Potential of Deep Learning in drought assessment by extracting information from hydrometeorological precursors

Rajib Maity, Mohd Imran Khan, Subharthi Sarkar, Riya Dutta, Subhra Sekhar Maity, Manali Pal and Kironmala Chanda

ABSTRACT

This study explores the potential of the Deep Learning (DL) approach to develop a model for basin-scale drought assessment using information from a set of primary hydrometeorological precursors, namely air temperature, surface pressure, wind speed, relative humidity, evaporation, soil moisture and geopotential height. The novelty of the study lies in extracting the information from the hydrometeorological precursors through the efficacy of the DL algorithm, based on a one-dimensional convolutional neural network. Drought-prone regions, from where our study basins are selected, often suffer from the vagaries of rainfall that leads to drought-like situations. It is established that the proposed DL-based model is able to capture the underlying complex relationship between rainfall and the set of aforementioned hydrometeorological variables and, subsequently, shows its promise for the basin-scale meteorological drought assessment as revealed through different performance metrics and skill scores. The accuracy of simulating the correct drought category, among the seven categories, is also high (>70%). Moreover, in general, the skill of any climate model is much higher for the primary meteorological variables as compared with other secondary or tertiary variables/phenomena, like droughts. Thus, the novelty of the proposed DL-based model also lies in the improved assessment of ensuing basin-scale meteorological droughts using the projected meteorological precursors and may lead to new research directions.

Key words | one-dimensional convolutional neural network (Conv1D), Deep Learning, drought, hydrometeorology, machine learning, Standardized Precipitation Anomaly Index (SPAI)

HIGHLIGHTS

- This study explores the potential of Deep Learning (DL) approach to capture the hidden complex hydrometeorological association.
- A DL-based model is developed for basin-scale drought assessment using the hidden complex relationship from a set of primary hydrometeorological precursors.
- DL may be effective to use simulated primary hydrometeorological variables from climate models to assess more complex phenomena.

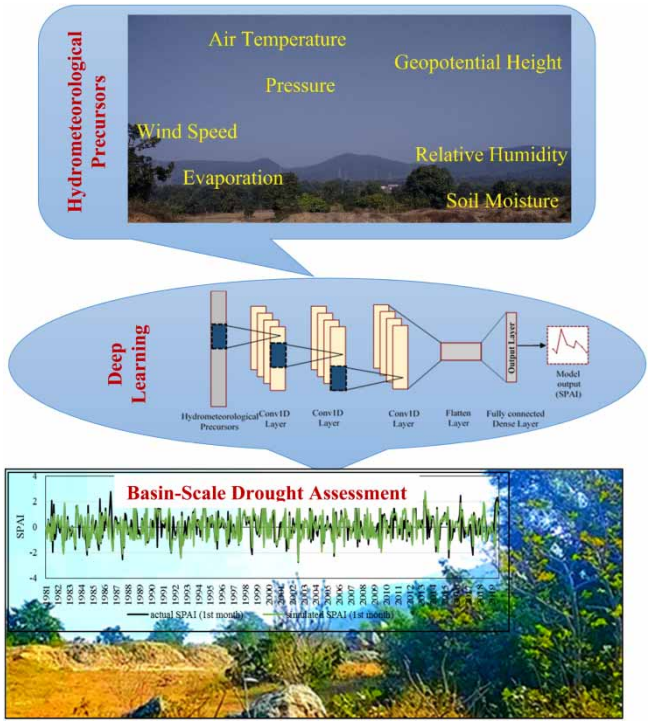
Rajib Maity (corresponding author)
Mohd Imran Khan
Subharthi Sarkar
Riya Dutta
Subhra Sekhar Maity
Department of Civil Engineering,
Indian Institute of Technology Kharagpur,
Kharagpur, 721302 West Bengal,
India
E-mail: *rajib@civil.iitkgp.ac.in*

Manali Pal
Department of Civil Engineering,
National Institute of Technology Warangal,
Warangal, 506004 Telangana,
India

Kironmala Chanda
Department of Civil Engineering,
Indian Institute of Technology (Indian School of Mines),
Dhanbad, Jharkhand 826004,
India

doi: 10.2166/wcc.2021.062

GRAPHICAL ABSTRACT



INTRODUCTION

Occurrences of droughts have a profound impact on water stress affecting both surface and groundwater resources causing reduced water supply and deteriorated water quality (Mishra & Singh 2010; Adnan *et al.* 2018; Mukherjee *et al.* 2018). The impact of drought is more profound in monsoon-dominated countries like India, where the agricultural production heavily depends on monsoon rainfall and the economic growth is dominantly dependent on agriculture (FAO 2019). In the last five decades, India has witnessed a drought at least once in every three years, making it one of the most vulnerable drought-prone countries. Prolonged, widespread and more frequent droughts have occurred in consecutive years for the last two decades (FAO 2002; World Bank 2003). However, droughts should not be mixed up with aridity, which is defined as a permanent imbalance in the water availability in a region characterized by low average annual precipitation along with low moisture availability and low carrying capacity of the ecosystems (Sanderson 1990; Pereira *et al.* 2002). On the other hand, droughts are the temporary imbalance of water availability due to prolonged below normal precipitation resulting in a shortage of soil moisture availability and other sources of water at a region with uncertain duration, frequency and severity. Such uncertainties lead to difficulties in the reliable prediction of upcoming drought status (Pereira *et al.* 2002).

The complexity associated with different types of drought, their dynamic nature and widespread impacts calls for an urgent need to establish early warning systems with reliable drought prediction mechanisms for the most vulnerable communities to be ready with a hands-on drought mitigation plan (World Meteorological Organization, WMO 2006). The probabilistic and risk-based drought monitoring and prediction information are also crucial for effective drought relief management throughout an extreme event. Yet, to date, the task of developing a reliable prediction model remains challenging due to the

random and nonlinear nature of drought variables, complexities in their origins and spatiotemporal scales of their occurrences (Hao *et al.* 2018).

Nevertheless, the multifold socioeconomic as well as hydroclimatological impacts of drought have led to the development of several drought forecasting models such as regression analysis, auto-regressive integrated moving average, Markov chain, Artificial Neural Network (ANN) (Mishra & Desai 2006; Mishra *et al.* 2007; Morid *et al.* 2007; Le *et al.* 2017; Ghorbani *et al.* 2018; Khan *et al.* 2020), Support Vector Regression (SVR) (Maity *et al.* 2010; Belayneh *et al.* 2014; Ganguli & Reddy 2014; Khan *et al.* 2020; Pal *et al.* 2020), random forest, regressions trees (Feng *et al.* 2019; Granata 2019), extreme learning machine (Deo & Sahin 2015; Mouatadid *et al.* 2018), *k*-nearest neighbor (Khan *et al.* 2020), gene expression programming and tree model (Shamshirband *et al.* 2020). The outcomes from these large varieties of drought forecasting models indicate that the Machine Learning (ML)/artificial intelligence approaches are gradually replacing the linear approaches, as they cannot ideally capture the nonlinearity component in the time series (Fung *et al.* 2020). The ML algorithms can predict the drought events that do not have a good, straightforward mathematical solution capturing the white noise, nonstationarity and nonlinearity in the time series (Fung *et al.* 2020). These ML algorithms are also able to capture the complex interactions involved in natural phenomena owing to their ability to automatically learn from the observed data (Lantz 2015; Sachindra & Kanae 2019).

However, the development of these models primarily requires multiple trials to determine weights and biases, run into restrictions while handling large data and consider the drought-causing factors to be limited. Moreover, these ML-based models suffer from overfitting in calibration, underfitting in validation and trapping at local minima, especially in the case of ANN and support vector machine (Khan *et al.* 2020). These lead to a pressing need for more advanced methods for the hydrometeorological analysis of complex processes like extreme events including droughts. Although in some studies, ANN is used for complex time-series modeling with a sufficient number of hidden layers and a specified number of units (Zhang *et al.* 1998), yet the complex hydrometeorological phenomena, such as droughts, summon more hidden layers leading to the problem of non-convex optimization. The issue can be dealt with Deep Learning (DL) algorithms with an unsupervised greedy layer-wise training for Deep Neural Networks (DNNs) (Hinton & Salakhutdinov 2006), which perform better than conventional ML approaches by avoiding getting stuck in the wrong local solutions (Lecun *et al.* 2015). The DL approach is able to extract the data features from raw data using multiple hierarchical layers (Khan & Maity 2020). The ability to learn from exposure to data without any human expertise enables it to effectively study the nonlinear, complex and hidden information involved in hydroclimatological processes and thus helps to develop the models at various spatiotemporal scales (Khan & Maity 2020).

A few studies have shown that DL-based approaches can extract more useful features from highly nonlinear physical processes involved in drought during a comprehensive drought model development. A Deep Belief Network (DBN) was used for short-term drought prediction in the Huaihe River Basin of China and was found to be superior to the standard backpropagation neural network (Chen *et al.* 2012). Another DBN algorithm-based approach has been implemented for the prediction of long-term (with 6- and 12-month lead times) drought conditions using the Standardized Streamflow Index in the Upper Colorado River Basin (Agana & Homaifar 2017). The results indicate the better prediction efficiency of DL-based algorithms compared with multilayer perceptron and SVR. Shen *et al.* (2019) used a DL-based model to construct a comprehensive drought monitoring model in the Henan Province of China considering the various hazard factors in drought development. In another study, Kaur & Sood (2020) used the DNN to predict the level of drought severity for different climate blocks and different time frames with various drought indices. The DNN also outperformed the ANN with genetic algorithm and ANN in assessing the drought conditions. In another study, Xu & Mo (2020) also used DBN to predict the Standardized Precipitation Index at different time scales. The DBN was shown to perform better as compared with other traditional methods. However, none of the aforementioned studies attempt to extract the information from a set of hydrometeorological precursors for drought assessment. This is important and beneficial because the performance of any climate model is generally highly skillful for the primary meteorological variables as compared with other

secondary or tertiary variables/phenomena. Thus, drought being a tertiary phenomenon, the ability of any climate model to perform drought assessment is very limited. On the other hand, while the efficient applications of DL-based approaches have been proved, its potential for drought assessment using the information from hydrometeorological precursors has yet to be measured. Herein lies the motivation of this study. Moreover, there are practically no such studies available in any Indian basins to the authors' best knowledge.

The objective of the study is to develop a DL-based model for drought assessment utilizing a set of hydrometeorological precursors. We picked out two medium-sized, rainfed river basins that are located in the central belt of India and frequently stricken by droughts due to the vagaries of precipitation. Further details about the study basins are provided in the 'Study area and data' section. Following that, the 'Methodology' section provides details on the methodology including drought characterization, data preparation and the proposed DL approach. Results and related discussions are presented in the 'Results and discussion' section along with more specific details of the DL model. The performance of the DL-based model is also compared with another popular ML-based approach, i.e., SVR, to explore the additional benefits against an existing approach keeping all other conditions the same. Finally, the major findings and conclusions are drawn in the 'Conclusions' section.

STUDY AREA AND DATA

Two river basins lying in different climatic zones of India with comparable catchment area, namely Damodar River Basin up to Tenughat Dam in eastern India (TRB) and Wardha River Basin up to upper Wardha Dam in western India (WRB), are considered for the study (Figure 1). TRB is a part of the Damodar River Basin with a catchment area of 4,939 km^2. The Damodar River rises in the Palamau hills of Chotanagpur Plateau at an elevation of about 609.75 m. The river flows across two Indian states, namely Jharkhand and West Bengal, and the dam is located at Tenughat in Petarwar block of Bokaro district in Jharkhand.

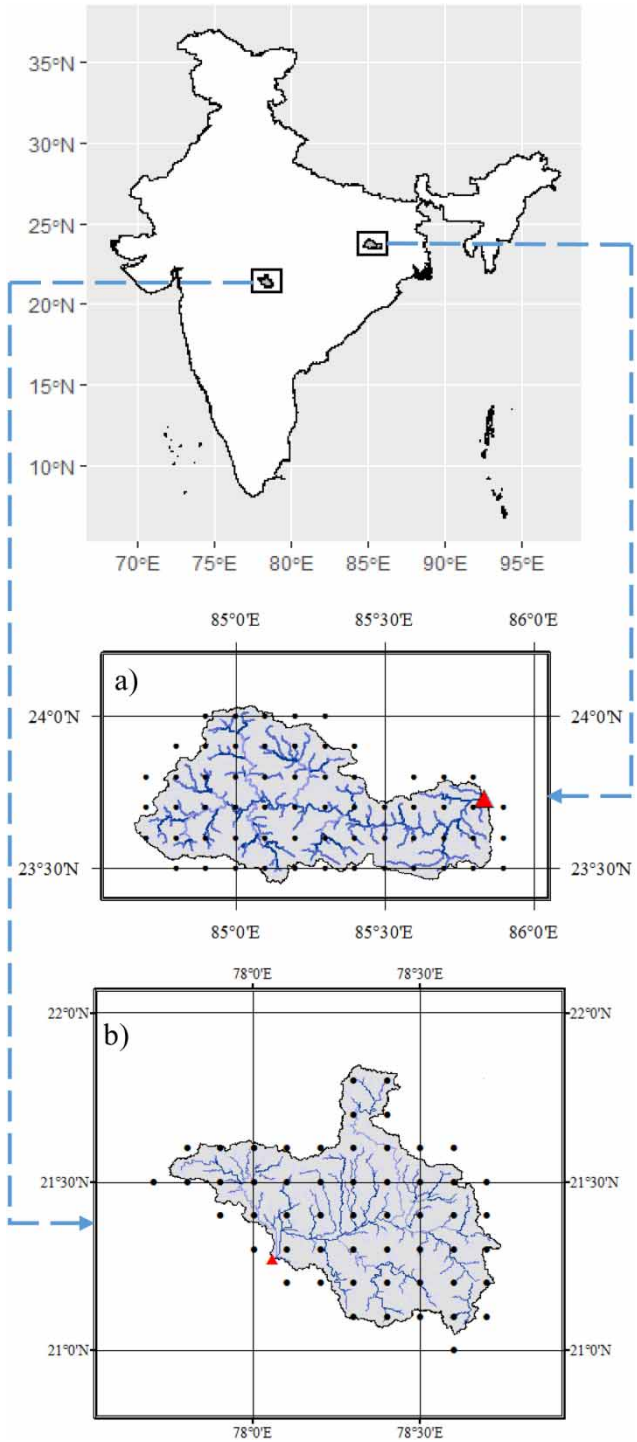


Figure 1 | Location of study basins with approximate drainage network: (a) Damodar River Basin up to Tenughat Dam (TRB) and (b) Wardha River Basin up to Upper Wardha Dam (WRB). The basin outlets are shown as red triangles and the black dots designate the grid points (0.1° × 0.1°) utilized for analysis. Please refer to the online version of this paper to see this figure in colour: doi:10.2166/wcc.2021.062.

The catchment comprises rough hilly areas denuded of forest and vegetal cover and is subjected to erosion. WRB is a part of the Godavari River Basin with a catchment area of 4,326 km^2. The Wardha River originates at Satpura Range near Khairwani village in Multai Tehsil, Betul District, Madhya Pradesh at an altitude of 777 m. The Upper Wardha Dam is built across the Wardha River, and the catchment area is hilly and forested. The path of monsoon depressions that originate in the Bay of Bengal descends directly on this catchment area. Both these study basins are located in the central belt of Indian mainland that is prone to droughts due to the vagaries of rainfall with some variation between eastern and western sides. It is hypothesized that the vagaries of rainfall are forced by several hydrometeorological factors in an unknown and complex way that may vary over space. Thus, such basins provide a unique scope to explore the potential of DL-based approaches.

All the datasets are obtained from the fifth generation of the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis product (ERA5, https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5, accessed April 2021) for the period of 1981–2020. ERA5 provides high-resolution estimates for a large number of atmospheric, land and oceanic variables at various temporal scales (sub-daily to monthly). It combines a vast amount of historical observations into global estimates using advanced modeling and data assimilation systems. The precipitation data are used for the evaluation of the drought index, the details of which are provided in the 'Methodology' section. The set of hydrometeorological variables includes air temperature (2 m height), surface pressure, wind speed (resultant of zonal and meridional wind at 10 m height), relative humidity, evaporation, surface soil moisture (0–7 cm) and geopotential height. Further details of the datasets as obtained from ERA5 are provided in Table 1. All the datasets are spatially averaged across the basin and converted to a monthly scale before further processing. The grid points lying within and in the proximity of the basin, as shown in Figure 1, are considered.

METHODOLOGY

A flowchart summarizing the complete methodological concept along with a schematic diagram showing the proposed DL model architecture using a one-dimensional Convolutional Neural Network (CNN), henceforth referred to as the Conv1D, is shown in Figure 2. Different components of the methodology are explained in the following subsections.

Drought characterization for Indian hydroclimatology

Standardized Precipitation Anomaly Index (SPAI) is a generalized anomaly-based index for characterizing meteorological drought in monsoon-dominated climatology (Chanda & Maity 2015; Makokha *et al.* 2016; Das & Chanda 2020; Maity *et al.* 2020; Monish & Rehana 2020). As the rainfall in India is strongly seasonal due to its monsoon-dominated climatology, the SPAI is used as the drought characterization index. For a monthly scale

Table 1 | Details of the meteorological dataset obtained from ERA5

Dataset	Variables	Spatial Resolution	Vertical/Pressure Level	Units (ERA5)	Units (Converted)
ERA5-Land monthly averaged data	Total precipitation	$0.1^\circ \times 0.1^\circ$	Surface	m	mm
	2-m temperature	$0.1^\circ \times 0.1^\circ$	2-m above surface	K	°C
	Pressure	$0.1^\circ \times 0.1^\circ$	Surface	Pa	kPa
	Evaporation	$0.1^\circ \times 0.1^\circ$	Surface	m (of water equivalent)	mm
	Volumetric soil water	$0.1^\circ \times 0.1^\circ$	Sub-surface (0–7 cm)	m^3/m^3	–
	10-m u-wind	$0.1^\circ \times 0.1^\circ$	10-m above surface	m/s	–
	10-m v-wind	$0.1^\circ \times 0.1^\circ$	10-m above surface	m/s	–
ERA5 monthly averaged data on pressure levels	Geopotential height	$0.25^\circ \times 0.25^\circ$	1,000 hPa	m^2/s^2 (geopotential)	m (by dividing geopotential by gravitational constant)
	Relative humidity	$0.25^\circ \times 0.25^\circ$	1,000 hPa	Percent (%)	–

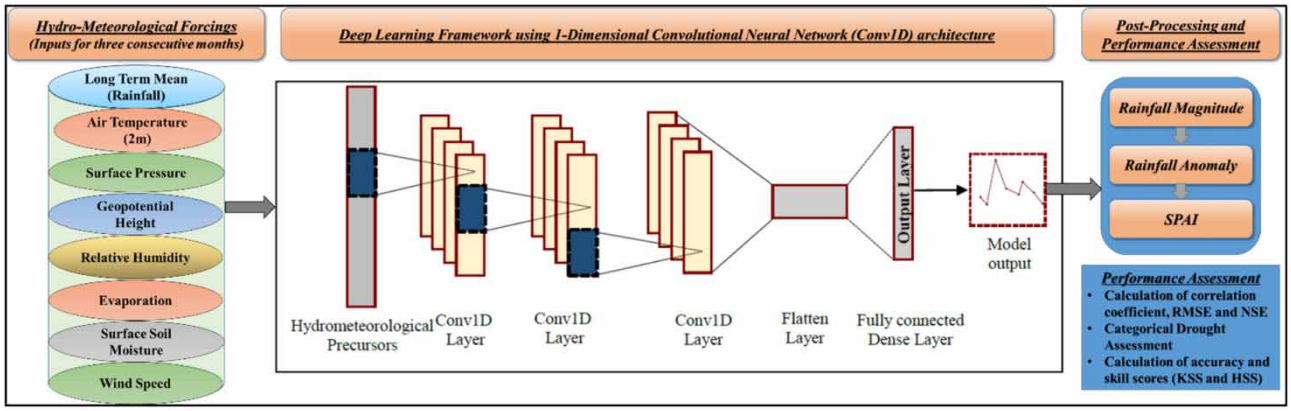


Figure 2 | Methodological flowchart along with a schematic diagram showing different layers of Conv1D model architecture. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/wcc.2021.062.

analysis, the SPAI should be computed in three steps. First, the monthly precipitation anomalies are calculated from the raw precipitation values. The monthly anomalies are obtained as follows:

$$y_{i,j} = (x_{i,j} - \bar{x}_j) \tag{1}$$

where $y_{i,j}$ is the precipitation anomaly for the ith year and jth time step of the year, $x_{i,j}$ is the precipitation value for the ith year and jth time step of the year and $\bar{x}_j$ is the long-term mean precipitation for the jth time step of the year. Next, a probability distribution (parametric or nonparametric) is fitted across the monthly anomaly series to obtain the probability quantiles. If a nonparametric distribution is considered, then the empirical Cumulative Distribution Function (CDF) may be obtained by using the Weibull plotting position formula as follows:

$$p = \frac{m}{N+1} \tag{2}$$

where p is the cumulative probability, m is the rank of the dataset arranged in descending order and N is the sample size. Finally, the quantiles obtained in the previous step are transformed to standard normal variates (z-score), which give the SPAI. Thus, the SPAI is obtained as follows:

$$z = F^{-1}(p) \tag{3}$$

where $F^{-1}(p)$ represents the inverse of the standard normal distribution for a CDF value of p.

It may be mentioned here that there are several alternative indices for the characterization of droughts, such as Palmer Drought Severity Index (PDSI) (Palmer 1965), Standard Precipitation Index (SPI) (McKee *et al.* 1993), Standardized Nonstationary Precipitation Index (SnsPI) (Russo *et al.* 2013), Joint Deficit Index (JDI) (Kao & Govindaraju 2010) and Copula-based Joint Drought Index (CJDI) (Won *et al.* 2020). Many of these drought indices (PDSI, JDI and CJDI) require multiple meteorological inputs for computation. However, SPI, SnSPI and SPAI are calculated from precipitation data only. Even among these, the SPAI is established to be more suitable than the SPI or SnsPI for a periodic precipitation series such as that observed in India (Chanda & Maity 2015). The SPAI is able to distinguish (statistically) similar inter-seasonal deficits, which have contrasting socioeconomic implications in a monsoon-dominated climatology (Chanda & Maity 2015; Park *et al.* 2019). As the SPAI is found to be a robust meteorological drought index, it is adopted in several recent studies (Amrit *et al.* 2018; Das & Chanda 2020; Rehana & Monish 2020) and also selected in the present study.

Next, the categorical classification of droughts is done based on the different ranges of SPAI values. These are as follows: if the SPAI value ranges between 1 and −1, i.e., $-1 \leq \text{SPAI} \leq 1$, it is considered as a 'near-normal' (N) situation. It is characterized as 'Moderate Drought (D0)' and 'Moderately Wet (W0)' conditions for $-1.5 \leq \text{SPAI} < -1$ and $1 < \text{SPAI} \leq 1.5$, respectively. Likewise, for $-2 \leq \text{SPAI} < -1.5$ and

$1.5 < \text{SPAI} \leq 2$, it is categorized as 'Severe Drought (D1)' and 'Severely Wet (W1)' conditions, respectively. Finally, 'Extreme Drought (D2)' and 'Extremely Wet (W2)' conditions are denoted by $\text{SPAI} < -2$ and $\text{SPAI} > 2$, respectively.

Data preparation

Entire data preparation and handling are carried out in the scientific python development environment (Spyder) notebook. It starts with spatially averaging the gridded datasets across the basins following the area weightage method and converting to a monthly scale honoring their units (Table 1).

In the proposed DL model, the values of the seven hydrometeorological precursors are used from three consecutive months to simulate the rainfall values for those 3 months, and subsequently, SPAI values are computed with the simulated values of rainfall. The analysis also includes the use of the k-fold cross-validation technique to evaluate the simulation/prediction skill of the model. Accordingly, k (here $k = 5$) approximately equal folds (division of the available data length) are considered. This leads to the consideration of approximately 80% of the data as the training set and the remaining 20% as the testing set. The model is trained on the four $(k - 1)$ folds, and its performance is checked on the remaining fold. This is repeated for all the folds.

One-dimensional CNN

The proposed DL model is based on the Conv1D model. It is developed in the Spyder notebook using Keras library built on the top of TensorFlow. The CNN architecture generally comprises an input layer, an output layer and between them, there is some arbitrary number of hidden layers. The proposed DL model is of a sequential type that can be used for simulation/prediction problems using one or multiple time series. A schematic diagram is shown in Figure 2.

The function of the input layer is to receive the signal (input data) and transfer it to the hidden layer(s). Hidden layers are the computational engine of the model. They may have one or more Conv1D layer, max-pooling layer, dropout layer and a flatten layer depending on the problem. The Conv1D layer is the main building block of CNN. It consists of one-dimensional filters/kernels to extract features from the input signal, kernel size to specify the length of the filter and an activation function to set a threshold limit for the neurons. The model is trained on the defined causal and target variables and extracts the hidden information of the sequence. A max-pooling layer (if used) is generally used after the convolutional layer. It helps in reducing the dimension of the output matrix and also the chance of overtraining. The function of the dropout layer (if used) is to randomly assign zero weights to the neurons of the network, which makes it less sensitive toward smaller variation, thus improving the accuracy of the model on unseen data (Srivastava *et al.* 2014). A flatten layer converts the output of the convolutional/pooling/dropout layers to one dimension and transfers the data to the output layer. The output layer is a fully connected dense layer, which is connected to all the neurons of the previous layer and is responsible for generating the output of the model. There are several hyper-parameters that needed to be specified to fix the model architecture. This is problem-specific, and details are provided in the 'Results and discussion' section. After configuring the layers, they are generally trained on a set of input and target data to learn the association involved between them (Zou *et al.* 2008). Training involves the tuning of several hyper-parameters to minimize the error (loss). The error is measured in several ways, namely Mean Squared Error (MSE), Mean Absolute Error (MAE) and log loss. Once the model is trained, it is used to generate output from the unseen/testing data. A more detailed background on the one-dimensional convolutional network can be found in Kiranyaz *et al.* (2021).

Model parameters/configuration

Several combinations of layers were evaluated with different sets of model hyper-parameters (such as filter size, number of filters, dropout percentage and number of hidden layers) and optimization hyper-parameters (such as batch size, activation function, number of epochs, learning rate, loss function, momentum and decay rate) so as to ascertain the best possible model configuration. The finalized architecture of the Conv1D model comprises six layers (Figure 2), whose configurations are discussed below.

The prepared dataset of seven precursors (also known as input features) along with long-term mean rainfall (hence, a total of eight features) is fed as input to the first Conv1D

layer. The input shape consists of three time-steps (months) along with their respective values of the eight features arranged in two-dimensional matrix form. The first convolutional layer comprises 124 filters, filter/kernel size 1, stride 1, Glorot uniform kernel initializer (aka Xavier uniform initializer) and Rectified Linear Unit (ReLU) activation function. The convolutional filters help to identify and excerpt the hidden information in the input data. After achieving the threshold value assigned by the ReLU activation function, the output of the layer moves to the next unit of the model. Subsequently, two more convolutional layers are added to the model having the same configuration as that of the first Conv1D layer. After providing three Conv1D layers, a flatten layer is added to the model to reshape the multidimensional input received from the previous Conv1D layer to one dimension. Followed by the flatten layer, the output layer is added comprising three neurons and a linear activation function. These three neurons are meant for the simulated rainfall values for three months.

Based on the above architectural configuration and by several trials, a batch size of 375 and 200 epochs (without shuffling) were selected for training the model. The model architecture uses the MAE to calculate the loss and an efficient Adam version of stochastic gradient descent with a learning rate of 0.0001, momentum rate of 0.9 and decay rate of 1×10^{-7}, which are adopted for the best possible configurations (Kingma & Ba 2014). After successful completion of training, the performance of the model was assessed on the testing dataset. The model configuration is kept unchanged across different folds during k-fold cross-validation. To ensure neither overfitting nor underfitting, it is recommended to observe that the loss (MAE) obtained from the training and testing data are more or less the same.

Comparison with other approaches

The proposed DL-based model is compared with another popular ML approach, i.e., SVR, which is recurrently used in the field of hydrology (Maity *et al.* 2010; Achieng 2019; Qasem *et al.* 2019; Maity *et al.* 2020; Pal *et al.* 2020). The SVR model is developed using the scikit-learn library in the spyder notebook with the same proportion of training and testing dataset, and using 5-fold cross-validation, the same as in the case of the DL-based model. The SVR is a supervised ML algorithm, which uses a nonlinear mapping function to map the input features to the target variable. In this study, two regularization parameters, Gamma (γ) and cost function (C) of Radial Basis Function (RBF) are tuned by trial and error to optimize the SVR model ($\gamma = 1 \times 10^{-5}$ and $C = 1{,}500$). The RBF is a nonlinear optimization function, which is commonly used in SVR (Choy & Chan 2003). More details on SVR can be found in the literature (Drucker *et al.* 1997; Choy & Chan 2003; Maity *et al.* 2010).

Performance evaluation for drought assessment

Once the simulated rainfall values are obtained from the DL model, the SPAI values are computed outside the DL domain, followed by categorical classification of droughts according to the method explained earlier in the 'Drought characterization for Indian hydroclimatology' section. The performance of the proposed DL model in simulating SPAI values is assessed through three standard statistical performance metrics, namely Correlation Coefficient (CC) (Pearson & Henrici 1896), Root-Mean-Square Error (RMSE) (Chai & Draxler 2014) and Nash–Sutcliffe Efficiency (NSE) (Nash & Sutcliffe 1970). If X_O = observed value, X_P = predicted/simulated value, σ_O = standard deviation of observed value, N = number of data points, $\overline{X_O} = \frac{1}{N}\sum_{i=1}^{N} X_O$ and $\overline{X_P} = \frac{1}{N}\sum_{i=1}^{N} X_P$, then CC, RMSE and NSE may be determined as follows:

$$\mathrm{CC} = \frac{\sum_{i=1}^{N}[(X_O - \overline{X_O})(X_P - \overline{X_P})]}{\sqrt{\sum_{i=1}^{N}(X_O - \overline{X_O})^2 \sum_{i=1}^{N}(X_P - \overline{X_P})^2}} \tag{4}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_O - X_P)^2} \tag{5}$$

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}(X_P - X_O)^2}{\sum_{i=1}^{N}(X_O - \overline{X_O})^2} \tag{6}$$

The simulation skill is determined through k-fold (here, $k = 5$) cross-validation, and the average values of aforementioned metrics across all k-folds are obtained for training and testing periods.

Next, the performance of the DL model in simulating the drought category is verified by compiling a typical $c \times c$ contingency table (here, $c = 7$), where the frequencies of observations and simulations are binned in relevant cells as illustrated in Figure 3. In this table, n_{ij} denotes the number of observations in category i that have been simulated as category j, NO_i denotes the total number of observations (marginal) in category i, NS_j denotes the total number of simulations (marginal) in category j and N is the total number of observations/simulations. Ideally, a perfect forecasting system would only have entries along the diagonal of the contingency table, with all of the other cells being empty (=0). To quantify the performance of the model in the case of such multi-category simulations, three standard statistical measures are used, namely accuracy, Kuipers Skill Score (KSS) and Heidke Skill Score (HSS). The relevant equations of these three performance measures are shown in Figure 3 in reference to the typical $c \times c$ contingency table. The accuracy indicates the overall fraction of the simulations in the 'correct' category. Its value ranges between 0 and 1, where 1 indicates the perfect score. On the other hand, the skill represents the accuracy of a model performance against a base model performance that is random or simple to produce or already be available to users. The KSS and HSS both measure the fraction of correct simulations after eliminating those simulations, which would be correct purely due to random chance. The estimation of randomness (expressed by the denominator) is different in these two skill scores. Theoretically, KSS may range between −1 and 1, where 0 indicates no skill and the perfect score is 1. Similarly, the theoretical range of HSS is (−∞, 1], where 0 indicates no skill and the perfect score is 1.

RESULTS AND DISCUSSION

At the outset, the SPAI values obtained from the spatially averaged ERA5 reanalysis precipitation data are compared with that of the observed precipitation data (Pai *et al.* 2014) over the two study basins. The scatter plots are shown in Figure 4 for both the study basins. The CC is found to be 0.71 for TRB and 0.68 for WRB. Given the limitations of any reanalysis product, such correspondence with the observed data can be considered as good. The proposed DL-based model is trained and tested based on its efficacy to capture the rainfall variation using the hydrometeorological precursors. It was targeted to consider as many months as possible simultaneously so as to avoid running the model for each month separately and at the same time not to be burdened by an increased computational effort. We found that the model is able to consider three consecutive months with comparable accuracy for all the months and at a reasonable computational effort. Henceforth, in general,

Typical Contingency table

Observed	Simulated				
	S_1	S_2	…	S_c	Sum
O_1	n_{11}	n_{21}	…	n_{c1}	NO_1
O_2	n_{12}	n_{22}	…	n_{c2}	NO_2
:	:	:	…	:	:
O_c	n_{1c}	n_{2c}	…	n_{cc}	NO_c
Sum	NS_1	NS_2	…	NS_c	N

Mathematical Formulations

$$Accuracy = \frac{1}{N}\sum_{i=1}^{k} n_{kk}$$

$$HSS = \frac{\frac{1}{N}\sum_{i=1}^{k} n_{kk} - \frac{1}{N^2}\sum_{i=1}^{k} NO_k NS_k}{1 - \frac{1}{N^2}\sum_{i=1}^{k} NO_k NS_k}$$

$$KSS = \frac{\frac{1}{N}\sum_{i=1}^{k} n_{kk} - \frac{1}{N^2}\sum_{i=1}^{k} NO_k NS_k}{1 - \frac{1}{N^2}\sum_{i=1}^{k} NO_k{}^2}$$

Figure 3 | A typical $c \times c$ contingency table and the mathematical formulations of three performance metrics, namely accuracy, KSS and HSS. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/wcc.2021.062.

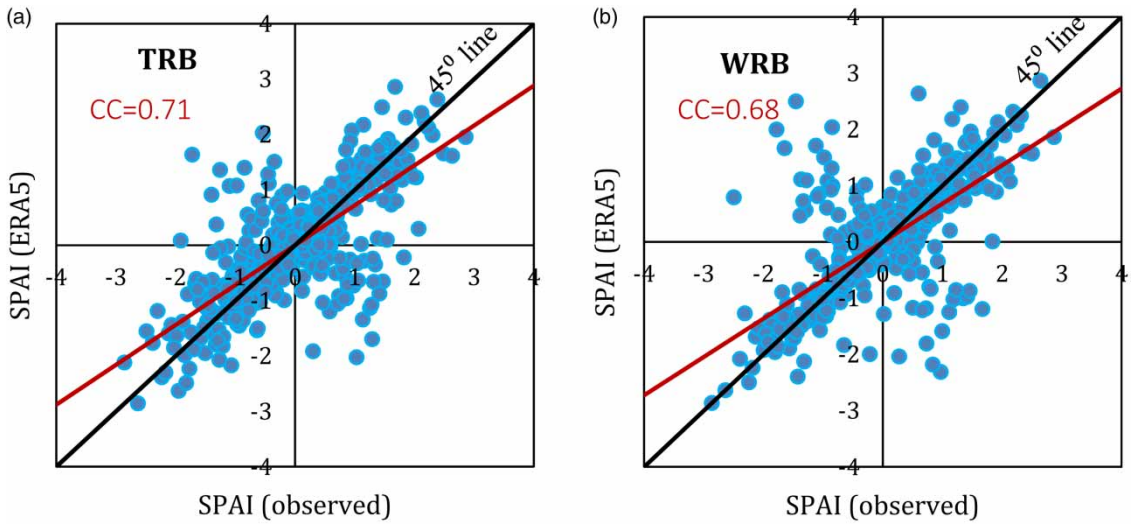


Figure 4 | Scatter plots between SPAI values calculated from observed precipitation data and that from ERA5 reanalysis precipitation data for (a) TRB and (b) WRB study basins. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/wcc.2021.062.

these 3 months are designated as 1st, 2nd and 3rd months in the following discussion.

Model performance in capturing the drought status

To assess the potential of the proposed DL-based model to capture the drought status in terms of SPAI, the observed and simulated rainfall values are converted to SPAI indices. Figure 5 presents the comparisons between the observed and simulated values of SPAI through time-series (left) and scatter plots (right) for TRB. The same for WRB is presented in Figure 6. In these figures, only the cases of the 1st months are shown, and the testing periods (all 5-fold cases) are shown separately within the continuous time series for an easy comparison. A visual inspection reveals that the predictions are promising in terms of direction, i.e., below- or above-normal (D or W, as per the categorization explained before), as well as the magnitude of SPAI. This is true for all the folds for both the study basins. More specifically, in the case of TRB, the performances of the model for all the folds, except the 1st fold, are good, as all the peaks are captured almost accurately both in training and testing periods. In the case of WRB, a comparable performance is noticed for all 5 folds. This indicates the reliable potential of the DL-based model in the basin-scale drought assessment using information from a set of primary hydrometeorological precursors. Comparable performances are noticed for the 2nd and 3rd months as well. Though the figures are not included owing to space limitations, the performances of the simulations for all three consecutive months are quantified through aforementioned performance metrics, i.e., CC, RMSE and NSE. These values for all the folds and for both the basins are summarized in Tables 2 and 3. A graphical presentation is also shown in Figure 7. The average performance of the DL-based model during the training period ranges between 0.78–0.84 (CC), 0.58–0.65 (RMSE) and 0.58–0.67 (NSE), and that for the testing period ranges between 0.70–0.77 (CC), 0.58–0.70 (RMSE) and 0.43–0.56 (NSE). Comparable performances during the training and testing periods indicate the proper training of the model without overtraining or undertraining. A comparison between two study basins indicates a marginally better performance in the case of TRB as compared with WRB. Moreover, the performances are comparable for all three consecutive months that indicate the ability of the model to simulate drought status for all 3 months simultaneously.

As mentioned before, the DL-based model performance is also compared against another popularly used ML approach, i.e., SVR, keeping all other conditions the same. The results are presented in the same tables and figure, i.e., Tables 2 and 3 and Figure 7, for an easy side-by-side comparison. The higher potential of the DL-based model is clearly established for all the cases. It is true for both the study basins. However, one point is to be mentioned

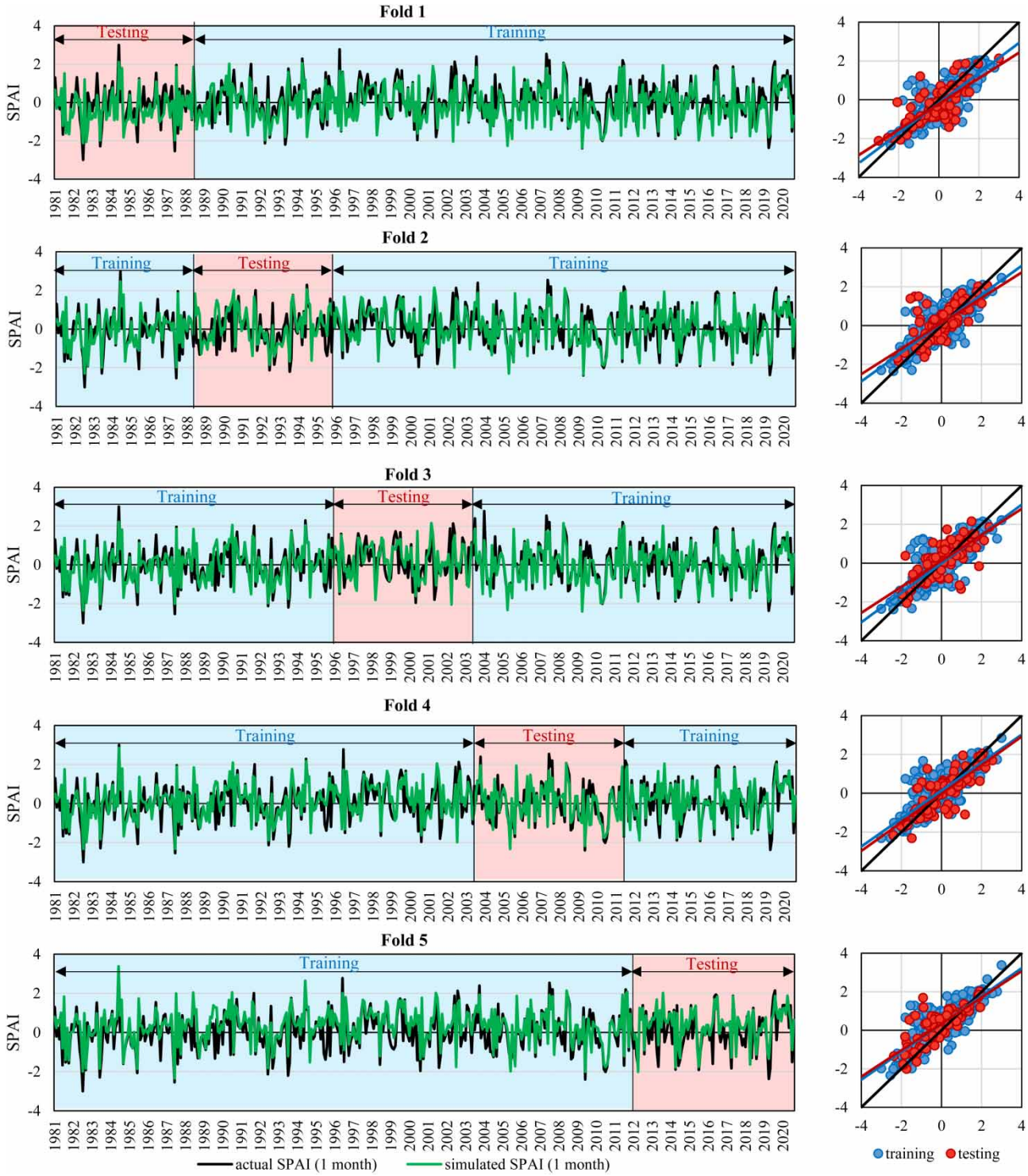


Figure 5 | Model performance in drought identification for TRB considering Conv1D model. Observed and simulated (1st month) SPAI values are shown through time-series (left) and scatter plots (right) for all 5 folds. Training and testing periods are shown in the time-series plots for different folds. In the scatter plots, the *x*-axis and *y*-axis represent the observed and simulated SPAI values, respectively. The solid black lines show the 45° line (line of perfect simulation), and the other two lines show best-fit lines for the scatter plots. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/wcc.2021.062.

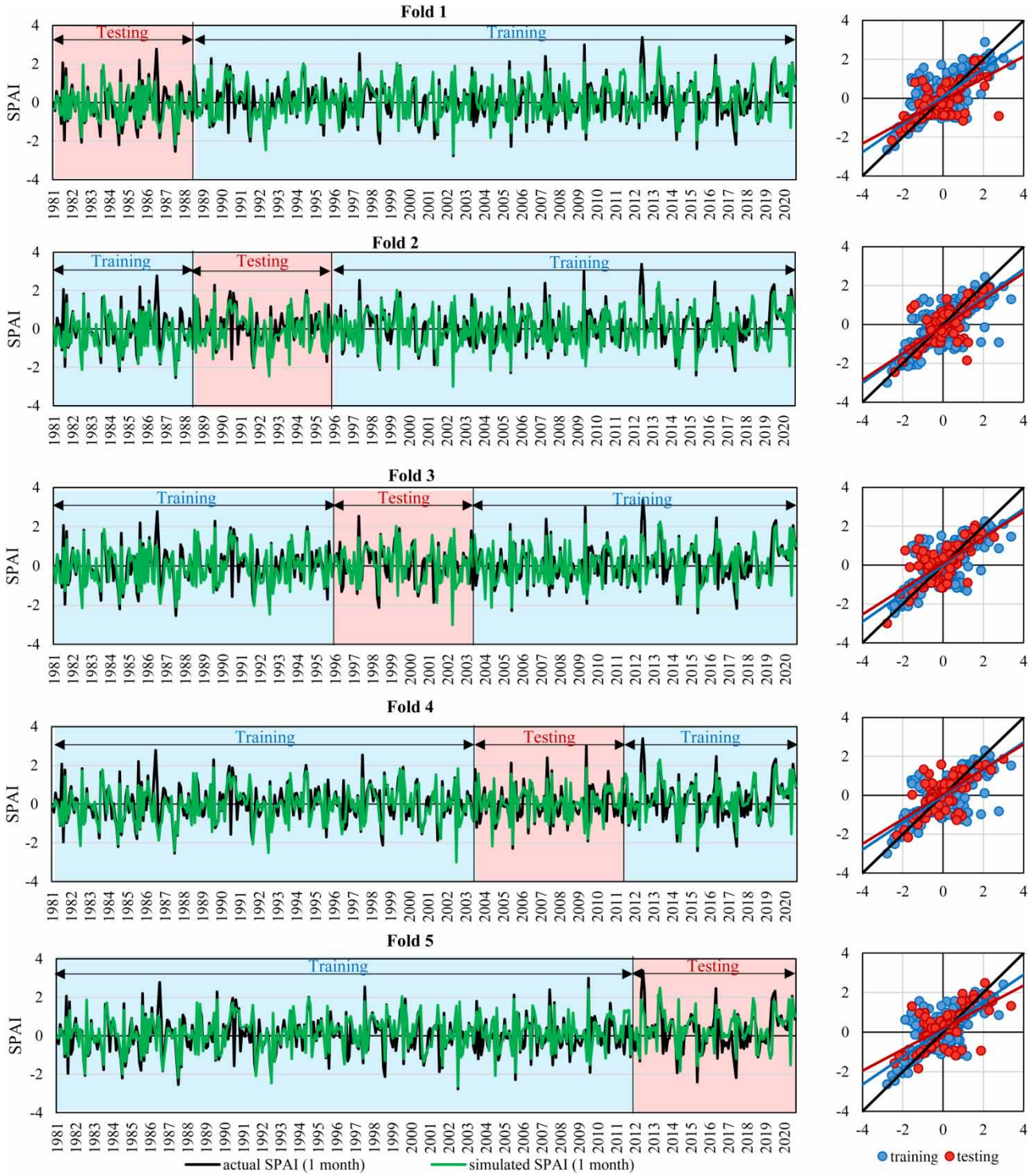


Figure 6 | Same as Figure 5, but for the WRB study basin. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/wcc.2021.062.

Table 2 | Fold-wise comparison of performance metrics between observed and simulated SPAI during training (trn) and testing (tst) periods obtained by DL and SVR models in 1st, 2nd and 3rd months for the TRB study basin

Study basin: TRB

		Correlation coefficient						RMSE						NSE					
		1st month		2nd month		3rd month		1st month		2nd month		3rd month		1st month		2nd month		3rd month	
Fold	Model	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst
Fold 1	DL	0.84	0.74	0.85	0.76	0.84	0.79	0.59	0.72	0.54	0.67	0.55	0.63	0.66	0.50	0.71	0.55	0.70	0.61
	SVR	0.78	0.70	0.79	0.66	0.77	0.68	0.63	0.82	0.62	0.86	0.64	0.85	0.60	0.49	0.62	0.43	0.60	0.45
Fold 2	DL	0.83	0.67	0.84	0.69	0.84	0.71	0.59	0.76	0.59	0.75	0.58	0.73	0.67	0.34	0.67	0.35	0.68	0.43
	SVR	0.77	0.66	0.76	0.70	0.77	0.65	0.67	0.72	0.65	0.63	0.67	0.74	0.59	0.41	0.62	0.55	0.59	0.42
Fold 3	DL	0.83	0.72	0.84	0.71	0.84	0.68	0.58	0.70	0.57	0.71	0.57	0.74	0.68	0.46	0.69	0.43	0.69	0.38
	SVR	0.78	0.62	0.80	0.69	0.79	0.66	0.66	0.76	0.63	0.68	0.65	0.71	0.61	0.37	0.64	0.47	0.62	0.43
Fold 4	DL	0.81	0.80	0.83	0.79	0.83	0.76	0.60	0.65	0.60	0.67	0.58	0.72	0.63	0.63	0.64	0.60	0.66	0.55
	SVR	0.78	0.73	0.79	0.71	0.78	0.76	0.65	0.73	0.64	0.75	0.65	0.69	0.60	0.53	0.61	0.51	0.61	0.58
Fold 5	DL	0.83	0.81	0.84	0.81	0.82	0.78	0.65	0.69	0.61	0.63	0.61	0.67	0.58	0.56	0.63	0.62	0.64	0.58
	SVR	0.77	0.71	0.78	0.75	0.79	0.76	0.66	0.72	0.65	0.68	0.65	0.67	0.60	0.51	0.61	0.56	0.61	0.57
Average	DL	**0.83**	**0.75**	**0.84**	**0.75**	**0.83**	**0.74**	**0.60**	**0.70**	**0.58**	**0.69**	**0.58**	**0.70**	**0.65**	**0.50**	**0.67**	**0.51**	**0.67**	**0.51**
	SVR	**0.78**	**0.68**	**0.78**	**0.70**	**0.78**	**0.70**	**0.66**	**0.75**	**0.64**	**0.72**	**0.65**	**0.73**	**0.60**	**0.46**	**0.62**	**0.51**	**0.61**	**0.49**

Table 3 | Fold-wise comparison of performance metrics between observed and simulated SPAI during training (trn) and testing (tst) periods obtained by DL and SVR models in 1st, 2nd and 3rd months for the WRB study basin

Study basin: WRB

		Correlation coefficient						RMSE						NSE					
		1st month		2nd month		3rd month		1st month		2nd month		3rd month		1st month		2nd month		3rd month	
Fold	Model	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst
Fold 1	DL	0.76	0.69	0.81	0.78	0.81	0.76	0.66	0.79	0.59	0.69	0.59	0.71	0.55	0.47	0.64	0.60	0.64	0.58
	SVR	0.73	0.60	0.76	0.64	0.76	0.56	0.68	0.86	0.65	0.84	0.64	0.91	0.52	0.36	0.57	0.41	0.57	0.32
Fold 2	DL	0.79	0.70	0.83	0.74	0.82	0.80	0.65	0.73	0.59	0.66	0.59	0.59	0.60	0.38	0.67	0.50	0.66	0.60
	SVR	0.72	0.66	0.73	0.71	0.75	0.77	0.71	0.70	0.70	0.66	0.69	0.61	0.52	0.43	0.53	0.50	0.55	0.58
Fold 3	DL	0.77	0.70	0.83	0.78	0.83	0.79	0.66	0.76	0.58	0.64	0.58	0.63	0.57	0.42	0.67	0.59	0.67	0.59
	SVR	0.66	0.63	0.69	0.61	0.66	0.55	0.76	0.79	0.73	0.82	0.76	0.85	0.44	0.37	0.48	0.32	0.43	0.26
Fold 4	DL	0.78	0.69	0.84	0.72	0.84	0.70	0.65	0.71	0.57	0.70	0.56	0.71	0.60	0.42	0.69	0.43	0.70	0.43
	SVR	0.71	0.65	0.71	0.70	0.70	0.63	0.73	0.71	0.72	0.69	0.74	0.73	0.50	0.42	0.50	0.46	0.48	0.38
Fold 5	DL	0.78	0.70	0.82	0.77	0.81	0.77	0.64	0.78	0.60	0.69	0.61	0.69	0.58	0.45	0.63	0.57	0.62	0.57
	SVR	0.70	0.58	0.73	0.61	0.71	0.54	0.70	0.86	0.68	0.84	0.70	0.90	0.49	0.33	0.53	0.37	0.50	0.28
Average	DL	**0.78**	**0.70**	**0.83**	**0.76**	**0.82**	**0.77**	**0.65**	**0.76**	**0.58**	**0.67**	**0.59**	**0.67**	**0.58**	**0.43**	**0.66**	**0.54**	**0.66**	**0.56**
	SVR	**0.70**	**0.63**	**0.72**	**0.65**	**0.72**	**0.61**	**0.72**	**0.79**	**0.70**	**0.77**	**0.71**	**0.80**	**0.50**	**0.38**	**0.52**	**0.41**	**0.51**	**0.36**

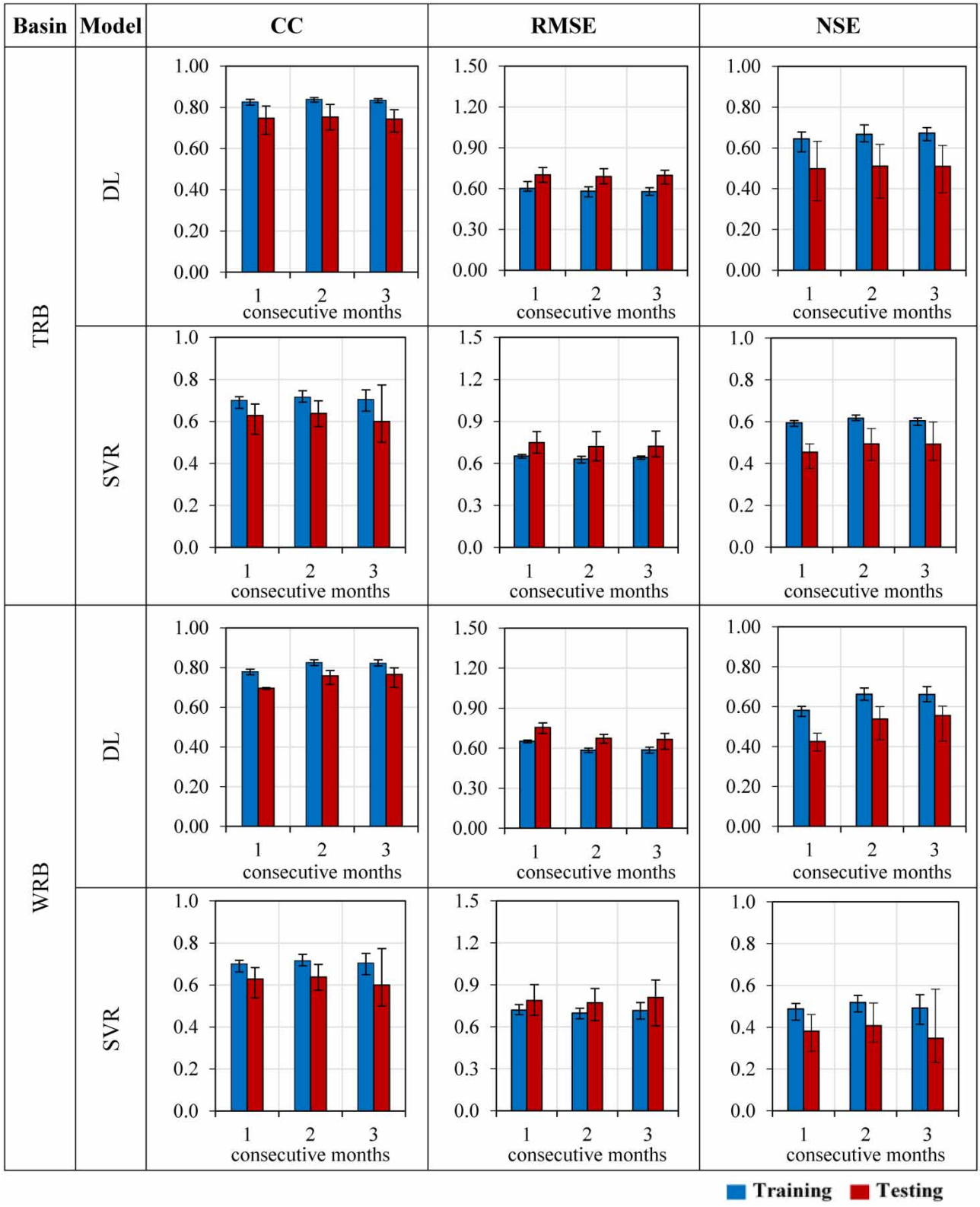


Figure 7 | Average CC, RMSE and NSE, along with range bars at the top, show the maximum to minimum value, across different folds for drought category identification for both the study basins using DL and SVR models for 1st, 2nd and 3rd months. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/wcc.2021.062.

here. SVR needs to be trained for each month separately, while the DL-based model is able to provide the output for all three consecutive months simultaneously. Thus, computational effort is less in the case of the DL-based approach. The effectiveness of the DL-based approach in analyzing and capturing the upcoming drought status with higher efficacy may be attributed to the convolutional feature of Conv1D, i.e., each layer contains a set of filters whose parameters need to be learned and the neurons are connected to the local region instead of being connected to all the neurons of the previous layer (Haidar & Verma 2018).

Model performance in drought category identification

While the overall prediction performance reflects the model ability over the entire range, it will be interesting to inspect the performances for different drought categories. Thus, the potential of the proposed DL model is also assessed by examining its skill to accurately simulate the category. Toward this, two-way (here, 7×7) contingency tables are prepared between observed and simulated drought categories. Referring to Figure 3, the number in the cell, n_{ij} in the contingency table, refers to the number of drought events falling in the ith observed and jth simulated categories. Such contingency tables are prepared for all the folds separately for both training and testing periods to calculate the performance metrics, i.e., accuracy, KSS and HSS, evaluated as per the mathematical formulations provided in Figure 3. These values for both the study basins and for all three consecutive months are reported in Table 4. For comparison, results from the SVR model are also shown in this table. The average values across the folds are also shown in this table in the last row.

As a first observation, the results indicate that the performance of both the models (DL and SVR) is much better than an unskilled random performance. In fact, the performances are very good, considering that the accuracy is almost always greater than 0.7 and KSS/HSS is much higher than random performance ($\gg 0$). The performances are more or less uniform across different folds and across two study basins. A graphical presentation of performances, summarizing all the folds, along with the range is shown in Figure 8 for three consecutive months and both the study basins, separately. Separate plots are also shown in the same figure for SVR models as well. Overall, it is noticed that the performances during training and testing periods are comparable with marginally better performance in training periods. However, a relative comparison clearly favors the DL-based model. In many cases, the improvement is as much as 100% in the case of the DL-based model as compared with the SVR model, particularly in cases of skill scores (KSS and HSS). This indicates a robust, consistent and improved performance of the proposed DL-based model. The performances are also comparable for all three consecutive months that reaffirm the ability of the model to simulate drought categories for all 3 months simultaneously. In a nutshell, the observation of the tables and figures, representing the skill scores, indicates the potential of the DL-based model in extracting the drought status from hydrometeorological precursors.

Next, all the testing period performances are considered to prepare the contingency tables. This helps to assess the model testing performance for the assessment of drought categories throughout the time period of analysis as a testing period only. However, this exercise is carried out only for the DL-based model, as the performance of SVR is already proved to be inferior. Secondly, the performance during the training period is kept aside while preparing the contingency tables, which will anyway be better than that during the testing period. Thus, the contingency tables only for the DL-based model and only for the performance during the testing period are shown in Tables 5 and 6 for TRB and WRB, respectively. In the case of TRB, almost every event is predicted accurately at least in the broad categories of dry and wet events. The total of nine extreme drought events occurred in the five testing folds. Out of these, two events are correctly identified as the extreme ones (D2), four events are identified as severe droughts (D1) and the rest are identified as moderate (D0) and near-normal (N) each. Overall, a heavy forward diagonal of these contingency tables indicates the correctness of drought category identification. In fact, 334 out of 470 (71%) categories are accurately simulated in the case of TRB, and 351 out of 470 (75%) categories are accurately simulated in the case of WRB. Performance metrics (accuracy, KSS and HSS) are also computed for these overall contingency tables. For TRB, accuracy, KSS and HSS are

Table 4 | Fold-wise comparison of performance measures between observed and simulated drought category during training (trn) and testing (tst) periods in 1st, 2nd and 3rd months for both the study basins (TRB and WRB) by Conv1D (bold font) and SVR models

		TRB						WRB					
		1st month		2nd month		3rd month		1st month		2nd month		3rd month	
Fold	Performance measures	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst	trn	tst
Fold 1	Accuracy	**0.72**	**0.76**	**0.76**	**0.74**	**0.75**	**0.75**	**0.78**	**0.77**	**0.78**	**0.69**	**0.77**	**0.66**
		0.72	0.71	0.74	0.74	0.72	0.74	0.75	0.67	0.75	0.66	0.75	0.66
	KSS	**0.42**	**0.49**	**0.49**	**0.37**	**0.48**	**0.46**	**0.50**	**0.46**	**0.53**	**0.32**	**0.51**	**0.31**
		0.34	0.20	0.37	0.27	0.33	0.31	0.21	0.12	0.24	0.10	0.21	0.09
	HSS	**0.45**	**0.48**	**0.53**	**0.39**	**0.51**	**0.46**	**0.52**	**0.52**	**0.54**	**0.36**	**0.52**	**0.34**
		0.40	0.24	0.44	0.31	0.40	0.35	0.28	0.17	0.32	0.14	0.29	0.13
Fold 2	Accuracy	**0.74**	**0.68**	**0.74**	**0.70**	**0.75**	**0.73**	**0.76**	**0.80**	**0.78**	**0.75**	**0.76**	**0.81**
		0.71	0.73	0.73	0.77	0.75	0.71	0.72	0.71	0.74	0.74	0.73	0.73
	KSS	**0.43**	**0.36**	**0.45**	**0.39**	**0.47**	**0.44**	**0.51**	**0.55**	**0.55**	**0.42**	**0.52**	**0.60**
		0.30	0.28	0.34	0.33	0.39	0.22	0.17	0.16	0.20	0.24	0.17	0.19
	HSS	**0.46**	**0.36**	**0.48**	**0.40**	**0.50**	**0.45**	**0.52**	**0.59**	**0.55**	**0.45**	**0.52**	**0.62**
		0.35	0.34	0.34	0.33	0.39	0.22	0.24	0.23	0.28	0.32	0.23	0.26
Fold 3	Accuracy	**0.72**	**0.69**	**0.76**	**0.70**	**0.76**	**0.66**	**0.78**	**0.73**	**0.80**	**0.76**	**0.79**	**0.80**
		0.71	0.67	0.73	0.69	0.72	0.70	0.73	0.73	0.74	0.74	0.74	0.71
	KSS	**0.41**	**0.40**	**0.48**	**0.44**	**0.49**	**0.29**	**0.55**	**0.38**	**0.59**	**0.47**	**0.60**	**0.61**
		0.30	0.16	0.34	0.21	0.32	0.18	0.24	0.20	0.28	0.22	0.27	0.12
	HSS	**0.43**	**0.41**	**0.52**	**0.44**	**0.52**	**0.30**	**0.56**	**0.42**	**0.60**	**0.50**	**0.59**	**0.61**
		0.35	0.21	0.40	0.27	0.39	0.24	0.31	0.28	0.35	0.30	0.34	0.18
Fold 4	Accuracy	**0.74**	**0.68**	**0.76**	**0.75**	**0.76**	**0.71**	**0.77**	**0.71**	**0.78**	**0.75**	**0.77**	**0.71**
		0.70	0.66	0.70	0.66	0.74	0.70	0.72	0.73	0.72	0.73	0.72	0.73
	KSS	**0.42**	**0.37**	**0.45**	**0.49**	**0.47**	**0.45**	**0.51**	**0.42**	**0.55**	**0.49**	**0.53**	**0.43**
		0.25	0.20	0.35	0.22	0.32	0.33	0.23	0.09	0.24	0.19	0.23	0.14
	HSS	**0.45**	**0.39**	**0.50**	**0.52**	**0.51**	**0.46**	**0.54**	**0.39**	**0.57**	**0.46**	**0.55**	**0.40**
		0.31	0.24	0.42	0.26	0.39	0.38	0.30	0.13	0.32	0.25	0.30	0.18
Fold 5	Accuracy	**0.73**	**0.72**	**0.74**	**0.75**	**0.74**	**0.71**	**0.77**	**0.74**	**0.78**	**0.76**	**0.77**	**0.75**
		0.72	0.64	0.73	0.68	0.72	0.66	0.72	0.75	0.73	0.74	0.72	0.74
	KSS	**0.42**	**0.44**	**0.44**	**0.49**	**0.44**	**0.37**	**0.49**	**0.47**	**0.52**	**0.47**	**0.50**	**0.45**
		0.26	0.21	0.31	0.30	0.26	0.26	0.20	0.22	0.23	0.20	0.21	0.21
	HSS	**0.44**	**0.48**	**0.46**	**0.52**	**0.46**	**0.42**	**0.53**	**0.46**	**0.55**	**0.49**	**0.52**	**0.46**
		0.33	0.26	0.37	0.35	0.33	0.30	0.27	0.29	0.31	0.26	0.28	0.28
Average	Accuracy	**0.73**	**0.71**	**0.75**	**0.73**	**0.75**	**0.71**	**0.77**	**0.75**	**0.78**	**0.74**	**0.77**	**0.75**
		0.71	0.68	0.73	0.71	0.73	0.70	0.73	0.72	0.74	0.72	0.73	0.71
	KSS	**0.42**	**0.41**	**0.46**	**0.44**	**0.47**	**0.40**	**0.51**	**0.45**	**0.55**	**0.43**	**0.53**	**0.48**
		0.29	0.21	0.34	0.26	0.33	0.26	0.21	0.16	0.24	0.19	0.22	0.15
	HSS	**0.45**	**0.42**	**0.50**	**0.45**	**0.50**	**0.42**	**0.53**	**0.48**	**0.56**	**0.45**	**0.54**	**0.49**
		0.35	**0.26**	**0.39**	**0.30**	**0.38**	**0.30**	**0.28**	**0.22**	**0.31**	**0.25**	**0.29**	**0.21**

computed as 0.71, 0.40 and 0.42, and for TRB, the same are computed as 0.75, 0.47 and 0.48, respectively.

It is also noticed that the model performance is comparatively low for higher extreme categories, i.e., D2 and W2, as compared with lower extreme categories, such as D1, D0, W0 and W1. However, in general, drought categories and wet categories are identified with almost equal skill. In other words, there is no selective bias toward any side of extremes. Regarding higher extreme categories, these are also captured by the DL-based model for many cases but not as excellently as lower extremes. This might be justifiable in the context of the tendency of the model to capture the central values by default. Higher extremes always need more accurate model performance and may

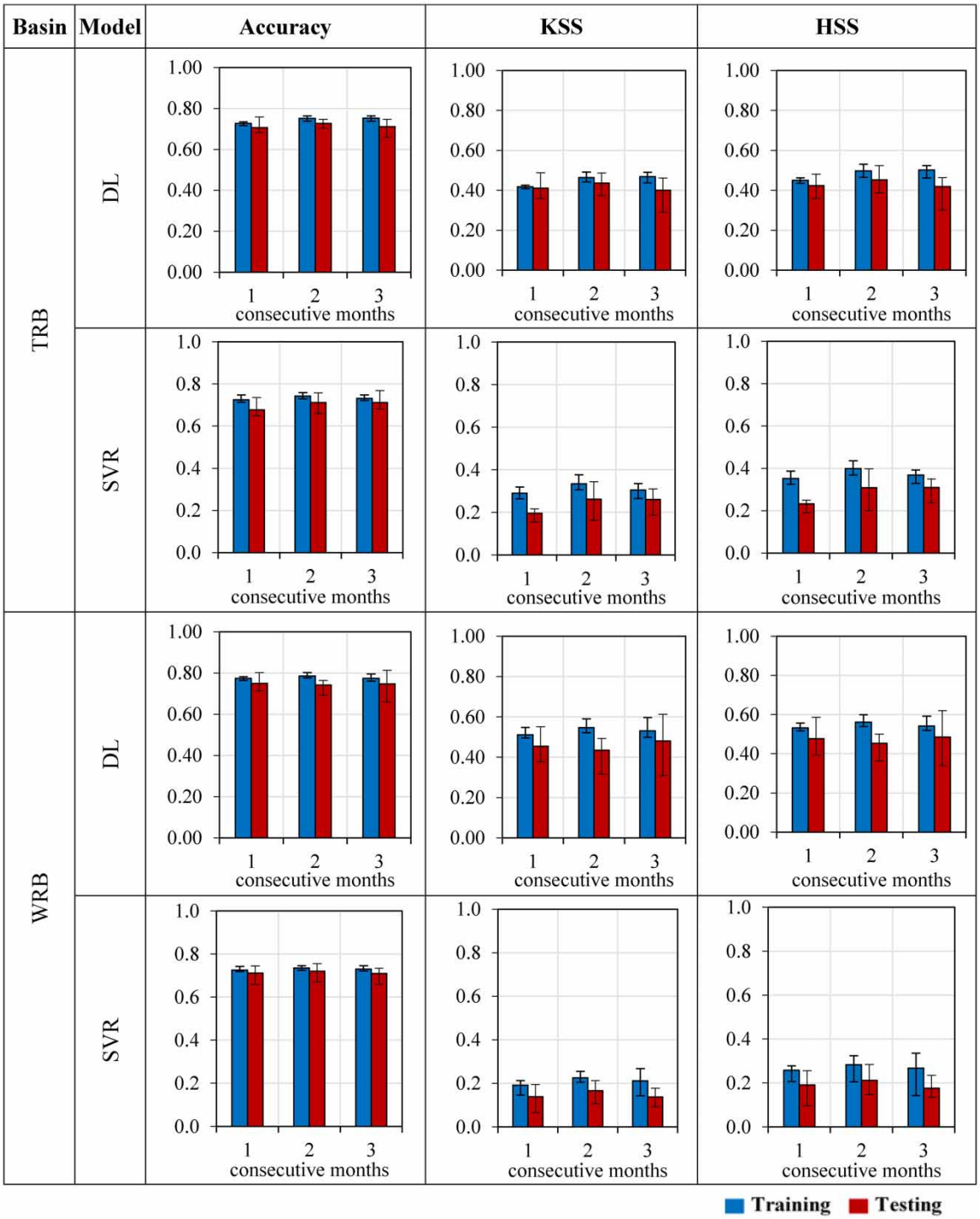


Figure 8 | Average accuracy, KSS and HSS, along with range bars at the top, show the maximum to minimum value, across different folds for drought category identification for both the study basins using DL and SVR models for 1st, 2nd and 3rd months. Please refer to the online version of this paper to see this figure in colour: http://dx.doi.org/10.2166/wcc.2021.062.

Table 5 | Contingency table for the observed and simulated drought category for the TRB study basin considering all the testing periods in different folds so as to consider the entire time period of analysis as testing period only

Drought category		Simulated							
		D2	D1	D0	N	W0	W1	W2	Sum
Observed	D2	2	4	2	1	0	0	0	**9**
	D1	1	10	6	7	1	0	0	**25**
	D0	1	1	13	21	3	0	0	**39**
	N	0	1	12	279	19	4	0	**315**
	W0	0	0	1	24	16	5	1	**47**
	W1	0	0	0	4	8	11	1	**24**
	W2	0	0	0	0	5	3	3	**11**
	Sum	**4**	**16**	**34**	**336**	**52**	**23**	**5**	**470**

Table 6 | Same as Table 5 but for the WRB study basin

Drought category		Simulated							
		D2	D1	D0	N	W0	W1	W2	Sum
Observed	D2	5	1	4	0	0	0	0	**10**
	D1	2	2	10	4	0	0	0	**18**
	D0	0	4	18	17	0	0	0	**39**
	N	0	0	12	295	14	3	1	**325**
	W0	0	1	1	12	20	6	0	**40**
	W1	0	0	0	3	9	7	1	**20**
	W2	0	0	0	1	6	7	4	**18**
	Sum	**7**	**8**	**45**	**332**	**49**	**23**	**6**	**470**

be controlled by more number of factors in much more complex ways. In other words, the higher the extremes the more complexity is involved. Thus, the model performs marginally weaker for the higher extremes.

Summarizing the results, the superiority of DL algorithms over traditional ML models in foreseeing the drought status can be emphasized. Similar findings were indicated in the recent literature as well (e.g., Reichstein *et al.* 2019; Xiao *et al.* 2019; Dikshit *et al.* 2021). Our analysis of comparing the DL-based model with the SVR model reaffirmed the same in the case of drought assessment. However, despite its better performance, only a handful of studies have used the DL algorithm for assessing the droughts across the globe, as discussed before (Chen *et al.* 2012; Agana & Homaifar 2017; Shen *et al.* 2019). For the Indian subcontinent also, the drought study using DL is rare, and as of now, there are no studies for the study basins considered in this study. Next, the method of extracting the information from the hydrometeorological precursors for drought assessment may lead to a new research direction, and the potential of DL algorithms may be beneficially used. Furthermore, the study basins used in this study are located in the drought-prone regions in India. The suitability of the DL model application can also be justified with the results obtained from contingency tables, where most of the drought and wet events are characterized accurately for both the basins. Overall, the DL-based model exhibits reliable potential in extracting the drought status from hydrometeorological precursors. It is true that a large data requirement for proper training is a limitation for its applicability to the data-scarce regions. Thus, the spatial transferability of the DL-based approaches may be another future research direction.

CONCLUSIONS

This study explores the potential of the DL-based model (here, Conv1D) for basin-scale drought assessment using the complex association between rainfall variation and hydrometeorological precursors. The performance of the proposed model is evaluated at two river basins in India, namely Damodar River Basin up to Tenughat Dam in eastern India (TRB) and Wardha River Basin up to upper Wardha Dam in western India (WRB). Being located in the drought-prone central belt of India, these basins are frequently stricken by droughts due to vagaries of precipitation. The potential of the DL-based model is used to extract the information from a set of hydrometeorological precursors, namely air temperature, surface pressure, wind speed, relative humidity, evaporation, soil moisture and geopotential height for the assessment of drought status, characterized through SPAI. The following conclusions are drawn from the study:

- DL has the potential to successfully capture the complex relationship between different hydrometeorological precursors and rainfall variation. Simultaneous modeling of 3 months is possible with a suitable model architecture that can be run with standard computing facility yielding reasonable accuracy.
- Findings of this study emphasize that the potential of the DL-based approach to extract the hidden complex

hydrometeorological association will be highly beneficial because, in general, climate model-simulated rainfall estimates are of inferior quality than other meteorological simulations, and on the other hand, the observed records of meteorological variables are either not available or sparsely available at many places over India.

- Demonstration with droughts in this study, however, using reanalysis products of hydrometeorological variables, shows the efficacy of the DL-based model. It may be noted that drought characterization adds another level of nonlinearity with respect to rainfall deviations from long-term mean. We used SPAI as a drought characterization index and different statistical performance measures (CC, RMSE, NSE, accuracy, KSS and HSS) that confirm a well correspondence between the observed and modeled SPAI values for both study basins.
- A comparison against another popularly used ML approach, i.e., SVR, clearly favors the DL-based model. Though all the conditions were kept same as that in the DL-based model, SVR needs to be trained for each month separately, while the DL-based model is able to provide the output for all three consecutive months simultaneously. Thus, computational ease also favors the DL-based approach, and the superiority of DL algorithms over traditional ML models in assessing the drought status is established.
- The method of information extraction from the hydrometeorological precursors for drought assessment using the potential of the DL-based method may lead to a new research direction. Demonstration with two study basins, located in the drought-prone central belt of India, indicates the promise of the DL-based model, as most of the drought and wet events are characterized accurately for both the basins.

Findings of this study also lead to a possible scope for future research to utilize the potential of climate models in simulating/predicting the primary hydrometeorological variables and, thereafter, to utilize the potential of DL to assess the status of secondary/tertiary hydrometeorological variables. In general, a proper training of such models needs a large amount of data, which is a shortcoming. To utilize the potential of such models in data-scarce regions, the spatial transferability of aforementioned models needs to be explored. This is kept as a future scope of this study.

ACKNOWLEDGEMENT

This work is partially supported by the Department of Science and Technology, Climate Change Programme (SPLICE), Government of India (Ref. No. DST/CCP/CoE/79/2017(G)) through a sponsored project.

DATA AVAILABILITY STATEMENT

The data used in this study are available from online repositories. The observed daily precipitation records are obtained from IMD (https://www.imdpune.gov.in/Clim_Pred_LRF_New/Grided_Data_Download.html). Similarly, ERA5 reanalysis data are downloaded from https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5.

REFERENCES

Achieng, K. O. 2019 Modelling of soil moisture retention curve using machine learning techniques: artificial and deep neural networks vs support vector regression models. *Comput. Geosci.* **133**, 104320. https://doi.org/10.1016/j.cageo.2019.104320.

Adnan, S., Ullah, K., Shuanglin, L., Gao, S., Khan, A. H. & Mahmood, R. 2018 Comparison of various drought indices to monitor drought status in Pakistan. *Clim. Dyn.* **51**, 1885–1899. https://doi.org/10.1007/s00382-017-3987-0.

Agana, N. A. & Homaifar, A. 2017 A deep learning based approach for long-term drought prediction. In: *Conf. Proc. – IEEE SOUTHEASTCON 2017*, March 30–April 2, 2017, Charlotte, NC. https://doi.org/10.1109/SECON.2017.7925314.

Amrit, K., Pandey, R. P. & Mishra, S. K. 2018 Characteristics of meteorological droughts in northwestern India. *Nat. Hazards* **94**, 561–582. https://doi.org/10.1007/s11069-018-3402-0.

Belayneh, A., Adamowski, J., Khalil, B. & Ozga-Zielinski, B. 2014 Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural networks and wavelet support vector regression models. *J. Hydrol.* **508**, 418–429. https://doi.org/10.1016/j.jhydrol.2013.10.052.

Chai, T. & Draxler, R. R. 2014 Root mean square error (RMSE) or mean absolute error (MAE)? – arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **7**, 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014.

Chanda, K. & Maity, R. 2015 Meteorological drought quantification with standardized precipitation anomaly index

for the regions with strongly seasonal and periodic precipitation. *J. Hydrol. Eng.* **20**, 06015007. https://doi.org/10.1061/(asce)he.1943-5584.0001236.

Chen, J., Jin, Q. & Chao, J. 2012 Design of deep belief networks for short-term prediction of drought index using data in the Huaihe River Basin. *Math. Probl. Eng.* **2012**, 1–16. https://doi.org/10.1155/2012/235929.

Choy, K. Y. & Chan, C. W. 2003 Modelling of river discharges and rainfall using radial basis function networks based on support vector regression. *Int. J. Syst. Sci.* **34** (14–15), 763–773. https://doi.org/10.1080/00207720310001640241.

Das, P. & Chanda, K. 2020 Bayesian Network based modeling of regional rainfall from multiple local meteorological drivers. *J. Hydrol.* **591**, 125563. https://doi.org/10.1016/j.jhydrol.2020.125563.

Deo, R. C. & Sahin, M. 2015 Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern Australia. *Atmos. Res.* **153**, 512–525. https://doi.org/10.1016/j.atmosres.2014.10.016.

Dikshit, A., Pradhan, B. & Alamri, A. M. 2021 Long lead time drought forecasting using lagged climate variables and a stacked long short-term memory model. *Sci. Total Environ.* **755**, 142638. https://doi.org/10.1016/j.scitotenv.2020.142638.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A. & Vapnik, V. 1997 Support vector regression machines. *Adv. Neural Inf. Process. Syst.* **9**, 155–161.

Feng, P., Wang, B., Liu, D. L. & Yu, Q. 2019 Machine learning-based integration of remotely-sensed drought factors can improve the estimation of agricultural drought in South-Eastern Australia. *Agric. Syst.* **173**, 303–316. https://doi.org/10.1016/j.agsy.2019.03.015.

FAO 2002 *Report of FAO-CRIDA Expert Group Consultation on Farming System and Best Practices for Drought-Prone Areas of Asia and the Pacific Region*. Food and Agricultural Organisation of United Nations. Central Research Institute.

FAO 2019 *The State of Food and Agriculture 2019 Moving Forward on Food Loss and Waste Reduction*. Food and Agricultural Organization of United Nations, Rome. ISBN: 978-92-5-131789-1.

Fung, K. F., Huang, Y. F., Koo, C. H. & Soh, Y. W. 2020 Drought forecasting: a review of modelling approaches 2007–2017. *J. Water Clim. Change* **11** (3), 771–799. https://doi.org/10.2166/wcc.2019.236.

Ganguli, P. & Reddy, M. J. 2014 Ensemble prediction of regional droughts using climate inputs and the SVM-copula approach. *Hydrol. Process.* **28**, 4989–5009. https://doi.org/10.1002/hyp.9966.

Ghorbani, M. A., Khatibi, R., Karimi, V., Yaseen, Z. M. & Zounemat-Kermani, M. 2018 Learning from multiple models using artificial intelligence to improve model prediction accuracies: application to river flows. *Water Resour. Manag.* **32** (13), 4201–4215. https://doi.org/10.1007/s11269-018-2038-x.

Granata, F. 2019 Evapotranspiration evaluation models based on machine learning algorithms – a comparative study. *Agric. Water Manag.* **217**, 303–315. https://doi.org/10.1016/j.agwat.2019.03.015.

Haidar, A. & Verma, B. 2018 Monthly rainfall forecasting using one-dimensional deep convolutional neural network. *IEEE Access* **6**, 69053–69063. https://doi.org/10.1109/ACCESS.2018.2880044.

Hao, Z., Singh, V. P. & Xia, Y. 2018 Seasonal drought prediction: advances, challenges, and future prospects. *Rev. Geophys.* **56**, 108–141. https://doi.org/10.1002/2016RG000549.

Hinton, G. E. & Salakhutdinov, R. R. 2006 Reducing the dimensionality of data with neural networks. *Science* **313** (5786), 504–507. https://doi.org/10.1126/science.1127647.

Kao, S. C. & Govindaraju, R. S. 2010 A copula based joint deficit index for droughts. *J. Hydrol.* **380**, 121–134. https://doi.org/10.1016/j.jhydrol.2009.10.029.

Kaur, A. & Sood, S. K. 2020 Deep learning based drought assessment and prediction framework. *Ecol. Inform.* **57**, 101067. https://doi.org/10.1016/j.ecoinf.2020.101067.

Khan, M. I. & Maity, R. 2020 Hybrid deep learning approach for multi-step-ahead daily rainfall prediction using GCM simulations. *IEEE Access* **8**, 52774–52784. https://doi.org/10.1109/ACCESS.2020.2980977.

Khan, N., Sachindra, D. A., Shahid, S., Ahmed, K., Shiru, M. S. & Nawaz, N. 2020 Prediction of droughts over Pakistan using machine learning algorithms. *Adv. Water Resour.* **139**, 103562. https://doi.org/10.1016/j.advwatres.2020.103562.

Kingma, D. P. & Ba, J. 2014 Adam: a method for stochastic optimization. In: *3rd International Conference on Learning Representations (ICLR 2015)*, May 7–9, 2015, San Diego, CA, USA.

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M. & Inman, D. J. 2021 1D convolutional neural networks and applications: a survey. *Mech. Syst. Signal Process.* **151**, 107398. https://doi.org/10.1016/j.ymssp.2020.107398.

Lantz, B. 2015 *Machine Learning with R*, Second Edition. Packt Publishing Ltd., Birmingham. https://doi.org/10.1002/9781119642183.ch14.

Le, J. A., El-Askary, H. M., Allali, M. & Struppa, D. C. 2017 Application of recurrent neural networks for drought projections in California. *Atmos. Res.* **188**, 100–106. https://doi.org/10.1016/j.atmosres.2017.01.002.

Lecun, Y., Bengio, Y. & Hinton, G. 2015 Deep learning. *Nature* **521**, 436–444. https://doi.org/10.1038/nature14539.

Maity, R., Bhagwat, P. P. & Bhatnagar, A. 2010 Potential of support vector regression for prediction of monthly streamflow using endogenous property. *Hydrol. Process. Int. J.* **24** (7), 917–923. https://doi.org/10.1002/hyp.7535.

Maity, R., Chanda, K., Dutta, R., Ratnam, J. V., Nonaka, M. & Behera, S. 2020 Contrasting features of hydroclimatic teleconnections and the predictability of seasonal rainfall over east and west Japan. *Meteorol. Appl.* **27**, 1–20. https://doi.org/10.1002/met.1881.

Makokha, G. O., Wang, L., Zhou, J., Li, X., Wang, A., Wang, G. & Kuria, D. 2016 Quantitative drought monitoring in a typical

cold river basin over Tibetan Plateau: an integration of meteorological, agricultural and hydrological droughts. *J. Hydrol.* **543**, 782–795. https://doi.org/10.1016/j.jhydrol.2016.10.050.

McKee, T. B., Doesken, N. J. & Kleist, J. 1993 The relationship of drought frequency and duration to time scales. In: *Proceedings of the 8th Conference on Applied Climatology*, pp. 179–183. https://doi.org/10.1002/jso.23002.

Mishra, A. K. & Desai, V. R. 2006 Drought forecasting using feed-forward recursive neural network. *Ecol. Modell.* **198**, 127–138. https://doi.org/10.1016/j.ecolmodel.2006.04.017.

Mishra, A. K. & Singh, V. P. 2010 A review of drought concepts. *J. Hydrol.* **391**, 202–216. https://doi.org/10.1016/j.jhydrol.2010.07.012.

Mishra, A. K., Desai, V. R. & Singh, V. P. 2007 Drought forecasting using a hybrid stochastic and neural network model. *J. Hydrol. Eng.* **12**, 626–638. https://doi.org/10.1061/(asce)1084-0699(2007)12:6(626).

Monish, N. T. & Rehana, S. 2020 Suitability of distributions for standard precipitation and evapotranspiration index over meteorologically homogeneous zones of India. *J. Earth Syst. Sci.* **129**, 1–19. https://doi.org/10.1007/s12040-019-1271-x.

Morid, S., Smakhtin, V. & Bagherzadeh, K. 2007 Drought forecasting using artificial neural networks and time series of drought indices. *Int. J. Climatol.* **27**, 2103–2111. https://doi.org/10.1002/joc.

Mouatadid, S., Raj, N., Deo, R. C. & Adamowski, J. F. 2018 Input selection and data-driven model performance optimization to predict the Standardized Precipitation and Evaporation Index in a drought-prone region. *Atmos. Res.* **212**, 130–149. https://doi.org/10.1016/j.atmosres.2018.05.012.

Mukherjee, S., Mishra, A. & Trenberth, K. E. 2018 Climate change and drought: a perspective on drought indices. *Curr. Clim. Chang. Rep.* **4**, 145–163. https://doi.org/10.1007/s40641-018-0098-x.

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.

Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S. & Mukhopadhyay, B. 2014 Development of a new high spatial resolution (0.25 × 0.25) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam* **65** (1), 1–18.

Pal, M., Maity, R., Ratnam, J. V., Nonaka, M. & Behera, S. K. 2020 Long-lead prediction of ENSO modoki index using machine learning algorithms. *Sci. Rep.* **10** (1), 1–13. https://doi.org/10.1038/s41598-019-57183-3.

Palmer, W. C. 1965 *Meteorological Drought*. Res. Paper No. 45. Weather Bureau, Washington, DC.

Park, J., Sung, J. H., Lim, Y. J. & Kang, H. S. 2019 Introduction and application of non-stationary standardized precipitation index considering probability distribution function and return period. *Theor. Appl. Climatol.* **136**, 529–542. https://doi.org/10.1007/s00704-018-2500-y.

Pearson, K. & Henrici, O. M. F. E. 1896 VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. *Philos. Trans. R. Soc. London. A* **187**, 253–318. https://doi.org/10.1098/rsta.1896.0007.

Pereira, L. S., Cordery, I. & Iacovides, I. 2002 *Coping with Water Scarcity*. UNESCO IHP VI, Technical Documents in Hydrology 58. UNESCO, Paris.

Qasem, S. N., Samadianfard, S., Kheshtgar, S., Jarhan, S., Kisi, O., Shamshirband, S. & Chau, K. W. 2019 Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural networks in arid and humid climates. *Eng. Appl. Comput. Fluid Mech.* **13** (1), 177–187. https://doi.org/10.1080/19942060.2018.1564702.

Rehana, S. & Monish, N. T. 2020 Characterization of regional drought over water and energy limited zones of India using potential and actual evapotranspiration. *Earth Sp. Sci.* **7**. https://doi.org/10.1029/2020EA001264.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J. & Carvalhais, N. 2019 Deep learning and process understanding for data-driven Earth system science. *Nature* **566** (7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1.

Russo, S., Dosio, A., Sterl, A., Barbosa, P. & Vogt, J. 2013 Projection of occurrence of extreme dry-wet years and seasons in Europe with stationary and nonstationary Standardized Precipitation Indices. *J. Geophys. Res. Atmos.* **118**, 7628–7639. https://doi.org/10.1002/jgrd.50571.

Sachindra, D. A. & Kanae, S. 2019 Machine learning for downscaling: the use of parallel multiple populations in genetic programming, Stochastic Environmental Research and Risk Assessment. *Stoch. Environ. Res. Risk Assess.* https://doi.org/10.1007/s00477-019-01721-y.

Sanderson, M. 1990 *UNESCO Sourcebook in Climatology for Hydrologists and Water Resource Engineers (No. 551.5 U56)*. UNESCO, París (Francia).

Shamshirband, S., Hashemi, S., Salimi, H., Samadianfard, S., Asadi, E., Shadkani, S., Kargar, K., Mosavi, A., Nabipour, N. & Chau, K. W. 2020 Predicting standardized streamflow index for hydrological drought using machine learning models. *Eng. Appl. Comput. Fluid Mech.* **14** (1), 339–350. https://doi.org/10.1080/19942060.2020.1715844.

Shen, R., Huang, A., Li, B. & Guo, J. 2019 Construction of a drought monitoring model using deep learning based on multi-source remote sensing data. *Int. J. Appl. Earth Obs. Geoinf.* **79**, 48–57. https://doi.org/10.1016/j.jag.2019.03.006.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. 2014 Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15** (1), 1929–1958.

WMO 2006 *Drought Monitoring and Early Warning: Concepts, Progress and Future Challenges*. World Meteorological Organization, Geneva, Switzerland. ISBN: 978-92-63-11006-0.

Won, J., Choi, J., Lee, O. & Kim, S. 2020 Copula-based joint drought index using SPI and EDDI and its application to climate change. *Sci. Total Environ.* **744**, 140701. https://doi.org/10.1016/j.scitotenv.2020.140701.

World Bank 2003 *Report on Financing Rapid Onset Natural Disaster Losses in India: A Risk Management Approach.* Report No. 26844-IN, Washington, DC.

Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J. & Chen, Z. 2019 Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. *Remote Sens. Environ.* **233**, 111358. https://doi.org/10.1016/j.rse.2019.111358.

Xu, L. & Mo, K. C. 2020 A preliminary study of deep learning based drought forecast. Climate prediction S&T digest. In *Science and Technology Infusion Climate Bulletin NOAA's National Weather Service, 44th NOAA Annual Climate Diagnostics and Prediction Workshop*, 22–24 October 2019, Durham, NC.

Zhang, G., Patuwo, B. E. & Hu, M. Y. 1998 Forecasting with artificial neural networks: the state of the art. *Int. J. Forecast.* **14** (1), 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7.

Zou, J., Han, Y. & So, S. S. 2008 Overview of artificial neural networks. *Methods Mol. Biol*. https://doi.org/10.1007/978-1-60327-101-1_2.

First received 27 February 2021; accepted in revised form 7 May 2021. Available online 21 May 2021