IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 24, 2020, accepted March 6, 2020, date of publication March 16, 2020, date of current version March 25, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2980977

Hybrid Deep Learning Approach for Multi-Step-Ahead Daily Rainfall Prediction Using GCM Simulations

MOHD IMRAN KHAN AND RAJIB MAITY
Department of Civil Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India
Corresponding author: Rajib Maity (rajib@civil.iitkgp.ac.in)

This work was supported in part by the Department of Science and Technology, Climate Change Programme (SPLICE), Government of India, through a Sponsored Project under Grant DST/CCP/CoE/79/2017(G).

ABSTRACT Deep Learning (DL) is an effective technique for dealing with complex systems. This study proposes a hybrid DL approach, a combination of one-dimensional Convolutional Neural Network (Conv1D) and Multi-Layer Perceptron (MLP) (hereinafter referred to as hybrid Conv1D-MLP model), for multi-step-ahead (1-day to 5-day in advance) daily rainfall prediction. Nine meteorological variables, closely associated with daily rainfall variation, are used as inputs to the hybrid model. The causal variables are obtained from a General Circulation Model (GCM). In general, simulation of meteorological variables from GCM is much better than rainfall estimate and observed records of meteorological variables is sparsely available, if not completely unavailable at many locations. Thus, proposed scheme helps to establish the effectiveness of the DL approach in augmenting the quality of rainfall prediction, exploiting the potential of GCM in simulating meteorological variables. The developed hybrid model is applied to twelve different locations in different climatic regimes in terms of daily precipitation characteristics. The proposed hybrid approach is compared with a DL approach namely, Multi-Layered Perceptron (deep MLP) and another machine learning approach namely, Support Vector Regression (SVR). It is also found that the performance of the model gradually decreases as the prediction lead time (in days) increases. Overall, this study establishes the fact that the hybrid Conv1D-MLP model is more effective in capturing the complex relationship between the causal variables and daily variation of rainfall. The benefit is due to the unification of potentials of individual approaches for extracting the hidden features of hydrometeorological association.

INDEX TERMS Deep learning (DL), multi-layer perceptron (MLP), hybrid DL models (Conv1D-MLP), multi-step-ahead rainfall prediction, one-dimensional convolutional neural network (Conv1D), support vector regression (SVR).

I. INTRODUCTION

Precipitation is one of the important components in the hydrologic system. It is also one of the six intrinsic parts of weather prediction. In the last few decades, the spatiotemporal distribution of precipitation is getting modified as an impact of changing climate. This leads to simultaneous drought and flood like situation within a spatial distance of a few hundred kilometers [1]. Considering the Indian mainland, many cases of extreme rainfall events had been recorded in the southern, east-central, northern and north-western parts of the country [2]. For instances, Gujarat and Maharashtra in 2005, Ladakh in 2010, Uttarakhand in 2013, Tamil Nadu and Puducherry in 2015 and Kerala in 2018 received unusually heavy rainfall that led to a huge loss of life and property [3]. Daily rainfall is affected by numerous uncertain factors, including climatic conditions, local meteorological factors, and atmospheric circulations. Quantitative prediction of daily rainfall is a challenging task and holds paramount importance for many operational and scientific applications.

There are several methods available with their own merits and demerits. Many of them require detailed information on the physical processes responsible for the occurrence of rainfall, and also their simulations are computationally challenging [4]–[8]. Recent developments in Artificial Intelligence (AI)/Machine Learning (ML) are proven to be highly

The associate editor coordinating the review of this manuscript and approving it for publication was Shaojun Wang.

potential approaches in understanding many such complex phenomena with a wide range of computational challenges, such as image processing, sequence learning, speech recognition, communication network etc. [5]–[9]. The application of AI/ML in hydrology has a long history since 1990s [10]. In these approaches, inherent physical processes are implicitly considered through AI without any explicit requirement as it is in physical and conceptual models [11], [12]. Such approaches are widely used in hydroclimatic modelling viz. rainfall and drought prediction, extreme climate event identification and water demand forecast. Methods like Artificial Neural Networks (ANNs), Support Vector Regression (SVR), Gene Expression Programming (GEP) etc. are being successfully used recently [12], [13]. For example, Sawale and Gupta [14] proposed an algorithm based on a neural network that forecasts atmospheric conditions. The authors used a hybrid architecture comprising of a Hopfield Network (HN) and a Back Propagation Network (BPN) approach. Charaniya and Dudul [15] proposed two different ANN models for forecasting consecutive rainfall values based on previous day lagged rainfall values. In this study, a pattern recognition approach of the neural network was adopted to extract the relevant spatiotemporal feature of historical rainfall data Lee *et al.* [16] presented an ANN algorithm for forecasting early summer rainfall for Geum River Basin in South Korea. In this analysis, the observed rainfall data was assumed to follow a normal distribution, and it was classified into three categories viz. below, near and above normal. The model gave very poor performance in classifying the above and below the normal category of rainfall. Furthermore, various conjunction models with ANN, such as the wavelet–neuro fuzzy model [17] and data pre-processing coupled ANN [18], have also been used successfully to predict precipitation. Several similar studies have effectively carried out rainfall forecasting and reported the applicability of Machine Learning (ML) algorithms [19]–[27].

In contrast to all the aforementioned computing techniques, Deep Learning (DL) is one of the recently popularized AI approach that is beneficial in many multi-dimensional aspects. It has the ability to utilize raw data and automatically extracts the data features using successive layer representation. These layers contribute information of the model and provide the flexibility to use multiple hierarchical layers and learn from exposure to data without any human expertise [9]. Several DL techniques viz. Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long-Short Term Memory (LSTM) etc., so far has been applied successfully in various fields such as in image recognition [22], [28], speech recognition [29], [30], medical science [31], language understanding [32], rainfall forecasting [33] etc., and it has outperformed the existing AI/ML algorithms. These results suggest that DL may have potential applicability in many different domains, which has motivated researchers to explore and apply it for forecasting hydrologic variables at various spatio-temporal scales. For instance, Liu *et al.* [34] presented a DL based deep MLP approach to process a huge weather dataset which was used to forecast the weather for next 24 hour (Liu *et al.* ([35] presented a DL based CNN approach to identify the extreme weather condition from the climate dataset. It was the first DL based study performed for detecting climate extreme Zhang *et al.* [36] presented a DL based deep belief network algorithm for forecasting next day precipitation using seven environmental factors from the previous day. They found a better accuracy in the forecast as compared with various ML and statistical algorithms. However, there were several days for which the forecast was not reasonably good. Ghaderi *et al.* [37] presented a DL based RNN approach for forecasting the next hour wind speed which outperformed various widely used benchmark ML models. Aswin *et al.* [38] presented DL based architectures, namely LSTM and CNN for prediction of rainfall magnitude. In this study, rainfall dataset for January month (1979-2018), from the Global Precipitation Climatology Project (GPCP) was used. Both the DL architectures were trained and optimized on this Global Average Monthly (GAM) dataset. The proposed architecture predicted the GAM rainfall value. However, both the architectures have a similar root mean squared error (RMSE) indicating a scope of improvement in both the architectures Hu *et al.* [39] presented a DL based LSTM approach for simulating rainfall-runoff process which outperforms the ANN models. In order to forecast the visibility at airports, Salman *et al.* [40] presented a hybrid approach consisting of DL based LSTM and ARIMA model using three atmospheric variables viz. temperature, dew point and humidity. Wang and Li [41] presented a hybrid approach to forecast multi-step-ahead wind speed using optimal feature extraction and DL based LSTM methodology. Haidar and Verma [33] used a DL based one-dimensional deep CNN approach (Conv 1D) to forecast the monthly rainfall at Innisfail, Australia. In this study, eleven climate indices and sunspot values were used as the predictors. The obtained result was compared with MLP and the forecasting model of the Bureau of Meteorology, Australia. The analysis revealed that Conv 1D model performance was better for the months having higher annual mean whereas it was not good for months having lower annual mean of rainfall. Likewise, different DL based researches have been attempted in recent past to forecast different hydroclimatic variables in different parts of the world ([27], [42]–[48] etc.). Therefore, daily rainfall prediction would be another potential field of application. It is established in the existing literature that DL primarily serves two functionalities: a) model building with higher prediction efficiency, greater processing capability, and reduced human intervention, and b) data mining to support discoveries that expand the current state of knowledge and capabilities [49]. The former can be effectively used in the field of hydroclimatology to study the nonlinear, complex, and hidden information of the hydrologic time series and develop prediction models at various spatio-temporal scale. This forms the motivation of this study, i.e. to explore the potential of the DL approaches in hydrometeorological

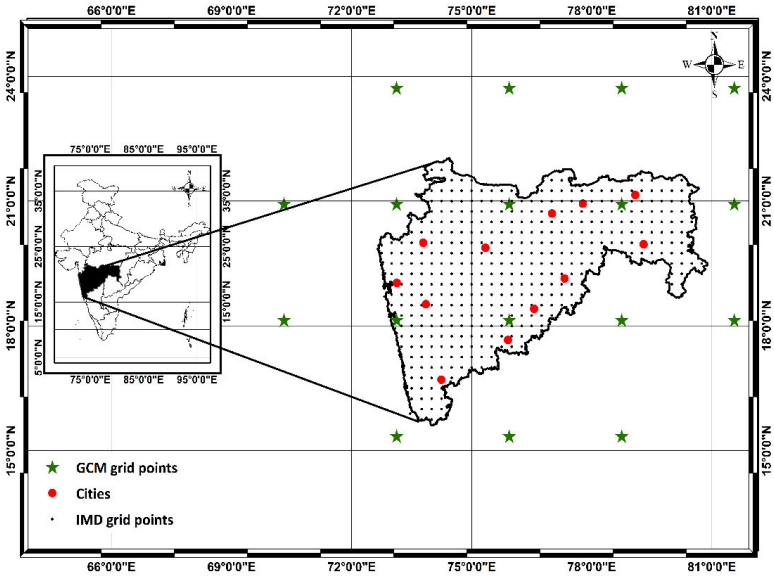

FIGURE 1. Study area map showing location of IMD grid points, GCM grid Points and cities of Maharashtra.

studies. Hydrometeorological prediction of daily rainfall prediction is considered in this study.

Objective of this study is to extract the hidden sequential information between rainfall (target) and associated meteorological (causal) variables in order to develop a multi-step-ahead rainfall prediction model at daily time-scale. Following specific contributions are made in this study.

- This study proposes a hybrid DL approach, a combination of one-dimensional Convolutional Neural Network (Conv1D) and Multi-Layer Perceptron (MLP) (hereinafter referred to as hybrid Conv1D-MLP model), for multi-step-ahead (1-day to 5-day-ahead) daily rainfall prediction.
- Nine meteorological variables, closely associated with daily rainfall variation, are used as inputs to the hybrid model. The causal variables are obtained from a General Circulation Model (GCM). In general, simulation of meteorological variables from GCM is much better than rainfall estimate and observed records of meteorological variables is sparsely available, if not completely unavailable at many locations. Thus, proposed scheme helps to establish the effectiveness of the DL approach in augmenting the quality of rainfall prediction, exploiting the potential of GCM in simulating meteorological variables.
- Next, the proposed hybrid approach is compared with a DL approach namely, Multi-Layered Perceptron (deep MLP) and another machine learning approach namely, Support Vector Regression (SVR).

The rest of the paper is organized as follows: Section II describes the study area, details of data along with their sources. Section III describes the methodology of data preparation and the proposed approach. Criteria for model performance evaluation are also discussed in this section. Section IV presents the results and discussion. First, the performance of the hybrid Conv1D-MLP model in the prediction of rainfall with different lead times (1-day to 5-day) is presented, and next, it is compared with other models (deep MLP and SVR). Conclusions are drawn in Section V.

II. STUDY AREA AND DATA COLLECTION

Twelve cities in the state of Maharashtra, India are considered in this study. These are Akola, Amravati, Aurangabad, Chandpur, Kolhapur, Latur, Nagpur, Nanded, Nashik, Navi Mumbai, Pune and Solapur. Maharashtra spans over a wide range of climatological conditions: the low rainfall regions of the state are under constant risk of droughts, whereas high rainfall zones of eastern and western regions are prone to flash floods and landslides. The state faces deficit in rainfall almost once in 5 years, severe droughts almost once in 8-9 years, and spells of extreme rainfall events in almost every year (source: https:// www.indiawaterportal.org/articles/droughts-maharashtra-lack-management-or-vagaries-climate-change accessed in February 2020). Due to its vulnerability to high disaster, the study area offers a good case to test the potential of DL in multi-step-ahead daily rainfall prediction across different climatic regimes.

The period of analysis is 1941-2005. During this entire period, daily values of nine variables viz. maximum air temperature, geopotential height, longwave radiation, maximum relative humidity, minimum relative humidity, u-wind speed, v-wind speed, sea level pressure and rainfall are obtained. Out of these variables, rainfall values are observed records, and obtained from India Meteorological Department (IMD) with a spatial resolution of 0.25° (latitude) × 0.25° (longitude). Other causal variables are the simulation (hindcast) outputs with a spatial resolution of 2.8125° (latitude) × 2.7906° (longitude) from the second generation Canadian Earth System

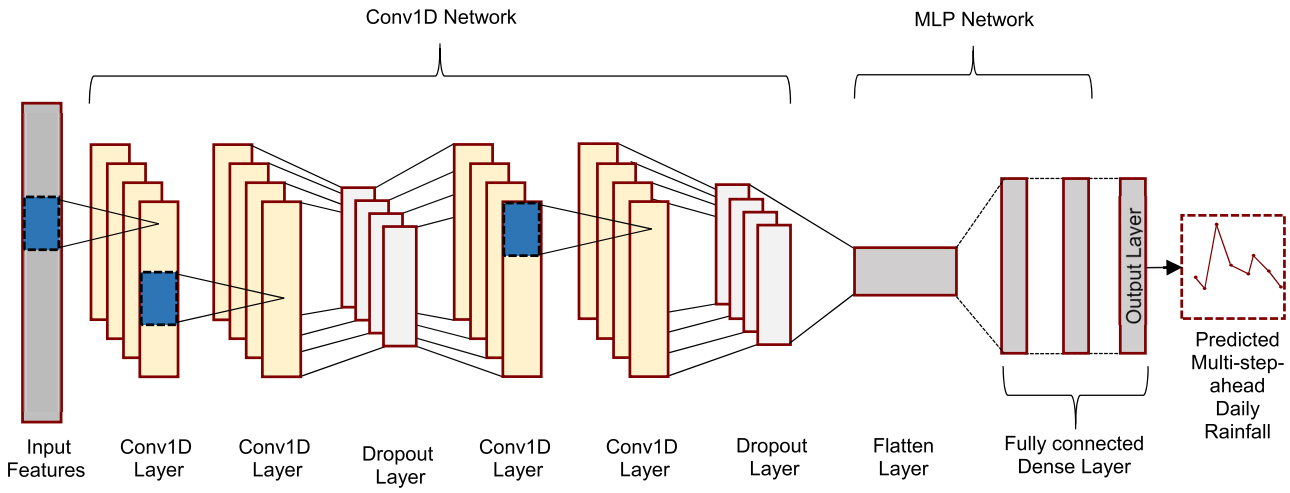

FIGURE 2. A schematic representation of the proposed hybrid Conv1D-MLP model architecture.

Model (CanESM2). The GCM grid points (green stars), IMD grid points (black dots), and the location of the aforementioned cities (red dots) are shown in Fig. 1.

III. METHODOLOGY

A. DATA PREPARATION

In this study, data preparation and handling are entirely done in Scientific Python Development Environment (spyder) notebook. Inverse Distance Weighting (IDW) method is used to obtain the values of all the causal and target variables at the specific city locations from its nearest neighboring four grid intersections.

The variables are scaled between 0 and 1, using Min-Max Scaler class from the sklearn preprocessing available in scipy library to avoid the problem of scaling. After scaling, the dataset is split into two parts, i.e. training and testing. A k-fold cross validation is used to evaluate the prediction skill of the model. Accordingly, seven (k) approximately equal folds are considered and the model is repeatedly trained on the remaining six (k−1) folds, and its performance is measured on the remaining fold. This leads to approximately 85% of the data as the training set and rest 15% as the testing set.

The values of the nine causal variables from previous five consecutive days are simultaneously used as input to predict the daily rainfall with a lead time of 1-day to 5-days.

B. MODEL ARCHITECTURE

The proposed hybrid Conv1D-MLP model is developed in the spyder notebook using Keras, which is a powerful library for large scale DL algorithms. The developed model is a sequential type. A schematic diagram is shown in Fig. 2. The first part consists of a Conv1D network and the second part consists of MLP network.

Conv1D is a type of CNN used in various applications including sequence prediction problems viz. time series analysis and forecasting [50], [51]. It comprises of the input layer, a fully connected output layer with activation function, and between them, there is an arbitrary number of hidden layer(s) along with the activation functions. The function of the input layer is to receive the signal (input data) in a three-dimensional format and transfer it to the hidden layer. Hidden layers are the computational engine of the model. These may have one or more layers of the Conv1D layer, max pooling layer and dropout layer depending on the need for the problem. The Conv1D layer is the main building block of CNN. It consists of filters to extract features from the input signal and kernels to specify the height of the filter. The model is trained on the defined dimension and extracts the hidden information of the sequence. A max-pooling layer (if used) is generally used after the Conv1D layer. It helps in reducing the complexity of the output and also the chance of overtraining. The function of the dropout layer (if used) is to randomly assign zero weights to neurons of the network, which makes it less sensitive towards smaller variation, thus improving the accuracy of the model on unseen data [52].

MLP network receives the input from the Conv1D model (Fig. 2). It is also a fully connected ANN that receives the data in a one-dimensional vector form. Therefore, after the convolutional layer, a flattened layer is added. Next, a fully connected dense layer along with the activation functions are added. The fully connected layers have a more number of neurons than output layers. In this way, neural networks are allowed to think wider before they converged to the output layer. Output layer is also a fully connected dense layer responsible for giving the prediction. There are several parameters to be specified to fix the model architecture. This is problem specific and details are provided in the results and discussion section of this paper.

After configuring the layers, the model is trained on a set of input and output data (numerical values) to learn the relationship between them [53]. Training involves adjustments of weights and biases (parameters) to minimize the error. It is achieved through backpropagation, which adjusts the parameters considering the error (loss) in the predictions.

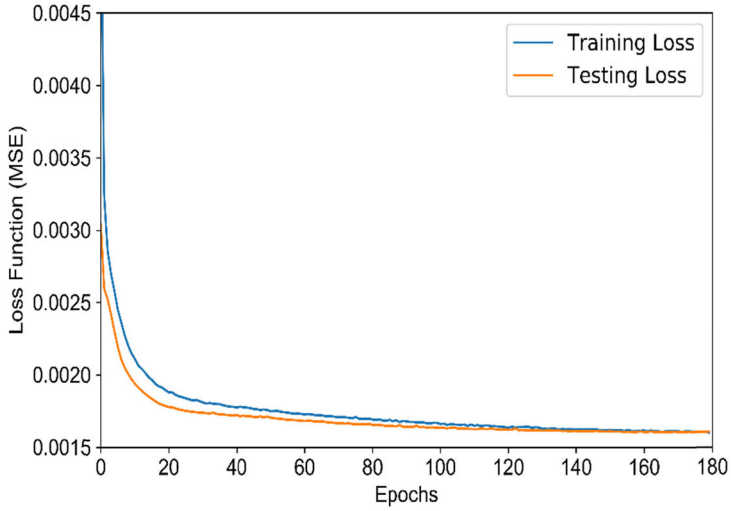

FIGURE 3. A typical plot showing the change in the value of the loss function (MSE) of the proposed hybrid network over the epochs for training and testing data (Navi-Mumbai).

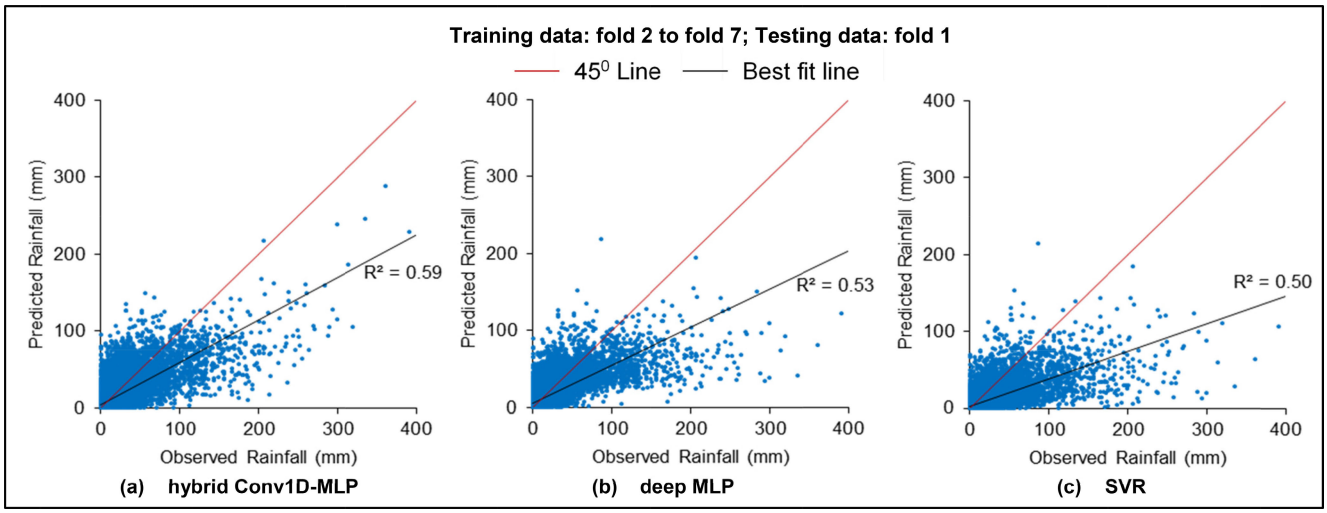

FIGURE 4. Comparative scatter plots between observed and 1-day-ahead predicted rainfall at Navi-Mumbai for the training data as indicated in the figure header: (a) hybrid Conv1D-MLP, (b) deep MLP and (c) SVR.

There are several error based metrics, e.g., mean squared error (MSE) and log loss etc. Once the model is properly trained, it is ready for further use.

C. OTHER MODELS FOR COMPARISON

The performance of the proposed hybrid Conv1D-MLP model is compared with other two models – i) deep learning based Multi-Layered Perceptron (deep MLP) and ii) machine learning approach namely, Support Vector Regression (SVR) [20], [33], [54]. These models are also developed using the same Keras package in the spyder notebook. In both cases, the same proportion of the training and testing datasets are used considering same 7-fold cross-validation.

The deep MLP contains an input layer to feed the data, a number of hidden layers for performing the computation and an output layer for predicting a value or a vector of values. MLPs also have an activation function in the hidden and output layers that maps the sum of the weighted input to the output of the neuron. After configuring the layers, the designed architecture is trained on a set of input and output data to learn the relationship between them [53]. Training of deep MLP model involves adjustments of weights and biases (parameters) to minimize the error using backpropagation algorithm. There are several error based metrics as mentioned before. Properly trained model is used for predictions.

SVR is a supervised ML algorithm, based on the support vectors of Support Vector Machine (SVM) [55]. It helps to map the input data using a nonlinear mapping function to a high dimensional space in order to make them linearly separable. The optimization of SVR was carried out using two

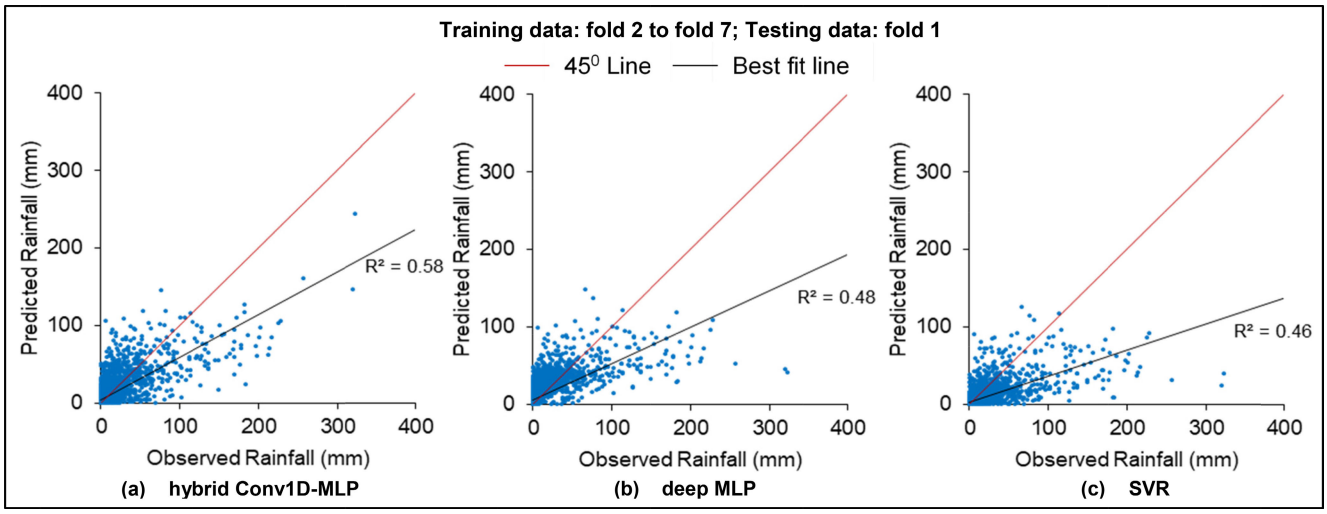

FIGURE 5. Same as Fig. 4 but for the testing data: (a) hybrid Conv1D-MLP, (b) deep MLP and (c) SVR.

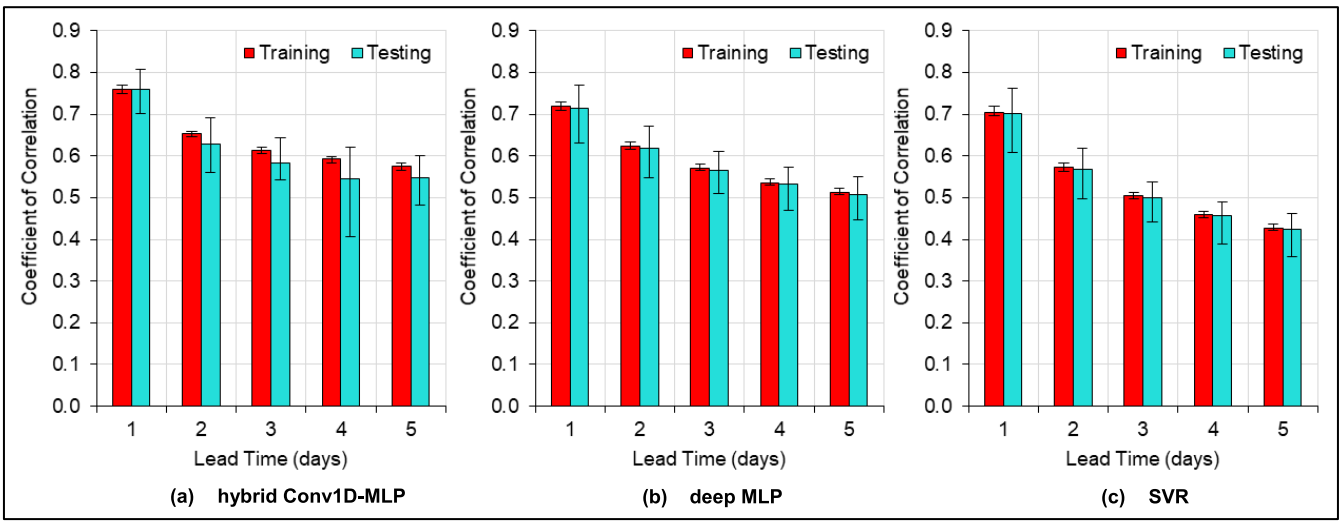

FIGURE 6. Average coefficient of correlation obtained for multi-step-ahead (1-day to 5-day) rainfall prediction during training and testing period using (a) hybrid Conv1D-MLP, (b) deep MLP and (c) SVR. The error bar shows the range (maximum to minimum) of the metric for each lead time obtained from 7 folds.

regularization parameter Gamma (γ) and cost function (C) of radial basis function (RBF). RBF is a nonlinear kernel function popularly used in SVR [56]. These two parameters of RBF are interdependent and their optimal value is found by trial and error analysis. More details on SVR can be found in the literature [12], [55].

D. PERFORMANCE EVALUATION CRITERIA

The performance of the hybrid Conv1D-MLP model is compared with SVR and deep MLP models through three statistical measures viz. Root Mean Squared Error (RMSE), coefficient of correlation (r), Nash–Sutcliffe Efficiency (NSE). The prediction skill is assessed for k-fold cross-validation and average values of aforementioned metrics for all k folds are obtained.

The coefficient of correlation is a measure of linear association between two variables. The value of r is computed as:

$$r = \frac{\sum_{t=1}^{n} \left(Y_t - \bar{Y}\right)\left(Y_t' - \bar{Y}'\right)}{\sum_{t=1}^{n} \left(Y_t - \bar{Y}\right) \sum_{t=1}^{n} \left(Y_t' - \bar{Y}'\right)} \tag{1}$$

where, Y_t(mm) and Y_t'(mm) represent the observed and predicted rainfall at the time t, (mm) and $\bar{Y}$(mm) are the means of predicted and observed rainfall, respectively, and n is the total number of data points. The value of r ranges between $[-1, 1]$, where -1 represents a perfect negative linear association, 0 denote no linear association, and 1 represents a perfect positive linear association. Higher the value of r, better the model performance.

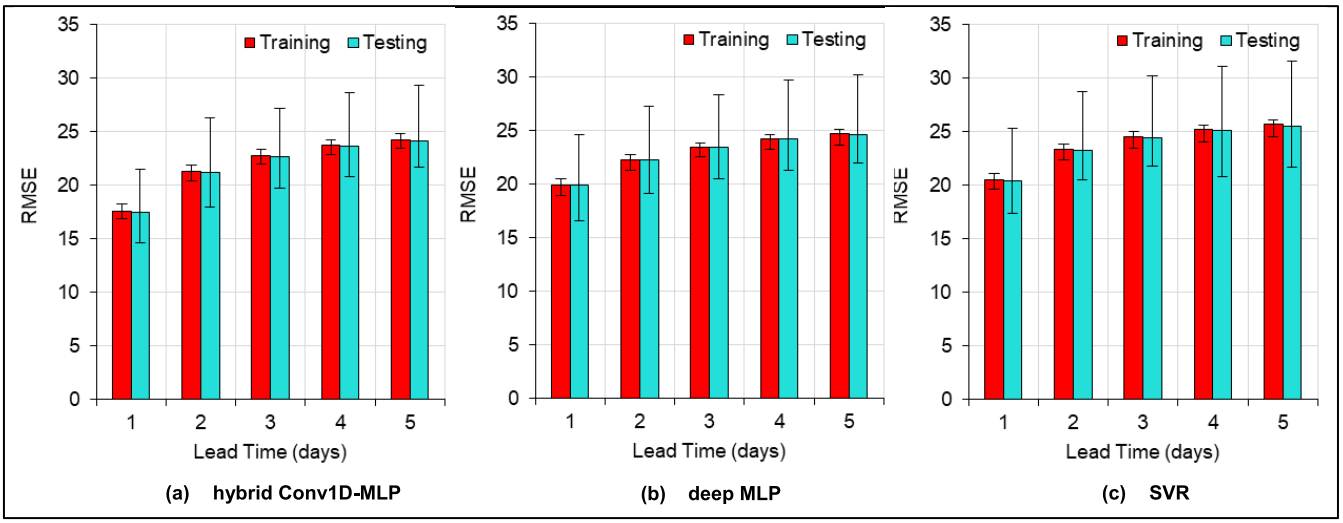

FIGURE 7. Average RMSE obtained for multi-step-ahead (1-day to 5-day) rainfall prediction during training and testing period using (a) hybrid Conv1D-MLP, (b) deep MLP and (c) SVR. The error bar shows the range (maximum to minimum) of the metric for each lead time obtained from 7 folds.

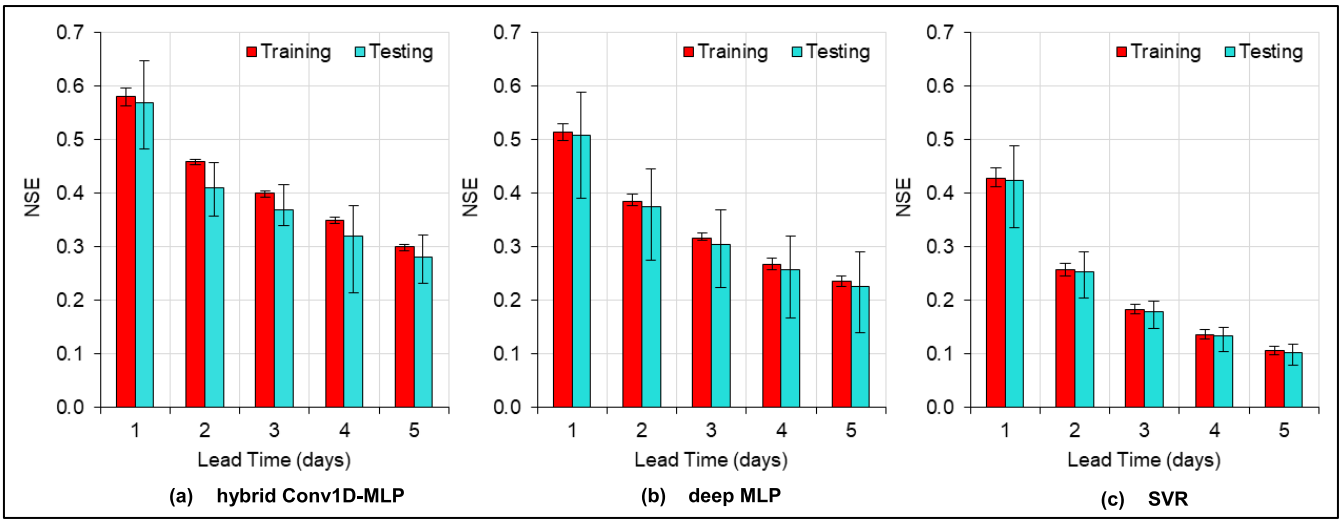

FIGURE 8. Average NSE obtained for multi-step-ahead (1-day to 5-day) rainfall prediction during training and testing period using (a) hybrid Conv1D-MLP, (b) deep MLP and (c) SVR. The error bar shows the range (maximum to minimum) of the metric for each lead time obtained from 7 folds.

RMSE is used frequently to measure the difference between observed and predicted values. It is always positive and a lower RMSE indicate a better model performance. RMSE is expressed as:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n} \left(Y_t - Y_t'\right)^2}{n}} \tag{2}$$

NSE is used to measure the efficiency of the proposed model [57]. The value NSE is computed by the eq. (3).

$$\text{NSE} = 1 - \frac{\sum_{t=1}^{n} \left(Y_t - Y_t'\right)^2}{\sum_{t=1}^{n} \left(Y_t - \bar{Y}\right)^2} \tag{3}$$

NSE ranges between $(-\infty, 1]$. A value greater than 0 indicates a better efficiency of the model as compared to a value equal to 0 that signifies the predicted values are as good as the mean of the observed values. NSE values less than zero indicates an unacceptable model performance.

IV. RESULTS AND DISCUSSIONS

A. MODEL PARAMETERS/CONFIGURATIONS

Several network architectures were evaluated by varying model parameters (viz. number of hidden layers, number of filters, kernel size, pooling size, and dropout percentage) and optimizing several hyperparameters (viz. learning rate, batch size, number of epochs, loss functions, and activation functions) in order to ascertain the best possible architectural configuration. The finalized architecture of the proposed

TABLE 1. Configurations of the proposed hybrid Conv1D-MLP model. Please refer Fig. 2 for a pictorial representation of different layers.

Layer No.	Layer	Type	Parameter of Layers				
			Activation Function.	Kernel size	No. of Filters	Dropout size	Neurons
1	Conv1D	Convolutional Layer	ReLU	1	120	-	-
2	Conv1D	Convolutional Layer	ReLU	1	96	-	-
3	Dropout	Dropout Layer	-	-	-	0.20	-
4	Conv1D	Convolutional Layer	ReLU	1	72	-	-
5	Conv1D	Convolutional Layer	ReLU	1	48	-	-
6	Dropout	Dropout Layer	-	-	-	0.20	-
7	Flatten	Flatten Layer	-	-	-	-	-
8	Dense	Fully Connected Layer	ReLU	-	-	-	200
9	Dense	Fully Connected Layer	ReLU	-	-	-	100
10	Dense	Fully Connected Layer (output layer)	Linear	-	-	-	5

hybrid model comprises of ten layers (Fig. 2). Details of the finalized configurations are shown in Table 1.

The entire set of nine causal variables (also known as input features) are fed as input to the first layer. The input shape consists of 5 previous time steps (days) with nine features, arranged in a single row. The first and second layer added to the model are Conv1D layers. The convolutional layers help to identify the pattern and excerpt hidden information in the input data [50]. The first and second Conv1D layers comprise of 120 and 96 filters, respectively. Kernel size of 1 and ReLU activation functions are used in these layers. Kernel size represents the height of a filter and a threshold value (bias) is set so that the output from a unit (also known as neuron) moves to the next layer if value of the activation function crosses the threshold value. The third layer is a dropout layer. It prevents overfitting and also helps to improve the performance of the model [52]. The dropout value is set as 0.20, i.e., 20 percent of neurons will be dropped from the second layer. The fourth and fifth layers are also the Conv1D layers, comprising of 72 and 48 filters, respectively. These layers also have the same kernel size and activation function as that of the previous Conv1D layers.

Subsequently, another dropout layer with dropout value of 0.20 is added as a sixth layer to the model. After providing a successive one dimension convolutional and dropout layers, a flatten layer is added as a seventh layer to the model. It receives the multidimensional input from the previous layer and converts them into a one-dimensional array. Further, the eighth and ninth layers added to the model are fully connected dense layers that are configured with 200 and 100 neurons, respectively. Finally, the output layer (tenth layer), consisting of five neurons, is added to the model. These five neurons are the 1-day to 5-day-ahead predicted values of rainfall.

Next, a batch size of 365 and 180 number of epochs are selected for training the aforementioned hybrid model architecture by trial and error method. The architecture uses the Mean Absolute Error (MAE) to calculate the loss function and efficient Adam version of stochastic gradient descent [30] with modified value of learning rate (5×10^{-5}) for selecting the best possible configuration.

An example of gradual change of the loss function over epochs is shown in Fig. 3 for the station Navi-Mumbai. It is found that the values of the loss function during training and testing are comparable at epoch 180. This ensures no overfitting/underfitting during the training process. It is checked for all other stations before selecting 180 as the number of epochs.

The model configuration was retained same during training and testing combinations for all the seven-folds. After successful completion of training, the input of the testing dataset was fed into the model to measure the model performance.

B. MODEL PERFORMANCE

Model performance (viz. r, RMSE and NSE) across all the 7 folds in case of 1-day-ahead rainfall prediction at all twelve cities are summarized in Table 2. It may be noted that the values of the reported performance metrics in Table 2 are averaged across all the folds. While checking fold-wise results (not reported), it was noticed that the performances during training and testing are comparable with training/testing performance marginally better than testing/training. This ensures that the model is neither overtrained nor undertrained.

Performance metrics for the proposed hybrid Conv1D-MLP model and other two models (deep MLP and SVR) are presented side by side for comparison. First, deep MLP and SVR are compared among themselves. It is noticed that in case of deep MLP, the average RMSE values range from 5.89 to 19.92, average r values range from 0.4 to 0.72, and average NSE values range from 0.16 to 0.51 during the training period of the deep MLP model. During the testing

TABLE 2. Performance statistics viz. *r*, RMSE and NSE respectively from top to bottom (averaged across different folds) at different cities.

City	hybrid Conv1D-MLP		deep MLP		SVR	
	Training	Testing	Training	Testing	Training	Testing
Akola	0.41	0.41	0.40	0.39	0.33	0.33
	7.82	7.82	7.79	8.26	8.74	8.71
	0.17	0.17	0.16	0.15	-0.03	-0.04
Amravati	0.47	0.47	0.45	0.44	0.41	0.41
	7.23	7.20	7.13	8.11	7.81	7.76
	0.23	0.22	0.22	0.19	0.10	0.10
Aurangabad	0.44	0.44	0.41	0.39	0.36	0.36
	6.57	6.55	6.64	6.47	7.46	7.41
	0.20	0.19	0.18	0.15	-0.04	-0.04
Chandrapur	0.47	0.47	0.42	0.39	0.39	0.40
	10.45	10.39	10.75	13.52	11.91	11.81
	0.22	0.21	0.20	0.14	-0.02	-0.02
Kolhapur	0.48	0.47	0.47	0.46	0.43	0.42
	6.28	6.28	6.38	6.57	6.66	6.62
	0.23	0.22	0.22	0.20	0.13	0.13
Latur	0.50	0.50	0.48	0.48	0.44	0.43
	6.25	6.23	6.32	6.57	7.34	7.30
	0.25	0.25	0.32	0.28	-0.03	-0.04
Nagpur	0.49	0.49	0.46	0.44	0.41	0.41
	8.47	8.45	9.76	9.92	9.87	9.83
	0.24	0.24	0.21	0.18	-0.03	-0.03
Nanded	0.48	0.47	0.43	0.41	0.41	0.41
	8.23	8.20	8.31	8.62	9.47	9.41
	0.23	0.22	0.20	0.16	-0.03	-0.03
Nashik	0.49	0.48	0.46	0.45	0.45	0.45
	6.29	6.26	6.32	6.30	6.74	6.67
	0.24	0.23	0.22	0.21	0.13	0.12
Navi-Mumbai	0.76	0.76	0.72	0.72	0.70	0.70
	17.61	17.45	19.92	19.87	20.52	20.41
	0.58	0.57	0.51	0.51	0.43	0.42
Pune	0.62	0.58	0.59	0.54	0.58	0.50
	7.30	7.35	7.36	7.48	7.22	7.68
	0.39	0.29	0.35	0.24	0.34	0.23
Solapur	0.49	0.48	0.46	0.45	0.44	0.44
	5.85	5.85	5.89	5.89	6.36	6.33
	0.24	0.23	0.21	0.19	0.10	0.10

period, these values are slightly less and range from 5.89 to 19.87, 0.39 to 0.72, and 0.14 to 0.51, respectively. In case of SVR, NSE values obtained during the training as well as the testing period at Akola, Aurangabad, Chandrapur, Latur, Nagpur and Nanded cities are below zero, indicating mean of observed rainfall would have been a better prediction than the modeled value. Also, for the rest of the cities, the performance of deep MLP was found better in terms of all three metrics. Thus, in general, performance of the deep MLP is better than SVR.

Next, the performance of proposed Conv1D-MLP model is compared against both deep MLP and SVR. In contrast to these models, the performance of the proposed hybrid Conv1D-MLP model is better. The improvements vary from marginally to reasonably across the cities. The average r value ranges from 0.41 to 0.76, average NSE values range from 0.17 to 0.58 and average RMSE values range from 5.85 to 17.61 during the training period whereas, during the testing period these values range from 0.41 to 0.76, 0.17 to 0.57 and 5.81 to 17.45 respectively. Evidently the performance of the hybrid Conv1D-MLP is better at all cities.

In particular, the range of daily rainfall is captured by the proposed hybrid model, whereas the performance of deep MLP and SVR is not able to capture the same. It is reflected through the high RMSE values and scatter plots between observed and predicted rainfall by the two models. One such case at Navi-Mumbai is considered and comparative scatter plots are shown at 1-day-ahead prediction for fold 1 as testing period and rest as training period (Fig. 4 and Fig. 5). Results for other folds as testing periods (and rests as training period) are shown in the supplementary document (Figs. 1 to 12). These figures show that the deep MLP and SVR models are not able to predict rainfall values exceeding 150 mm and 180 mm respectively during the testing period. In contrast, the hybrid of Conv1D-MLP is able to predict at least some of the extreme rainfall values indicating a better way of capturing the range of observed rainfall.

Next, the results for prediction with other lead times (2-day to 5-day-ahead) are explored. Average values of all the three metrics, their range (maximum and minimum) across all the seven folds during training and testing periods for different lead times are computed. For discussion, a typical plot of r, RMSE and NSE values with maximum and minimum values as an error bar are presented in Figs. 6, 7 and 8, respectively, for the station Navi Mumbai. It is noticed that the performance during training and testing is comparable, ensuring no overfitting/underfitting during model training for all the lead times. Secondly, the model performance gradually decreases with the increase in the lead time of prediction, i.e., 1-day to 5-day-ahead predictions. However, the performance of Conv1D-MLP is reasonably good even at 5-days-ahead prediction with the range of r as 0.25 to 0.55, RMSE as 6.6 to 24.19, and NSE as 0.28 to 0.11 during the testing period. It is also found that the performance of the hybrid Conv1D-MLP model is better than other models (deep MLP and SVR) for all the lead times.

V. CONCLUSION

This study presents the potential of a proposed hybrid DL approach, namely hybrid Conv1D-MLP model, for multi-step-ahead (1-day to 5-day in advance) prediction of daily rainfall using GCM simulated meteorological variables as inputs. The performance of two existing models, viz. deep MLP and SVR, are compared with the proposed model at

different stations located in different climate regimes in terms of rainfall characteristics. Following conclusions are drawn from the study:

- Deep learning has the potential to capture the non-linear relationship between the causal variables and rainfall variability. Therefore, it can be effectively used for its prediction with a couple of days in advance.
- Simulated meteorological variables by climate models can be beneficially used for better rainfall prediction than the climate models itself exploiting the potential of deep learning approach. This is even more beneficial considering the fact that the observed data of meteorological variables are either not available or sparsely available at many places.
- The correspondence between the observed and the predicted rainfall is better than the correspondence between the observed and the GCM simulated rainfall magnitude. This can be considered as the potential of the hybrid Conv1D-MLP in improving the rainfall prediction.
- The performance of the hybrid Conv1D-MLP model is better than the deep MLP and SVR model. It is further noticed that the extreme daily rainfall magnitudes are better captured by the hybrid model as compared to other models. In particular, daily rainfall magnitude greater than 200 mm is better captured by the hybrid Conv1D-MLP model as compared to the deep MLP and SVR model. Thus, the hybrid Conv1D-MLP model may be a better choice, considering the extreme rainfall events.
- It is also noticed that the performance of hybrid Conv1D-MLP model gradually reduces as the prediction lead time increases from 1-day to 5-day-ahead. This is as per expectation, however, it is noticed that the performance of Conv1D-MLP is reasonably good even at 5-days-ahead prediction. In general, the predictions obtained from proposed hybrid DL model can be helpful in agriculture, irrigation scheduling, and even flooding due to heavy rainfall.

There is a scope to apply the hybrid DL approach for simulation of rainfall field instead of point rainfall (as demonstrated in this paper). This is kept as the future scope of this study.

REFERENCES

[1] Core Writing Team, R. K. Pachauri, and L. A. Meyer, Eds., *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, document, IPCC, Geneva, Switzerland, 2014, p. 151.

[2] *Annual Report, 2017*, Ministry Earth Sci., New Delhi, India, 2017.

[3] S. Dash and R. Maity, "Temporal evolution of precipitation-based climate change indices across India: Contrast between pre- and post-1975 features," *Theor. Appl. Climatol.*, vol. 138, nos. 3–4, pp. 1667–1678, Nov. 2019.

[4] K. C. Luk, J. E. Ball, and A. Sharma, "A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting," *J. Hydrol.*, vol. 227, nos. 1–4, pp. 56–65, Jan. 2000.

[5] S. Hong, Z. Zhou, E. Zio, and K. Hong, "Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method," *Digit. Signal Process.*, vol. 27, no. 1, pp. 159–166, Apr. 2014.

[6] S. Hong, Z. Zhou, E. Zio, and W. Wang, "An adaptive method for health trend prediction of rotating bearings," *Digit. Signal Process.*, vol. 35, pp. 117–123, Dec. 2014.

[7] S. Hong, Z. Zhou, C. Lu, B. Wang, and T. Zhao, "Bearing remaining life prediction using Gaussian process regression with composite kernel functions," *J. Vibroengineering*, vol. 17, no. 2, pp. 695–704, 2015.

[8] S. Hong, H. Yang, E. Zio, and N. Huang, "A novel dynamics model of fault propagation and equilibrium analysis in complex dynamical communication network," *Appl. Math. Comput.*, vol. 247, pp. 1021–1029, Nov. 2014.

[9] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nat. Methods*, vol. 13, no. 1, p. 35, 2015.

[10] R. Govindaraju, A. Rao, D. Leib, Y. Najjar, H. Gupta, and A. Hjelmfelt, Jr., "Artificial neural networks in hydrology. II: Hydrologic application," *J. Hydrol. Eng.*, vol. 5, no. 2, pp. 115–123, 2000.

[11] K. Hsu, H. V. Gupta, and S. Sorooshian, "Artificial neural network modeling of the rainfall-runoff process," *Water Resour. Res.*, vol. 31, no. 10, pp. 2517–2530, 1995.

[12] R. Maity, P. P. Bhagwat, and A. Bhatnagar, "Potential of support vector regression for prediction of monthly streamflow using endogenous property," *Hydrol. Processes*, vol. 24, no. 7, pp. 917–923, Mar. 2010.

[13] M. Pal, R. Maity, J. V. Ratnam, M. Nonaka, and S. K. Behera, "Long-lead prediction of ENSO modoki index using machine learning algorithms," *Sci. Rep.*, vol. 10, no. 1, pp. 1–13, Dec. 2020.

[14] G. J. Sawale and S. R. Gupta, "Use of artificial neural network in data mining for weather forecasting," *Int. J. Comput. Sci. Appl.*, vol. 6, no. 2, pp. 383–387, Apr. 2013.

[15] N. A. Charaniya and S. V. Dudul, "Design of neural network models for daily rainfall prediction," *Int. J. Comput. Appl.*, vol. 61, no. 14, pp. 23–27, 2013.

[16] J. Lee, C.-G. Kim, J. Lee, N. Kim, and H. Kim, "Application of artificial neural networks to rainfall forecasting in the geum river basin, korea," *Water*, vol. 10, no. 10, p. 1448, 2018.

[17] T. Partal and Ö. Kişi, "Wavelet and neuro-fuzzy conjunction model for precipitation forecasting," *J. Hydrol.*, vol. 342, nos. 1–2, pp. 199–212, Aug. 2007.

[18] C. L. Wu, K. W. Chau, and C. Fan, "Prediction of rainfall time series using modular artificial neural networks coupled with data-preprocessing techniques," *J. Hydrol.*, vol. 389, nos. 1–2, pp. 146–167, Jul. 2010.

[19] K. P. Moustris, I. K. Larissi, P. T. Nastos, and A. G. Paliatsos, "Precipitation forecast using artificial neural networks in specific regions of greece," *Water Resour. Manage.*, vol. 25, no. 8, pp. 1979–1993, Jun. 2011.

[20] S. S. Kashid and R. Maity, "Prediction of monthly rainfall on homogeneous monsoon regions of India based on large scale circulation patterns using genetic programming," *J. Hydrol.*, vols. 454–455, pp. 26–41, Aug. 2012.

[21] F. Mekanik, M. A. Imteaz, S. Gato-Trinidad, and A. Elmahdi, "Multiple regression and artificial neural network for long-term rainfall forecasting using large scale climate modes," *J. Hydrol.*, vol. 503, pp. 11–21, Oct. 2013.

[22] R. V. Ramana, B. Krishna, S. R. Kumar, and N. G. Pandey, "Monthly rainfall prediction using wavelet neural network analysis," *Water Resour. Manage.*, vol. 27, no. 10, pp. 3697–3711, Aug. 2013.

[23] J. Abbot and J. Marohasy, "Input selection and optimisation for monthly rainfall forecasting in queensland, Australia, using artificial neural networks," *Atmos. Res.*, vol. 138, pp. 166–178, Mar. 2014.

[24] E. Hernández, V. Sanchez-Anguix, V. Julian, J. Palanca, and N. Duque, "Rainfall prediction: A deep learning approach," in *Proc. Int. Conf. Hybrid Artif. Intell. Syst.*, vol. 1, 2016, pp. 151–162.

[25] T. Li, H. Shen, Q. Yuan, X. Zhang, and L. Zhang, "Estimating ground-level PM2.5by fusing satellite and station observations: A geo-intelligent deep learning approach," *Geophys. Res. Lett.*, vol. 44, no. 23, pp. 11985–11993, Dec. 2017.

[26] D. Kumar, A. Singh, P. Samui, and R. K. Jha, "Forecasting monthly precipitation using sequential modelling," *Hydrol. Sci. J.*, vol. 64, no. 6, pp. 690–700, Apr. 2019.

[27] X. Shi, Z. Gao, L. Lausen, H. Wang, W.-K. Wong, W.-C. Woo, and D. Yeung, "Deep learning for precipitation nowcasting: A benchmark and a new model," in *Proc. 31st Conf. Neural Inf. Process. Syst. (NIPS)*, 2017, pp. 5617–5627.

[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. 25th Int. Conf. Neural Inf. Process. Syst.*, 2012, vol. 60, no. 6, pp. 1097–1105.

[29] T. Mikolov, A. Deoras, D. Povey, L. Burget, and J. Cernocky, "Strategies for training large scale neural network language models," in *Proc. IEEE Workshop Autom. Speech Recognit. Understand.*, Dec. 2011, pp. 196–201.
[30] T. N. Sainath, A.-R. Mohamed, B. Kingsbury, and B. Ramabhadran, "Deep convolutional neural networks for LVCSR," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 8614–8618.
[31] J. Ma, R. P. Sheridan, A. Liaw, G. E. Dahl, and V. Svetnik, "Deep neural nets as a method for quantitative structure-activity relationships," *J. Chem. Inf. Model.*, vol. 55, no. 2, pp. 263–274, 2015.
[32] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (Almost) from scratch," *J. Mach. Learn. Res.*, vol. 12, pp. 2493–2537, Aug. 2011.
[33] A. Haidar and B. Verma, "Monthly rainfall forecasting using one-dimensional deep convolutional neural network," *IEEE Access*, vol. 6, pp. 69053–69063, 2018.
[34] J. N. K. Liu, Y. Hu, J. J. You, and P. W. Chan, "Deep neural network based feature representation for weather forecasting," in *Proc. Int. Conf. Artif. Intell. (ICAI)*, 2014, p. 1.
[35] Y. Liu, E. Racah, Prabhat, J. Correa, A. Khosrowshahi, D. Lavers, K. Kunkel, M. Wehner, and W. Collins, "Application of deep convolutional neural networks for detecting extreme weather in climate datasets," 2016, *arXiv:1605.01156*. [Online]. Available: http://arxiv.org/abs/1605.01156
[36] P. Zhang, L. Zhang, H. Leung, and J. Wang, "A deep-learning based precipitation forecasting approach using multiple environmental factors," in *Proc. IEEE Int. Congr. Big Data (BigData Congress)*, Jun. 2017, pp. 193–200.
[37] A. Ghaderi, B. M. Sanandaji, and F. Ghaderi, "Deep forecast: Deep learning-based spatio-temporal forecasting," 2017, *arXiv:1707.08110*. [Online]. Available: http://arxiv.org/abs/1707.08110
[38] S. Aswin, P. Geetha, and R. Vinayakumar, "Deep learning models for the prediction of rainfall," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2018, pp. 657–661.
[39] C. Hu, Q. Wu, H. Li, S. Jian, N. Li, and Z. Lou, "Deep learning with a long short-term memory networks approach for rainfall-runoff simulation," *Water*, vol. 10, no. 11, p. 1543, 2018.
[40] A. G. Salman, Y. Heryadi, E. Abdurahman, and W. Suparta, "Weather forecasting using merged long short-term memory model (LSTM) and autoregressive integrated moving average (ARIMA) model," *J. Comput. Sci.*, vol. 14, no. 7, pp. 930–938, Jul. 2018.
[41] J. Wang and Y. Li, "Multi-step ahead wind speed prediction based on optimal feature extraction, long short term memory neural network and error correction strategy," *Appl. Energy*, vol. 230, pp. 429–443, Nov. 2018.
[42] E. Hernández, V. Sanchez-Anguix, V. Julian, J. Palanca, and N. Duque, "Rainfall prediction: A deep learning approach," in *Proc. Int. Conf. Hybrid Artif. Intell. Syst.*, 2016, pp. 151–162.
[43] M. Shi, F. Xie, Y. Zi, and J. Yin, "Cloud detection of remote sensing images by deep learning," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 701–704.
[44] M. Mudigonda, S. Kim, A. Mahesh, S. Kahou, K. Kashinath, D. Williams, V. Michalski, T. O'Brien, and M. Prabhat, "Segmenting and tracking extreme climate events using neural networks," in *Proc. 31st Conf. Neural Inf. Process. Syst.*, 2017, pp. 1–5.
[45] M. Le Goff, J.-Y. Tourneret, H. Wendt, M. Ortner, and M. Spigai, "Deep learning for cloud detection," in *Proc. 8th Int. Conf. Pattern Recognit. Syst. (ICPRS)*, no. 1, 2017, p. 10.
[46] D. Matsuoka, M. Nakano, D. Sugiyama, and S. Uchida, "Deep learning approach for detecting tropical cyclones and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model," *Progr. Earth Planet. Sci.*, vol. 5, no. 1, pp. 1–16, Dec. 2018.
[47] M. Swapna and N. Sudhakar, "A hybrid model for rainfall prediction using both parametrized and time series models," *Int. J. Pure Appl. Math.*, vol. 119, no. 14, pp. 1549–1556, 2018.
[48] T. Kurth, S. Treichler, J. Romero, M. Mudigonda, N. Luehr, E. Phillips, A. Mahesh, M. Matheson, J. Deslippe, M. Fatica, P. Prabhat, and M. Houston, "Exascale deep learning for climate analytics," in *Proc. Int. Conf. High Perform. Comput., Netw., Storage Anal.*, Nov. 2018, pp. 649–660.
[49] C. Shen, "A transdisciplinary review of deep learning research and its relevance for water resources scientists," *Water Resour. Res.*, vol. 54, no. 11, pp. 8558–8593, Nov. 2018.
[50] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: A survey," 2019, *arXiv:1905.03554*. [Online]. Available: http://arxiv.org/abs/1905.03554
[51] S. Harbola and V. Coors, "One dimensional convolutional neural network architectures for wind prediction," *Energy Convers. Manage.*, vol. 195, pp. 70–75, Sep. 2019.
[52] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
[53] J. Zou, Y. Han, and S. S. So, "Overview of artificial neural networks," in *Artificial Neural Networks* (Methods in Molecular Biology), vol. 458. 2008, pp. 15–23.
[54] R. Maity and D. N. Kumar, "Basin-scale stream-flow forecasting using the information of large-scale atmospheric circulation phenomena," *Hydrol. Processes*, vol. 22, no. 5, pp. 643–650, Feb. 2008.
[55] H. Drucker, C. J. C. Surges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," in *Proc. Adv. Neural Inf. Process. Syst.*, Nov. 1997, pp. 155–161.
[56] K. Y. Choy and C. W. Chan, "Modelling of river discharges and rainfall using radial basis function networks based on support vector regression," *Int. J. Syst. Sci.*, vol. 34, nos. 14–15, pp. 763–773, Nov. 2003.
[57] S. Pichuka, R. Prasad R, R. Maity, and H. Kunstmann, "Development of a method to identify change in the pattern of extreme streamflow events in future climate: Application on the Bhadra reservoir inflow in India," *J. Hydrol. Regional Stud.*, vol. 9, pp. 236–246, Feb. 2017.

MOHD IMRAN KHAN received the B.Tech. degree in civil engineering from Uttar Pradesh Technical University (UPTU), Lucknow, India, in 2013, and the M.Tech. degree in water resources engineering from the National Institute of Technology, Patna, India, in 2015. He is currently pursuing the Ph.D. degree with the Department of Civil Engineering, Indian Institute of Technology Kharagpur, India. His research interest includes hydroclimatology, specifically the application of machine learning (ML) and deep learning (DL) in hydroclimatic analysis.

RAJIB MAITY received the M.E. (First Class with Distinction) and Ph.D. degrees from the Indian Institute of Science (IISc), Bengaluru, India, in 2004 and 2007, respectively. He is currently an Associate Professor with the Civil Engineering Department, Indian Institute of Technology Kharagpur, India. He has published two books and more than 120 research articles in different peer-reviewed journals and conferences and chapters in books. His research interests include hydroclimatology, climate impacts on water resources, stochastic hydrology, hydrologic applications of satellite remote sensing, and so on. Some of his professional recognitions include the Humboldt Fellowship (Experienced Researchers), the James Rennell MoES Young Fellow (MoES, GoI), the Faculty Excellence Award, IIT Kharagpur, the Prof. R. J. Garde Research Award, the ASCE 2011 Outstanding Reviewer, USA, the Emerging Leaders Fellowship, Australia, the BOYSCAST Fellowship, the IEI Young Engineers Award, the DAAD Fellowship for IIT Faculty, Germany, the International ICE WaRM Fellowship, Australia, and the Prof. N. S. Govinda Rao Memorial Gold Medal, IISc. Very recently, he was awarded the Humboldt Research Fellowship for Experienced Researchers by the Alexander von Humboldt Foundation, Germany. He received the Gold Medal for his M.E. degree from the Indian Institute of Science (IISc). He is currently serves as an Associate Editor for the *Journal of Earth System Sciences (JESS)* (Springer) and *ISH Journal of Hydraulic Engineering* (Taylor and Francis).