewise, the RF, Adaboost, and GRRN models were also assessed to compare each model in hybrid and standalone versions. The robustness of the models was examined based on goodness-of-fit metrics (i.e., R, RMSE, MAE, NSE, KGE, and I_A) and various graphical interpretation tools. The $T_{VF\text{-}EMD}$-S_{VD}-$E_{DBi\text{-}LSTM}$ model was appeared to be the most accurate by achieving the most precise values in terms of R = 0.5870, RMSE = 118.4782 mm, MAE = 72.4846 mm, NSE = 0.3116, KGE = 0.3927, and I_A = 0.7000 for Haridwar site. Similarly, these metrics are accurate for Nainital site with R = 0.9698, RMSE = 44.3963 mm, MAE = 28.5545 mm, NSE = 0.9388, KGE = 0.9614, and I_A = 0.9846 to forecast monthly rainfall.

Ethical Approval

The manuscript is conducted within the ethical manner advised by the journal.

Consent to Participate

Not applicable.

Consent to Publish

The research is scientifically consented to be published.

Declaration of Competing Interest

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

Data availability

The authors do not have permission to share data.

References

Abadi, M., 2016. TensorFlow: learning functions at scale. Proc. 21st ACM SIGPLAN Int. Conf. Funct. Program. https://doi.org/10.1145/2951913.2976746.

Adaryani, F.R., Mousavi, S.J., Jafari, F., 2022. Short-term rainfall forecasting using machine learning-based approaches of PSO-SVR. LSTM and CNN. J. Hydrol. 614, 128463.

Ahmed, Z., Le, H.P., Shahzad, S.J.H., 2021. Toward environmental sustainability: how do urbanization, economic growth, and industrialization affect biocapacity in Brazil? Environ. Dev. Sustain. 1–21.

Alamgir, M., Khan, N., Shahid, S., Yaseen, Z.M., Dewan, A., Hassan, Q., Rasheed, B., 2020. Evaluating severity–area–frequency (SAF) of seasonal droughts in Bangladesh under climate change scenarios. Stoch. Environ. Res. Risk Assess. https://doi.org/10.1007/s00477-020-01768-2.

Ali, M., Prasad, R., Xiang, Y., Yaseen, Z.M., 2020. Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2020.124647.

Arnold, T.B., 2017. kerasR: R Interface to the Keras Deep Learning Library. J. Open Source Softw. 2, 296.

Ashwini, U., Kalaivani, K., Ulagapriya, K., Saritha, A., 2021. Time Series Analysis based Tamilnadu Monsoon Rainfall Prediction using Seasonal ARIMA. In: In: 6th International Conference on Inventive Computation Technologies, ICICT 2021. Institute of Electrical and Electronics Engineers Inc., Vels Institute of Science, Technology and Advanced Studies, Department of Computer Science and EngineeringChennai, India, pp. 1293–1297. https://doi.org/10.1109/ICICT50816.2021.9358615.

Basha, C.Z., Bhavana, N., Bhavya, P., Sowmya, V., 2020. Rainfall prediction using machine learning & deep learning techniques, in: 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC). IEEE, pp. 92–97.

Bian, C., He, H., Yang, S., Huang, T., 2020. State-of-charge sequence estimation of lithium-ion battery based on bidirectional long short-term memory encoder-decoder architecture. J. Power Sources 449, 227558. https://doi.org/10.1016/j.jpowsour.2019.227558.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

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

Bretherton, C.S., Smith, C., Wallace, J.M., 1992. An Intercomparison of Methods for Finding Coupled Patterns in Climate Data. J. Clim. 5, 541–560. https://doi.org/10.1175/1520-0442(1992)005<0541:AIOMFF>2.0.CO;2.

Cai, Q.-C., Hsu, T.-H., Lin, J.-Y., 2021. Using the General Regression Neural Network Method to Calibrate the Parameters of a Sub-Catchment. Water 13, 1089. https://doi.org/10.3390/w13081089.

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 1724–1734. https://doi.org/10.3115/v1/D14-1179.

Cong, F., Chen, J., Dong, G., Zhao, F., 2013. Short-time matrix series based singular value decomposition for rolling bearing fault diagnosis. Mech. Syst. Signal Process. 34, 218–230. https://doi.org/10.1016/j.ymssp.2012.06.005.

Cramer, S., Kampouridis, M., Freitas, A.A., Alexandridis, A.K., 2017. An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives. Expert Syst. Appl. 85, 169–181.

Danandeh Mehr, A., Nourani, V., Kahya, E., Hrnjica, B., Sattar, A.M.A., Yaseen, Z.M., 2018. Genetic programming in water resources engineering: A state-of-the-art review. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2018.09.043.

Dash, Y., Mishra, S.K., Panigrahi, B.K., 2018. Rainfall prediction for the Kerala state of India using artificial intelligence approaches. Comput. Electr. Eng. 70, 66–73.

Diez-Sierra, J., del Jesus, M., 2020. Long-term rainfall prediction using atmospheric synoptic patterns in semi-arid climates with statistical and machine learning methods. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2020.124789.

Ding, L., Rangaraju, P., Poursaee, A., 2019. Application of generalized regression neural network method for corrosion modeling of steel embedded in soil. Soils Found. 59, 474–483. https://doi.org/10.1016/j.sandf.2018.12.016.

L. Diop S. Samadianfard A. Bodian Z.M. Yaseen M.A. Ghorbani H. Salimi Annual Rainfall Forecasting Using Hybrid Artificial Intelligence Model: Integration of Multilayer Perceptron with Whale Optimization Algorithm. Water Resour. Manag 2020 10.1007/s11269-019-02473-8.

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

El Bilali, A., Taleb, A., Brouziyne, Y., 2021. Groundwater quality forecasting using machine learning algorithms for irrigation purposes. Agric. Water Manag. 245, 106625.

Fahimi, F., Yaseen, Z.M., El-shafie, A., 2017. Application of soft computing based hybrid models in hydrological variables modeling: a comprehensive review. Theor. Appl. Climatol. 128, 875–903. https://doi.org/10.1007/s00704-016-1735-8.

Fan, C., Wang, J., Gang, W., Li, S., 2019. Assessment of deep recurrent neural network-based strategies for short-term building energy predictions. Appl. Energy 236, 700–710. https://doi.org/10.1016/j.apenergy.2018.12.004.

Feng, D., Fang, K., Shen, C., 2020. Enhancing Streamflow Forecast and Extracting Insights Using Long-Short Term Memory Networks With Data Integration at Continental Scales. Water Resour. Res. https://doi.org/10.1029/2019WR026793.

Ferreira, L.B., da Cunha, F.F., 2020. Multi-step ahead forecasting of daily reference evapotranspiration using deep learning. Comput. Electron. Agric. 178, 105728 https://doi.org/10.1016/j.compag.2020.105728.

Freeze, R.A., 1982. Hydrogeological concepts in stochastic and deterministic rainfall-runoff predictions. Recent Trends Hydrogeol. (NARASIMHAN, TN, Ed.) Geol. Soc. Am. Spec. Pap. 189, 63–79.

Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55, 119–139.

Fu, M., Le, C., Fan, T., Prakapovich, R., Manko, D., Dmytrenko, O., Lande, D., Shahid, S., Yaseen, Z.M., 2021. Integration of complete ensemble empirical mode decomposition with deep long short-term memory model for particulate matter concentration prediction. Environ. Sci. Pollut. Res. 1–12.

Graves, A., Schmidhuber, J., 2005. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Networks 18, 602–610. https://doi.org/10.1016/j.neunet.2005.06.042.
Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377, 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.
Hadi, S.J., Abba, S.I., Sammen, S.S.H., Salih, S.Q., Al-ansari, N., Yaseen, Z.M., 2019. Non-Linear Input Variable Selection Approach Integrated With Non-Tuned Data Intelligence Model for Streamflow Pattern Simulation 1–16.
Halder, B., Haghbin, M., Farooque, A.A., 2021. An Assessment of Urban Expansion Impacts on Land Transformation of Rajpur-Sonarpur Municipality. Knowledge-Based Eng. Sci. 2, 34–53.
Herath, H., Chadalawada, J., Babovic, V., 2020. Hydrologically informed machine learning for rainfall-runoff modelling: towards distributed modelling. Hydrol. Earth Syst. Sci. Discuss. 2020, 1–42.
Hernández, E., Sanchez-Anguix, V., Julian, V., Palanca, J., Duque, N., 2016. Rainfall prediction: A deep learning approach, in: International Conference on Hybrid Artificial Intelligence Systems. Springer, pp. 151–162.
Hochreiter, S., Schmidhuber, J.J., 1997. Long short-term memory. Neural Comput. 9, 1–32. https://doi.org/10.1162/neco.1997.9.8.1735.
Hong, W.C., 2008. Rainfall forecasting by technological machine learning models. Appl. Math. Comput. 200, 41–57. https://doi.org/10.1016/j.amc.2007.10.046.
Hu, A., Zhang, K., 2018. Using Bidirectional Long Short-Term Memory Method for the Height of F2 Peak Forecasting from Ionosonde Measurements in the Australian Region. Remote Sens. 10, 1658. https://doi.org/10.3390/rs10101658.
Jamei, M., Karbasi, M., Olumegbon, I.A., Moshraf-Dehkordi, M., Ahmadianfar, I., Asadi, A., 2021. Specific Heat Capacity of Molten Salt-based Nanofluid in Solar Thermal Applications: A Paradigm of Two Modern Ensemble Machine Learning Methods. J. Mol, Liq, p. 116434.
Karbasi, M., Jamei, M., Ali, M., Abdulla, S., Chu, X., Yaseen, Z.M., 2022a. Developing a novel hybrid Auto Encoder Decoder Bidirectional Gated Recurrent Unit model enhanced with empirical wavelet transform and Boruta-Catboost to forecast significant wave height. J. Clean. Prod. 379, 134820 https://doi.org/10.1016/j.jclepro.2022.134820.
Karbasi, M., Jamei, M., Ali, M., Malik, A., Yaseen, Z.M., 2022b. Forecasting weekly reference evapotranspiration using Auto Encoder Decoder Bidirectional LSTM model hybridized with a Boruta-CatBoost input optimizer. Comput. Electron. Agric. 198, 107121.
Karbasi, M., Karbasi, M., Jamei, M., Malik, A., Azamathulla, H.M., 2022c. Development of a new wavelet-based hybrid model to forecast multi-scalar SPEI drought index (case study: Zanjan city, Iran). Theor. Appl. Climatol. 147, 499–522. https://doi.org/10.1007/s00704-021-03825-4.
Kim, H.I., Kim, B.H., 2020. Flood Hazard Rating Prediction for Urban Areas Using Random Forest and LSTM. KSCE J. Civ. Eng. https://doi.org/10.1007/s12205-020-0951-z.
Kurata, G., Xiang, B., Zhou, B., 2016. Labeled Data Generation with Encoder-Decoder LSTM for Semantic Slot Filling, in: Interspeech 2016. ISCA, ISCA, pp. 725–729. https://doi.org/10.21437/Interspeech.2016-727.
Li, H., Li, Z., Mo, W., 2017. A time varying filter approach for empirical mode decomposition. Signal Processing 138, 146–158. https://doi.org/10.1016/j.sigpro.2017.03.019.
Li, X., Xu, W., Ren, M., Jiang, Y., Fu, G., 2022. Hybrid CNN-LSTM models for river flow prediction. Water Supply. https://doi.org/10.2166/ws.2022.170.
Livieris, I.E., Pintelas, E., Pintelas, P., 2020. A CNN–LSTM model for gold price time-series forecasting. Neural Comput. Appl. 32, 17351–17360. https://doi.org/10.1007/s00521-020-04867-x.
Lu, K., Wang, L., 2011. A novel nonlinear combination model based on support vector machine for rainfall prediction. In: In: 2011 Fourth International Joint Conference on Computational Sciences and Optimization. IEEE, pp. 1343–1346.
Malik, A., Kumar, A., 2020. Spatio-temporal trend analysis of rainfall using parametric and non-parametric tests: case study in Uttarakhand. India. Theor. Appl. Climatol. 140, 183–207. https://doi.org/10.1007/s00704-019-03080-8.
Malik, A., Kumar, A., 2021. Application of standardized precipitation index for monitoring meteorological drought and wet conditions in Garhwal region (Uttarakhand). Arab. J. Geosci. https://doi.org/10.1007/s12517-021-07158-4.
Malik, A., Kumar, A., Salih, S.Q., Kim, S., Kim, N.W., Yaseen, Z.M., Singh, V.P., 2020. Drought index prediction using advanced fuzzy logic model: Regional case study over Kumaon in India. PLoS One. https://doi.org/10.1371/journal.pone.0233280.
Malik, A., Kumar, A., Kisi, O., Khan, N., Salih, S.Q., Yaseen, Z.M., 2021a. Analysis of dry and wet climate characteristics at Uttarakhand (India) using effective drought index. Nat. Hazards 105, 1643–1662.
Malik, A., Kumar, A., Rai, P., Kuriqi, A., 2021b. Prediction of Multi-Scalar Standardized Precipitation Index by Using Artificial Intelligence and Regression Models. Climate 9, 28. https://doi.org/10.3390/cli9020028.
Malik, A., Tikhamarine, Y., Al-Ansari, N., Shahid, S., Sekhon, H.S., Pal, R.K., Rai, P., Pandey, K., Singh, P., Elbeltagi, A., Sammen, S.S., 2021c. Daily pan-evaporation estimation in different agro-climatic zones using novel hybrid support vector regression optimized by Salp swarm algorithm in conjunction with gamma test. Eng. Appl. Comput. Fluid Mech. 15, 1075–1094. https://doi.org/10.1080/19942060.2021.1942990.
Malik, A., Tikhamarine, Y., Souag-Gamane, D., Rai, P., Sammen, S.S., Kisi, O., 2021d. Support vector regression integrated with novel meta-heuristic algorithms for meteorological drought prediction. Meteorol. Atmos. Phys. 133, 891–909. https://doi.org/10.1007/S00703-021-00787-0.
Mardani, A., Liao, H., Nilashi, M., Alrasheedi, M., Cavallaro, F., 2020. A multi-stage method to predict carbon dioxide emissions using dimensionality reduction, clustering, and machine learning techniques. J. Clean. Prod. 275 https://doi.org/10.1016/j.jclepro.2020.122942.
Mohsenipour, M., Shahid, S., Ziarh, G.F., Yaseen, Z.M., 2020. Changes in monsoon rainfall distribution of Bangladesh using quantile regression model. Theor. Appl. Climatol. 1–14.
Moon, S.H., Kim, Y.H., Lee, Y.H., Moon, B.R., 2019. Application of machine learning to an early warning system for very short-term heavy rainfall. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2018.11.060.
Moriasi, D.N., Gitau, M.W., Pai, N., Daggupati, P., 2015. Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria. Trans. ASABE 58, 1763–1785. https://doi.org/10.13031/trans.58.10715.
Nandargi, S., Gaur, A., Mulye, S.S., 2016. Hydrological analysis of extreme rainfall events and severe rainstorms over Uttarakhand. India. Hydrol. Sci. J. 61, 2145–2163. https://doi.org/10.1080/02626667.2015.1085990.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I - A discussion of principles. J. Hydrol. 10, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.
Oleiwi, S., Jalal, S., Hamed, S., Ozgur, S., Zaher, K., Yaseen, M., 2018. Precipitation pattern modeling using cross-station perception: regional investigation. Environ. Earth Sci. https://doi.org/10.1007/s12665-018-7898-0.
Omeje, O.E., Maccido, H.S., Badamasi, Y.A., Abba, S.I., 2021. Performance of Hybrid Neuro-Fuzzy Model for Solar Radiation Simulation at Abuja, Nigeria: A Correlation Based Input Selection Technique. Knowledge-Based Eng. Sci. 2, 54–66.
Park, S.H., Kim, B., Kang, C.M., Chung, C.C., Choi, J.W., 2018. Sequence-to-Sequence Prediction of Vehicle Trajectory via LSTM Encoder-Decoder Architecture, in: 2018 IEEE Intelligent Vehicles Symposium (IV). IEEE, pp. 1672–1678. https://doi.org/10.1109/IVS.2018.8500658.
Patel, J., Parekh, F., 2014. Forecasting rainfall using adaptive neuro-fuzzy inference system (ANFIS). Int. J. Appl. Innov. Eng. Manag 3, 262–269.
Patra, J.P., Mishra, A., Singh, R., Raghuwanshi, N.S., 2012. Detecting rainfall trends in twentieth century (1871–2006) over Orissa State. India. Clim. Change 111, 801–817.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., 2011. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.
Pham, B.T., Le, L.M., Le, T.-T., Bui, K.-T.-T., Le, V.M., Ly, H.-B., Prakash, I., 2020. Development of advanced artificial intelligence models for daily rainfall prediction. Atmos. Res. 237, 104845.
Prasad, R., Ali, M., Xiang, Y., Khan, H., 2020. A double decomposition-based modelling approach to forecast weekly solar radiation. Renew. Energy. https://doi.org/10.1016/j.renene.2020.01.005.
Prasetya, E.P., Djamal, E.C., 2019. Rainfall forecasting for the natural disasters preparation using recurrent neural networks, in: 2019 International Conference on Electrical Engineering and Informatics (ICEEI). IEEE, pp. 52–57.
Prudden, R., Adams, S., Kangin, D., Robinson, N., Ravuri, S., Mohamed, S., Arribas, A., 2020. A review of radar-based nowcasting of precipitation and applicable machine learning techniques. arXiv Prepr. arXiv2005.04988.
Quilty, J., Adamowski, J., 2018. 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. https://doi.org/10.1016/j.jhydrol.2018.05.003.
Ramirez, M.C.V., de Campos Velho, H.F., Ferreira, N.J., 2005. Artificial neural network technique for rainfall forecasting applied to the Sao Paulo region. J. Hydrol. 301, 146–162.
Salehin, I., Talha, I.M., Hasan, M.M., Dip, S.T., Saifuzzaman, M., Moon, N.N., 2020. An artificial intelligence based rainfall prediction using LSTM and neural network, in: 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE). IEEE, pp. 5–8.
Seng, D., Zhang, Q., Zhang, X., Chen, G., Chen, X., 2021. Spatiotemporal prediction of air quality based on LSTM neural network. Alexandria Eng. J. 60, 2021–2032. https://doi.org/10.1016/j.aej.2020.12.009.
Song, C., Chen, X., Ding, X., Zhang, L., 2021a. Sea level simulation with signal decomposition and machine learning. Ocean Eng. 241, 110109 https://doi.org/10.1016/j.oceaneng.2021.110109.
Song, C., Chen, X., Wu, P., Jin, H., 2021b. Combining time varying filtering based empirical mode decomposition and machine learning to predict precipitation from nonlinear series. J. Hydrol. 603, 126914 https://doi.org/10.1016/j.jhydrol.2021.126914.
Song, C., 2021. Performance Comparison of Machine Learning Models for Annual Precipitation Prediction Using Different Decomposition Methods.
Specht, D.F., 1991. A general regression neural network. Neural Networks, IEEE Trans. 2, 568–576. https://doi.org/10.1109/72.97934.
Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106, 7183–7192. https://doi.org/10.1029/2000JD900719.
Teegavarapu, R.S.V., Chandramouli, V., 2005. Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. J. Hydrol. 312, 191–206. https://doi.org/10.1016/j.jhydrol.2005.02.015.
Thamilselvan, R., Prabhu, M., Selvi, K.T., Karthik, V., Bhuvaneshwaran, T., 2022. An Extensive Review of Rainfall Prediction using Machine Learning and Deep Learning Techniques, in: 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS). IEEE, pp. 198–205.
Wang, K., Fu, W., Chen, T., Zhang, B., Xiong, D., Fang, P., 2020. A compound framework for wind speed forecasting based on comprehensive feature selection, quantile regression incorporated into convolutional simplified long short-term memory

network and residual error correction. Energy Convers. Manag. 222, 113234 https://doi.org/10.1016/j.enconman.2020.113234.
Wang, J., Zhang, C., 2018. Software reliability prediction using a deep learning model based on the RNN encoder–decoder. Reliab. Eng. Syst. Saf. 170, 73–82. https://doi.org/10.1016/j.ress.2017.10.019.
Willmott, C.J., 1981. On the validation of models. Phys. Geogr. https://doi.org/10.1080/02723646.1981.10642213.
Xiang, Z., Yan, J., Demir, I., 2020. A Rainfall-Runoff Model With LSTM-Based Sequence-to-Sequence Learning. Water Resour. Res. 56 https://doi.org/10.1029/2019WR025326.
Yaseen, Z.M., 2022. The next generation of soil and water bodies heavy metals prediction and detection: New expert system based Edge Cloud Server and Federated Learning technology. Environ. Pollut. 313, 120081 https://doi.org/10.1016/j.envpol.2022.120081.
Yaseen, Z.M., Ebtehaj, I., Kim, S., Sanikhani, H., Asadi, H., Ghareb, M.I., Bonakdari, H., Wan Mohtar, W.H.M., Al-Ansari, N., Shahid, S., 2019. Novel hybrid data-intelligence model for forecasting monthly rainfall with uncertainty analysis. Water (Switzerland). https://doi.org/10.3390/w11030502.
Yaseen, Z.M., Shahid, S., 2020. Drought Index Prediction Using Data Intelligent Analytic Models: A Review. In: intelligent Data Analytics for Decision-Support Systems in Hazard Mitigation. Springer, pp. 1–27.
Yin, J., Deng, Z., Ines, A.V.M., Wu, J., Rasu, E., 2020. Forecast of short-term daily reference evapotranspiration under limited meteorological variables using a hybrid bi-directional long short-term memory model (Bi-LSTM). Agric. Water Manag. 242, 106386 https://doi.org/10.1016/j.agwat.2020.106386.
Yu, P.-S., Yang, T.-C., Chen, S.-Y., Kuo, C.-M., Tseng, H.-W., 2017. Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. J. Hydrol. 552, 92–104. https://doi.org/10.1016/j.jhydrol.2017.06.020.
Zaw, W.T., Naing, T.T., 2008. Empirical statistical modeling of rainfall prediction over Myanmar. World Acad. Sci. Eng. Technol. 2, 500–504.
Zeleke, K., Raes, D., 1999. Test of homogeneity, frequency analysis of rainfall data and estimate of drought probabilities in Dire Dawa, eastern Ethiopia. Ethiop. J. Nat. Resour. 1, 125.
Zhang, X., Liu, Z., Miao, Q., Wang, L., 2018. An optimized time varying filtering based empirical mode decomposition method with grey wolf optimizer for machinery fault diagnosis. J. Sound Vib. 418, 55–78. https://doi.org/10.1016/j.jsv.2017.12.028.
Zhang, B., Zou, G., Qin, D., Lu, Y., Jin, Y., Wang, H., 2021a. A novel Encoder-Decoder model based on read-first LSTM for air pollutant prediction. Sci. Total Environ. 765, 144507 https://doi.org/10.1016/j.scitotenv.2020.144507.
Zhang, B., Zou, G., Qin, D., Lu, Y., Jin, Y., Wang, H., 2021b. Science of the Total Environment A novel Encoder-Decoder model based on read- fi rst LSTM for air pollutant prediction. Sci. Total Environ. 765, 144507 https://doi.org/10.1016/j.scitotenv.2020.144507.
Zhang, S., Xu, F., Hu, M., Zhang, L., Liu, H., Li, M., 2021c. A novel denoising algorithm based on TVF-EMD and its application in fault classification of rotating machinery. Measurement 179, 109337. https://doi.org/10.1016/j.measurement.2021.109337.