Water Resources Management
https://doi.org/10.1007/s11269-021-02899-z

Modeling Long-term Groundwater Levels By Exploring Deep Bidirectional Long Short-Term Memory using Hydro-climatic Data

Sangita Dey[1] · Arabin Kumar Dey[2] · Rajesh Kumar Mall[1]

Received: 11 November 2020 / Accepted: 28 June 2021

Abstract
Inevitable issues concerning the sustainability of groundwater resources are crucial under the present climatic situation. Therefore, the prevision of groundwater environments may able to reinforce the management system. In this respect present study considered a new method to predict long-term groundwater level framework as an alternative option of expensive physical models. The proposed Bidirectional Long Short-Term Memory (BLSTM) model can efficiently capture Spatio-temporal features from historical data. A highway LSTM network is also introduced within the architecture of the model to optimize the analysis. The relative performance of the proposed BLSTM with the highway LSTM (BHLSTM) network compared with simple BLSTM. Stack size increment of the BHLSTM and BLSTM layers can enhance the learning ability and improve by incorporating straight LSTM at the top of the architecture. The proposed model was applied to predict the groundwater level exemplary of the Varuna River basin for twenty years. The model incorporates the historical annual average of total precipitation, temperature, relative humidity, actual evapotranspiration, and groundwater level data to develop and validate the models. The result shows that the signals are captured reasonably well by a stack of four BHLSTM and straight LSTM models in forecasting groundwater levels. The predicted water level range (0—20 mbgl) has four categories low, medium, high, and very high which eventually, illustrates the water-threatened situation in upcoming years in the study area. It is also recommended to exploring this proposed method for further improvements and extensions towards interpreting spatial features.

Keywords Bidirectional LSTM · Highway LSTM · Spatio-temporal prediction · Climatic parameters · Groundwater level

✉ Rajesh Kumar Mall
rkmall@bhu.ac.in

[1] DST- Mahamana Centre of Excellence in Climate Change Research, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi 221005, India

[2] Department of Mathematics, Indin Institute of Technology, Guwahati, Guwahati, India

1 Introduction

Artificial intelligence and machine learning modeling have shown efficiency in most predictive analyses. The present study focuses on predicting groundwater levels based on certain factors observed around a particular river basin over a certain period. Groundwater level change is an important aspect that can be affected due to climate variation and human activities. Predictive analysis of world climate variation suggests plunges in mean precipitation and upsurges in temperature and solar radiation which escalates potential severity and frequency of drought events or scarcity of water resources (IPCC 2014). Hence the deficiency or water scarcity condition will initiate surface water resources meagerness, extra burden on groundwater resources to fulfill the societal needs. Therefore, groundwater storage is perceived as a dependent feature of climate variability and reflects groundwater level variation. However, long-term groundwater level premonition is an essential critical task; it is necessary as climate change leads to appalling significances for agriculture and environment sustainability, including food and water scarcity (Mall et al. 2006).

Over the past few decades, artificial intelligence models have gained considerable attention due to their advantages over numerical models and have been proved efficient in predicting complex hydrologic systems (Banerjee et al. 2009; Adamowski and Chan 2011; Garg 2014; Mirzavand et al. 2015; Gong et al. 2016; Alizamir et al. 2018; Afzaal et al. 2020; Bai et al. 2021; Feng et al. 2008). Coulibaly et al. (2001) and Style (2009) reviewed a comprehensive study on the application of artificial neural networks (ANN) in the prediction of groundwater fluctuation. Lallahem et al. (2005), for the first time, used ANN in a Multilayer Perceptron (MLP) to evaluate the dynamic water level in a karstic aquifer. They confirmed the ability of ANN in simulating groundwater level fluctuations of karstic aquifer compared to numerical models using the climate parameters. Several researchers have investigated the merit of the Spatio-temporal ANN model through distinct approaches such as the Fuzzy Inference System (FIS) (Zhu and Zhou 2009), Adaptive Neuro-Fuzzy Inference System (ANFIS) (Djurovic et al. 2015), Support Vector Regression (SVR) (Suryanarayana et al. 2014), Nonlinear Autoregressive Networks with Exogenous Input (NARX) (Guzman et al. 2017). A long-term groundwater level prediction had carried out using the WNN model that is trained by a novel improved algorithm to forecast one-year groundwater level conditions by Rakhshandehroo et al. (2017). The study lacks the Spatio-temporal variation and considers one deep and shallow well data. Coulibaly et al. (2001), first attempted the RNN model to predict groundwater level in a Spatio-temporal setup. The Recurrent Neural Network (RNN) is the most appropriate model for model-dependent data. However, direct RNN has drawbacks of vanishing gradient and exploding gradient problems, discouraging practitioners from using this model. Moreover, their approach raises questions regarding the spatial contributions within the network.

The Long Short-Term Memory (LSTM) networks successfully overcome such a problem. These models are capable of learning long-term dependencies of sequence data (Hochreiter et al. 1997). Recently some attempts are made to use long-short term memory (LSTM) based on recurrent neural networks (Zhang et al. 2018; Jeong and Park 2019; Wunsch et al. 2020) to predict groundwater level. Modeling a multidimensional sequence in Spatio-temporal data is attempted by Cui et al. 2018. They propose to use output dimensions as the spatial dimension for one particular variable and the time series model set up to predict the variables at all locations together. All the work mentioned related to LSTM considered a sufficient number of temporal data. Therefore, using straight LSTM was not a problem. However, the present dataset is slightly different. It does not have enough

temporal data sets, but it has a spatial structure, which demands a different approach to predict groundwater levels with minimum parameterization. Simultaneously, such a model should be able to handle expected climate variability. Considering all these factors, a more advanced, simple, robust, and efficient BLSTM model used to overcome the hitches. This study designs and evaluates an innovative methodology based on BLSTM to capture the Spatio-temporal behavior of an aquifer system or dynamic nature of groundwater with a limited data source. The Varuna river basin in Uttar Pradesh, India, was chosen as the study area. The area is situated in the central part of the Ganga river basin and is a highly agriculturally productive region. As a result, 23% and 31% of the area includes over-exploited and semi-critical groundwater resources (CGWB 2019), and the groundwater level has been declining steadily. This new approach can provide new insights for socially relevant predictions, attributions, and optimizations on a long-term basis, primarily to manage groundwater infrastructure.

2 Materials and Methods

2.1 Study Area

Varuna river basin is located in the central Ganga alluvial plain between 25°39°28.71 N to 25°19°44.61 N latitude and 81°45°57.46 E to 82°03°06.59 E longitude and covers an area of 3675 km^2. The basin encompasses fluvial sediments of Pleistocene to Recent, which forms a multi-layered aquifer system. The shallower aquifer has silt, clay, and silty clay horizon mainly forms the unconfined aquifer. The deeper aquifers are confined to semiconfined in nature and primarily made up of sand (Dey et al. 2021). The Varuna river basin (Fig. 1) experiences semiarid to sub-humid tropical climate and receives annual rainfall between 572 and 897 mm.

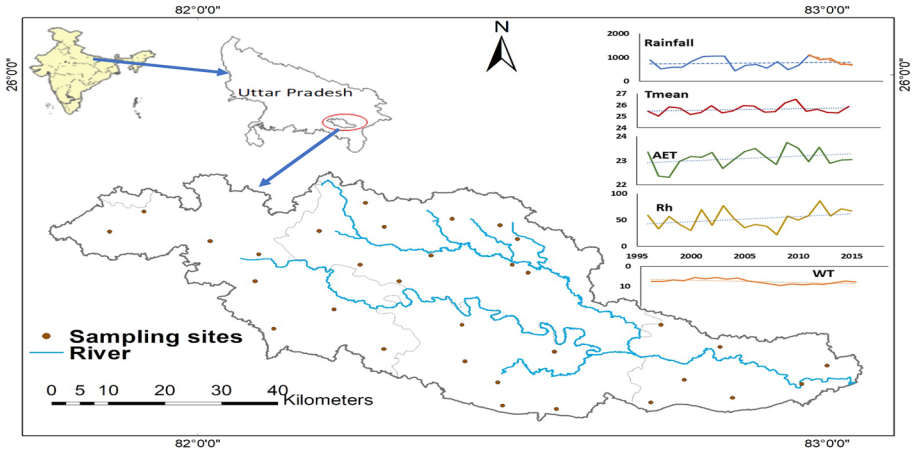

Fig. 1 Location map of Varuna river basin, Uttar Pradesh, India and variation of climatic parameters (Rainfall, Temperature mean (Tmean), Actual Evapotranspiration (AET), Relative Humidity (Rh)) and groundwater levels over the Varuna river basin

However, for the last two decades, the precipitation at the Varuna river basin (Avg. value of 31 locations) shows a slightly decreasing trend mainly since the last five years. As a result, the area receives a sharp diminution in amount (Fig. 1). Subsequently, the mean temperature and evapotranspiration show an increasing trend over the years, reflecting relative humidity (Fig. 1). Though a high level of relative humidity supports a hike in precipitation, an increase in temperature may hinder the tendency. Thus, particular circumstances create a low recharge condition for both the water resources. On the other hand, aquifers of this region manifest by decreasing temporal groundwater level trends over the years (Fig. 1) due to the rapid growth of human populations and expansion of agricultural, energy, and industrial sectors.

2.2 Dataset Description

The present dataset includes the time series data from 1996 to 2015. In addition, groundwater level data of bore wells of the Varuna River basin collected from India- WRIS (Water Resource Information System) web-portal (https://indiawris.gov.in/wris/#/groundWater). Furthermore, the Precipitation data and temperature data were collected from the India Meteorological Department in the same period, and relative humidity of the data obtained from the NASA POWER Data AccessViewer website (https://power.larc.nasa.gov/data-access-viewer/).

2.2.1 Evaluation of Potential Evapotranspiration (PET) uses the Hargreaves equation (NR under FAO 1998)

$$PET = 0.0023 \times (Tmean + 17.8) \times (Tmax - Tmin)0.5 * Ra \quad (1)$$

where PET = Potential Evapotranspiration, Tmax = Mean maximum temperature, Tmin = Mean minimum temperature, Tmean = Mean temperature calculated as (Tmax + Tmin)/2, Ra = Extra-terrestrial Radiation (mm/day) are obtained from standard graphs given in FAO (NR 1998). PET was further multiplied by 0.8 to get actual evapotranspiration (AET) (Kumar et al. 2011).

All models primarily depend on selecting significant input variables (Lallahem et al. 2005; Chitsazan et al. 2015; Djurovic et al. 2015). Precipitation is considered an essential recharging parameter to control the groundwater level. Evapotranspiration, relative humidity, and temperature are responsible for the discharge of groundwater. However, to make the selection design more concrete, it is necessary to ensure the dependency level, which is highly complex and explicitly not defined (Djurovic et al. 2015). Autocorrelations are a few frequently used methods (Coulibaly et al. 2000; Sudheer et al. 2002; Djurovic et al. 2015) to determine the dependency of parameters over the different time horizon. The correlation matrix in the present dataset shows that rainfall had a negative relationship with almost all the parameters, whereas other parameters had a strong positive relationship (Table 1). In an autocorrelation analysis, data related to rainfall clearly showed the serial dependence (Fig. 2). It is also very much possible that variance is time-varying. However, the presence of even one-time dependent variables put the whole multidimensional setup into a dependent structure.

Table 1 Correlation Matrix of different Climate Variables at a given time

	Rainfall	*Tmin*	*Tmax*	*Tmean*	*AET*	*Relative Humidity*
Rainfall	1					
Tmin	-0.227	1				
Tmax	-0.401	0.688	1			
Tmean	-0.374	0.839	0.972	1		
AET	-0.438	0.501	0.966	0.886	1	
Relative Humidity	-0.332	0.574	0.748	0.747	0.698	1

2.3 Methodology

Learning complex Spatio-temporal features of a large-scale aquifer network may be explored through applying dynamic and nondynamic predictions in an experimental model. The prediction pattern in the LSTM model depends on the construction of overlapping sets of observations. Suppose a sliding window of datasets passed through LSTM cells; it will gain an ability to predict only the next time step or filter observations of a particular variable in multidimensional set up at the end of the time steps. Nondynamic predictions generally filter some variable or predict the following time step observations based on some known input parameters. However, Observations at each time step may not always be available for the long- term in such prediction. Therefore, the dynamic prediction comes into the picture. Dynamic prediction sequentially uses the predicted groundwater level/other parameters as input parameters to predict groundwater level/other parameters next time. In this paper, the primary objective is to capture the Spatio-temporal behavior when only 20 years (annual) data point is available. BLSTM and its variations are proposed by taking the output dimension as a parameter dimension and sliding the widow of observations over different spaces. Usually, LSTM structures work based on forwarding dependences (Lipton et al. 2015). The chronologically arranged dataset used in the LSTM model passed the information from (t—1) time step to t time step in a positive direction only. But it is incapable of capturing the system's randomness or periodicity, which may solve by considering the backward dependencies. BLSTM is capable of dealing with both forward and backward

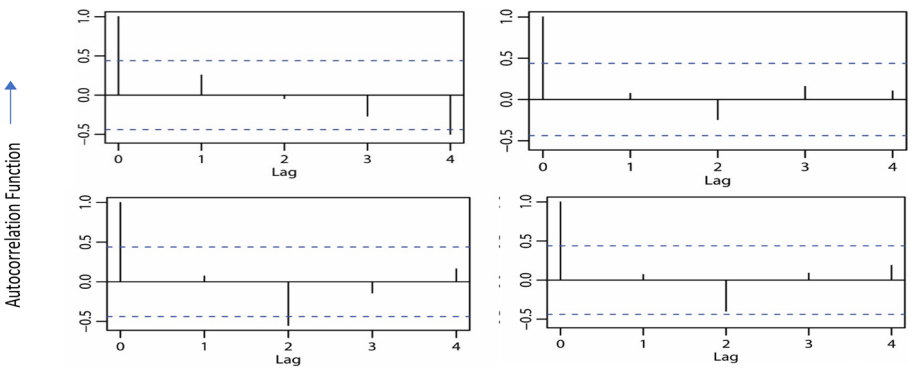

Fig. 2 Autocorrelation of the input parameters

dependencies and extracting concepts of multiple meanings. Therefore, BLSTM models learn both the spatial and temporal picture of datasets comprehensively. Also, BLSTM models combined with a straight LSTM network and highway network utilize Spatio-temporal information from the input data and achieve optimal performance.

2.3.1 Bidirectional Long Short-Term Memory (BLSTM)

BLSTM is a fabricated form of LSTM, a particular type of RNN. It converts the independent activations into dependent activations by assigning the same weight and biases to all the layers and memorizing each previous output to provide the next hidden layer as input. In simple RNN model (Fig. 3a), at each time iteration, t, the hidden layer continues a hidden state, h_t, and updates and accelerated it based on the layer input, x_t, and previous hidden state, h_{t-1}, using the following equation:

$$h_t = \sigma_h\left(Wx_t + Vh_t - 1 + b_h\right) \tag{2}$$

W is the weight matrix from the input layer to the hidden layer, V is the weight matrix between two consecutive hidden states (h_{t-1} and h_t), b_h is the bias vector of the hidden layer, and σ_h is the activation function to generate the hidden state. The network output can characterize as

$$y_t = \sigma_y\left(Uh_t + b_y\right) \tag{3}$$

U is the weight matrix from the hidden layer to the output layer, b_y is the bias vector of the output layer, and σ_y is the activation function of the output layer. Finally, the hidden layer generates the output y_t. The model's efficiency is reduced by its vanishing gradient problem during the propagation and is incapable of learning from long-term time lags (Gers et al. 2000). However, the advanced architecture of the LSTM efficiently controls the problem mentioned above. The motivation for the introduction of the LSTM cell is through input weight conflict and output weight conflict.

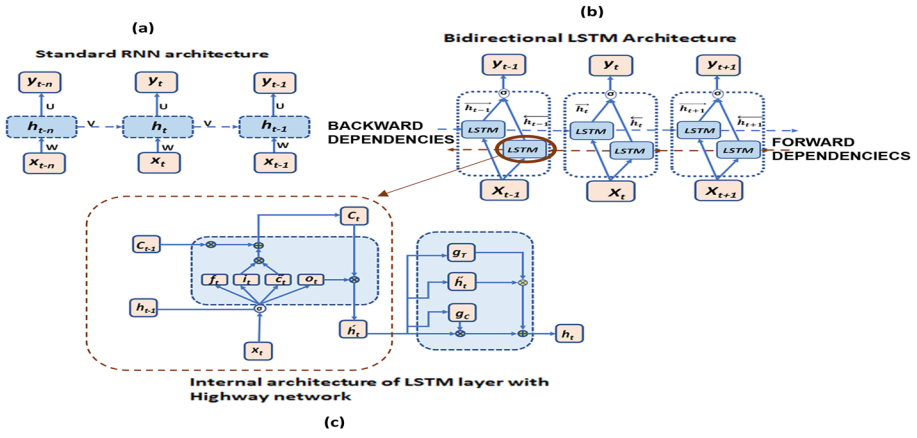

Fig. 3 Figure shows (**a**) standard RNN follows only forward dependencies and (**b**) BLSTM includes both forward and backward dependencies in which LSTM cell comprises of memory blocks represented within red rectangle region. (**c**)The LSTM layer is incorporated with the Highway network to form the BHLSTM layer

The LSTM architecture consists of current input x_t, the previous output h_{t-1}, and the last state cell c_{t-1}. Inside the cell, there are three gates, forget gate (f), input gate (i), and output gate (o). These multiplicative gates in LSTM memory cells store and access information over a long period thereby avoids the vanishing gradient problem. The following equations govern the calculation of the gate vectors

$$f_t = \sigma\left(Wfx_t + Ufh_t - 1 + b_f\right) \tag{4}$$

$$i_t = \sigma\left(W_i x_t + U_i h_t - 1 + b_i\right) \tag{5}$$

$$o_t = \sigma\left(W_o x_t + U_o h_t - 1 + b_o\right) \tag{6}$$

The new values of cell state c_t and output h_t (along with an intermediary variable $\tilde{c}_t$) are calculated by the following equations

$$c \sim_t = tanh\left(W_c x_t + U_c h_t - 1 + b_c\right) \tag{7}$$

$$c_t = i_t \odot c \sim_t + f_t \odot c_t - 1 \tag{8}$$

$$h_t = o_t \odot tanh\left(c_t\right) \tag{9}$$

Here, ⊙ denotes Hadamard product. W_f, W_i, W_o, W_c are the weights corresponding to input, and U_i, U_f, U_o, U_c are the weights corresponding to hidden units at previous time steps.b_f, b_i, b_o, and b_c denote different biases. σ is the activation function, usually the sigmoid function of the output layer, and the tanh is the hyperbolic tangent function. The final output of the LSTM layer takes the notation as $Y_T = [h_t, \ldots, h_{t-1}]$, which should be a vector of all the results.

Usually, the LSTM layers process sequence data uni-directionally and restrict it to capture the system's randomness or periodicity. However, introducing a backward LSTM layer into the network may fetch the problem, make the network bidirectional. Thus, the structure of an unfolded BLSTM layer containing a forward LSTM layer and a backward LSTM layer, which processes sequence data with two separate hidden layers and connects them to the same output layer (Fig. 3b). The inputs in a positive direction from the time frame t-n to t+n iteratively produce the forward layer output sequence. In contrast, the reverse data in the same period calculate the backward layer sequence. The layer outputs are computed using the standard LSTM Eqs. (3–4) and generate the final output y_t, derived from the following equation.

$$yt = \sigma\left(\overrightarrow{h_t}, h_t^{'}\right) \tag{10}$$

Where σ is the activation function to combine the two output sequences.

2.3.2 Highway LSTM

Highway networks was introduced to make the training of strong neural networks easy. The highway networks strengthen the LSTM layer to handle long-range dependencies. It increases the depth of the time to remember. The network consists of two gates, the Carry gate (g_C)

and the Transformation gate (g_T) (Fig. 3c). The transformation of input x to output y using highway network is given by.

$$g_C = \sigma\left(W_c x + b_C\right) \tag{11}$$

$$g_T = \sigma\left(W_T x + b_T\right) \tag{12}$$

$$y = x \odot C + tanh(Wx + b) \odot T \tag{13}$$

where W_C and W_T are the weight matrices for the Carry gate (g_C) and Transformation gate (g_T); b_C and b_T are the biases for the Carry gate and Transformation gate; W and b are the weight and bias in the model. Thus, the highway network can incorporate with the LSTM cell in two ways, one at the update of cell state c_t and the other at the update of the hidden unit h_t (Fig. 3c).

2.3.3 Stacked bidirectional LSTM with Highway LSTM network

A BLSTM architecture stacked with several hidden layers has enhanced the efficiency to predict more accurately than one layer (LeCun et al. 2015). When a Spatio-temporal dataset is fed to the BLSTM or BHLSTM, the spatial and temporal correlation of the input parameters in different locations can apprehend by the feature learning processes. The architecture includes several layers of BLSTM or BHLSTM and an LSTM layer at the top (Fig. 4). The bottommost layer of BLSTM or BHLSTM learns the valuable information from the Spatio-temporal dataset and passes into the next upper layer. Finally, the LSTM layer employed at the top for capturing the forward dependencies helps to improve the model performance in many cases.

2.4 Prediction performance index

We execute the goodness of fit based on RMSE values and C- index values. The comparison between different architectures can also evaluate by calculating the absolute RMSE and C-index. Lower RMSE indicates a better model, whereas a larger C-index represents better models or better fit to the data. In general, RMSE defined as,

$$RMSE = \sqrt{\frac{\sum_{i=0}^{n}\left(O_i - P_i\right)^2}{n}} \tag{14}$$

where, O_i = observed value for i^{th} data, P_i = predicted value for i^{th} data, and n = number of observations.

The concordance statistics, or the C-index (CI), measures the probability of concordance between the predicted and the observed values.

$$C(D, \varepsilon, f) = \frac{1}{|\varepsilon|}\sum\nolimits_{\varepsilon_{ij}} 1_{f(x_i)<f(x_j)} \tag{15}$$

with the indicator function $1_{a<b} = 1$ if $a < b$, and 0 if not; $|\varepsilon|$ denotes the number of edges in the order graph. $f(x_i)$ is the predicted value for subject i by model f. Perfect prediction accuracy is indicated by CI = 1, and CI = 0.5 is as good as a random predictor.

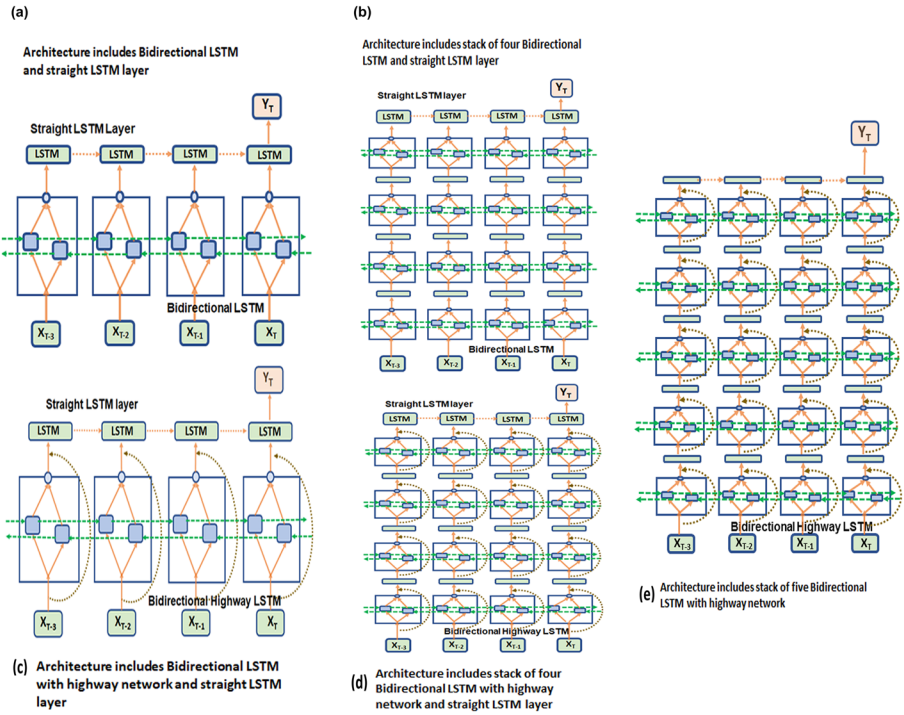

Fig. 4 Overview of the proposed model involves (**a**) single BLSTM and a straight LSTM, (**b**) stack of four BLSTM layers and a straight LSTM, (**c**) single BHLSTM and a straight LSTM, (**d**) stack of four a straight LSTM and (**e**) stack of five BHLSTM

2.5 Data Processing

The whole data sets are divided into two parts: training (75% data of time series dataset) and the other for testing (25% data of time series dataset) the trained models. Detrending, normalization, and outlier exclusion were carried out to improve the training efficiency and estimation accuracy. Feature scaling is performed in training machine learning models. The usual method for doing this is to normalize the entire data, feature-wise, to convert it to values between 0 and 1. Normalization of the data in this paper has made using the following relation of min–max scaling.

$$norm = \frac{X - Xmin}{Xmax - Xmin} \tag{16}$$

Since the study deals with a constant lookback while training the model (say, BLSTM), it makes sense to normalize the data in each lookback length window. Hence, each training sequence normalized between 0 to 1, and the prediction from the model is later denormalized. Thus, window feature scaling enhances training speed and enables the model to capture relative variance in the data in a small neighborhood.

2.6 Parameter Tuning

The input, output, and hidden cell dimensions were set as seven units. The training set with batch size 1. The total number of input variables was equal to 465, three-fourths of the dataset's overall quantity. The test set size is 155. We use the sliding window method with a window size of 20, which is also the lookback number or number of time steps taken during BLSTM training and testing for the data. The dropout value is 0.5 at every LSTM step.

2.7 Interpolation Method

The best-performed architecture was applied to predict the annual groundwater levels of 31 stations of the Varuna river basin. Then Inverse distance weighting (IDW) interpolation using ArcGIS 10.5 platform methods was used to estimate the overall groundwater level distribution. The computation of IDW uses a function of the distance between observed stations and the stations used for making a prediction (Gunnink and Burrough 1996). Thus, this interpolation may preserve some local variation of groundwater level.

3 Results and Discussion

3.1 Performance of Model Based On the Present Dataset: Observations and Inference

The analysis showed that the output groundwater level series largely depends on input climatic factors. Notably, the accuracy and trend of prediction are high with known climatic factors in dynamic and nondynamic setups. However, the dynamic prediction does not work well in the test set due to some overfitting problems. Hence the study chooses to use dropout as regularization, which prevents overfitting of the model. The parameter dropout is the fraction of the total units that will be randomly selected and turned off. The use of dropout enhanced the performance and provided a better graph approximating the test signal. For this experiment, the dropout is 0.5 at every LSTM step. However, the appropriate choice of dropout percentage at every level of the model demands more research. Adam optimizer had chosen during training. The training had done on 20 epochs. In highway LSTM, we consider the activation function as a rule. The loss function, in this case, is set as RMSE. In general, getting a performance with high accuracy is very difficult in the case of dynamic prediction. The paper carries information regarding tuning the parameters to get the best possible performance in dynamic prediction.

BLSTM has a riveting ability to learn spatial features inherently by the weights in LSTM layers at the training process, irrespective of the input data's spatial structure and spatial order. Therefore, the model performance is not affected by the magnitude of the spatial dimension of input data (Cui et al. 2018). Various datasets are apprehended between two locations by the LSTM weight matrices, storing established spatial correlation. Each of the sites under consideration took only 20-time steps in this study. Therefore, the analysis takes a lookback number value 20 to manage time and space-related features simultaneously.

The proposed model designs some architectures combining a single or several ascending (four stacks) BLSTM layers with a straight LSTM, BHLSTM layers with a straight LSTM, and a stack of five BHLSTM layers without a straight LSTM (Fig. 4). Computers with model: Intel(R) Core (TM) i5-6200U CPU 2.30 GHz. We use Python 3.5.2 to code all architectures.

The performances of the five architectures compared to choose the best architecture, to be used further for future forecasting. Table 2 illustrates the results of performance indices (RMSE and C-index) for both train and test scores. In terms of the influence of the stack size of the neural network, models achieved their best performance with five layers in non-dynamic prediction. RMSE and C- index improved as the number of layers increased from one to four, but minimum RMSE and maximum CI available in a stack of five BHLSTM only. Whereas in dynamic prediction, a stack of four layers outperformed the others (Table 2). The study indicates that one-layer BLSTM may be good enough for capturing features, but it is not satisfactory to predict the results. The nondynamic prediction is pretty good with stack size increment. Still, it restricts up to use for the historical period only because it is incapable of predicting the climate input parameters for long-term prediction and eventually getting the groundwater level's future output. In contrast, the dynamic forecast can predict input climate parameters and simultaneously output the groundwater level for an extended period. Hence, stack size increments of BHLSTM layers and a straight LSTM in dynamic prediction may improve the model performances more efficiently.

Proposed architectures adapted during training and testing of the data set, illustrated as exemplary of the Kashi subregion, which is analogous for each location considered in this study. The present study involved dynamic prediction with the best architecture (stack of four BHLSTM layers with a straight LSTM) for future forecasting. The proposed tuning is capable of capturing the groundwater level reasonably well. Results showed that original data and trained data almost overlapped with each other (Fig. 5). In fact, during testing, the ups and downs of the signal are well captured towards the last five years (Fig. 5). Hence, the proposed approach can deal with spatially distributed groundwater levels and works pretty well over a long period. In all cases, the prediction can capture the trend reflecting the mean of the time series at every time instance; surprisingly, variability has improved over time. Therefore, the proposed method is helpful in long-term future prediction. One of the drawbacks of such an estimate is measuring the confidence interval of the projection. Separate research needs to evaluate this issue in a Spatio-temporal setup and will address the same in an independent research article.

3.2 Case study: Water level long- term Prediction

The proposed model is applied to predict and visualize 20 years of groundwater circumstance over the Varuna river basin. The suggested network can predict the annual groundwater levels of 31 locations randomly selected over the study area. The anticipated results then interpolated spatially using ArcGIS 10.5 to represent the decadal changes over the basin. Spatial interpolation is carried for four decades, respectively, from 1996 – 2005, 2006 – 2015, 2016 – 2026, and 2027 – 2035 (Fig. 6). The entire range is categorized into four classes low (0 – 5 mbgl), medium (6 – 10 mbgl), high (11 – 15 mbgl), and very high (16 – 20 mbgl).

Table 2 Train and Test Score of different architectures during dynamic and nondynamic predictions of the proposed models

Architectures	*Benchmark based on RMSE*				*Benchmark based on C-Index*			
	Non-dynamic Prediction		*Dynamic Prediction*		*Non-dynamic Prediction*		*Dynamic Prediction*	
	Train Score	*Test Score*	*Train Score*	*Test Score*	*Train Score*	*Test Score*	*Train Score*	*Test Score*
Single BLSTM and a straight LSTM	*1.185*	*1.826*	*1.237*	*1.859*	*0.890*	*0.890*	*0.887*	*0.661*
Single BHLSTM and a straight LSTM	*1.219*	*1.853*	*1.219*	*1.769*	*0.887*	*0.888*	*0.892*	*0.654*
Stack of four BLSTM and a straight LSTM	*1.263*	*1.904*	*1.209*	*1.758*	*0.945*	*0.931*	*0.891*	*0.638*
Stack of four BHLSTM and a straight LSTM	*1.238*	*1.873*	*1.187*	*1.740*	*0.888*	*0.888*	*0.894*	*0.674*
Stack of five BHLSTM	*0.093*	*0.146*	*1.163*	*2.125*	*0.994*	*0.992*	*0.896*	*0.614*

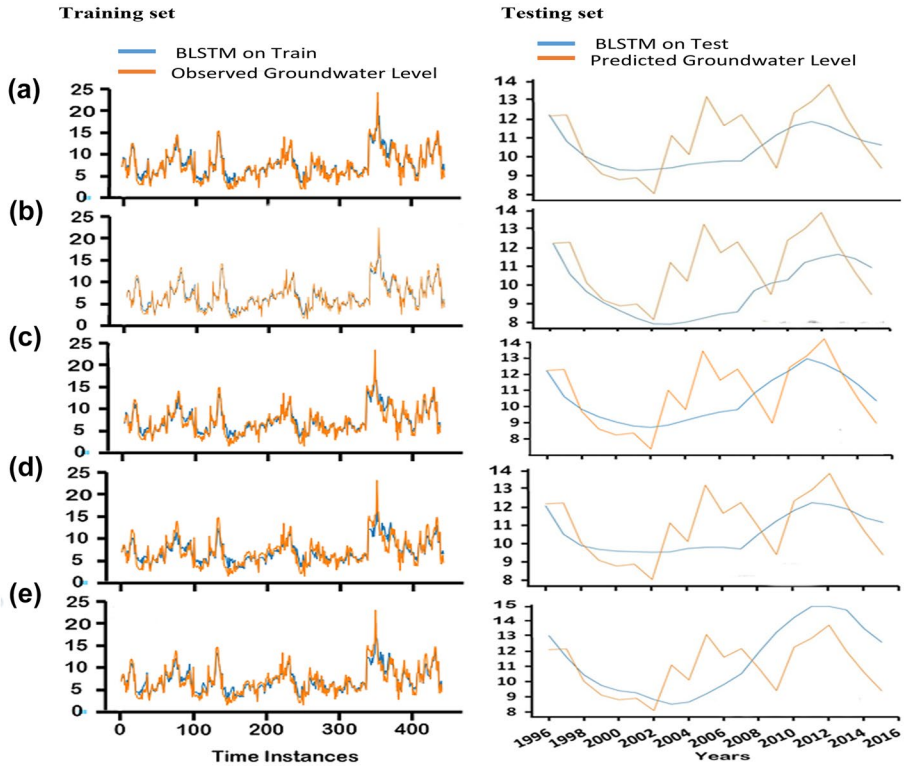

Fig. 5 Results of Training and Test Set for (**a**) single BLSTM and a straight LSTM (**b**) single BHLSTM and a straight LSTM (**c**) Stack of four BLSTM and a straight LSTM (**d**) Stack of four BHLSTM and a straight LSTM (**e**) Stack of five BHLSTM

The analysis result of the predicted values showed decreased amplitude and, eventually, a water-scarce scenario in the future. In the earlier two decades up to 2015, the north, northwestern, and southeastern zone of Jaunpur, Allahabad, and Sant Ravidas Nagar showed low groundwater level conditions. The western-most part of the Varanasi showed the transformation from low groundwater level conditions to very high groundwater level conditions. Still, in later decades the whole basin showed medium to very high groundwater level conditions. The result illustrates that the eco-environment of the groundwater system of the Varuna river basin will drop substantially in future periods. Subsequently, the acceleration in present society in economic, industrial, and agricultural growth claims a significant surge in groundwater demand in this region. Hence shortly, the decline rate and magnitude may be in an advanced form. In that case, the long-term prediction alerts about future damage may help initiate the proper water management through efficient planning. Based on the present groundwater level baseline, the forthcoming groundwater level needs a regular recharge to maintain. It may help to reverse the groundwater gradient and improve the damaged groundwater ecosystem. According to the physical and socio-economic characteristics of the study area, we implement feasible mitigation plans.

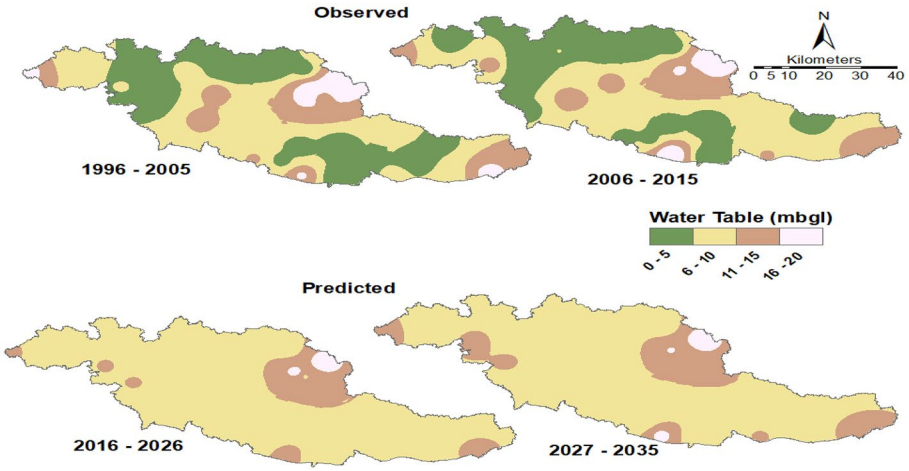

Fig. 6 Spatial distribution of Observed and Predicted water level over the study area

4 Conclusion

The present study was designed and attempted to predict the long-term groundwater level, which involved deep-stacked BHLSTM model architectures. The proposed model can learn and remember information for an extended period due to forward and backward dependencies. The presence of a highway network enhances its performance. The proposed model provides satisfactory performance in predicting the groundwater level. The increase in stack size in nondynamic prediction and stack size increments and a straight LSTM layer in dynamic prediction can improve the model's performance. The efficiency of dynamic prediction with a stack of four BHLSTM and a straight LSTM has proven to be the most appropriate architecture to predict the groundwater level for a long-term basis (20 years). The efficient architecture experimented over the Varuna river basin. Predicted values reveal that the river basin will face groundwater scarcity in the near and distant future, demanding proper sustainable groundwater planning to maintain prolific groundwater sources. Proper scientific conservation techniques for surface water and efficient irrigation techniques to reduce excess water stress should be implemented as early as possible. Low water consuming plants, water recycling facilities, water tariffs, domestic and industrial water-saving practices must sustain and maintain the equilibrium water balance condition so far and future.

Eventually, the proposed algorithm works pretty well in predicting the groundwater level. Predictions have not performed well if the number of data points decreases in these cases. Moreover, the proposed model can apply to other spatial and temporal prediction approaches, such as water quality changes, surface runoff, and base flow changes in hydrological studies.

Acknowledgements Sangita Dey thanks the Women Scientist Scheme (SR/WOS-A/EA-1004/2015), Department of Science and Technology, New Dealhi for support. Climate Data were acquired from India Meteorological Department, New Delhi, India. The authors are also thankful to DST- Mahamana Centre of Excellence in Climate Change Research (DST-MCECCR) for providing the lab facility and other support. The authors express their sincere gratitude to all reviewers and the Editor, Assistant Editor, for their constructive comments that have improved the paper.

Authors' Contributions All authors contributed to the study's conception and design. Material preparation, data collection, and analysis done by Arabin Kumar Dey and Sangita Dey. The first draft written by Sangita Dey and review and editing done by R. K. Mall and Sangita Dey. All authors read and approved the final manuscript.

Funding Women Scientist Scheme A (Reference no. SR/ WOS-A/EA-1004/2015), Department of Science and Technology, New Delhi.

Data Availability Data are available from the corresponding author by request.

Code Availability The codes used in this work are available from the corresponding author by request.

Declarations

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

References

Adamowski J, Chan H (2011) A wavelet neural network conjunction model for groundwater level forecasting. J Hydrol 407:28–40. https://doi.org/10.1016/j.jhydrol.2011.06.013

Afzaal H, Farooque AA, Abbas F, Acharya B, Esau T (2020) Groundwater estimation from major physical hydrology components using artificial neural networks and deep learning. Water 2(1):5

Alizamir M, Kisi O, Zounemat-Kermani M (2018) Modelling long-term groundwater fluctuations by extreme learning machine using hydro-climatic data. Hydrol Sci J 63(1):63–73

Bai Y, Bezak N, Zeng B, Li C, Sapač K, Zhang J (2021) Daily Runoff Forecasting Using a Cascade Long Short-Term Memory Model that Considers Different Variables. Water Resour Manag 19:1–5

Banerjee P, Prasad R, Singh V (2009) Forecasting of groundwater level in hard rock region using artificial neural network. Environ Geol 58(6):1239–1246

CGWB (2019) National Compilation on dynamic Ground Water Resources of India, 2017, Government of India, Ministry of Jal Shakti, Department of Water Resources, RD & GR, Central Ground Water Board, pp 298. http://www.cgwb.gov.in

Chitsazan M, Rahmani G, Neyamadpour A (2015) Forecasting groundwater level by artificial neural networks as an alternative approach to groundwater modeling. J Geol Soc India 85:98–106. https://doi.org/10.1007/s12594-015-0197-4

Coulibaly P, Anctil F, Bob'ee B (2000) Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. J Hydrol 230(3–4):244–257

Coulibaly P, Anctil F, Aravena R, Bobe'e B (2001) Artificial neural network modeling of water table depth fluctuations. Water Resour Res 37(4):885–896

Cui Z, Member S, Ke R, Member S, Wang Y (2018) 1801.02143 1–12. https://arxiv.org/abs/1801.02143

Dey S, Shukla UK, Mehrishi P, Mall RK (2021) Appraisal of groundwater potentiality of multilayer alluvial aquifers of the Varuna river basin, India, using two concurrent methods of MCDM. Environ Dev Sustain. https://doi.org/10.1007/s10668-021-01400-5

Djurovic N, Domazet M, Stricevic R, Pocuca V, Spalevic V, Pivic R, Gregoric E, Domazet U (2015) Comparison of Groundwater Level Models Based on Artificial Neural Networks and ANFIS. Sci World J 2015 https://doi.org/10.1155/2015/742138

Feng S, Kang S, Huo Z, Chen S, Mao X (2008) Neural networks to simulate regional groundwater levels affected by human activities. Groundwater 46:80–90. https://doi.org/10.1111/j.1745-6584.2007.00366.x

Garg V (2014) Modeling catchment sediment yield: a genetic programming approach. Nat Hazards 70(1):39–50

Gers FA, Schmidhuber J, Cummins F (2000) Learning to forget: Continual prediction with lstm. Neural Comput 12(10):2451–2471

Gong Y, Zhang Y, LanS WH (2016) A comparative study of artificial neural networks, support vector machines and adaptive neuro fuzzy inference system for forecasting groundwater levels near lake Okeechobee. Florida Water Resour Manag 30(1):375–391

Gunnink JL, Burrough PA (1996) Interactive spatial analysis of soil attribute patterns using exploratory data analysis (EDA) and GIS. In: Masser I, Salge F (eds) Spatial Analytical Perspectives on GIS. Taylor & Francis, New York, pp 87–99

Guzman SM, Paz JO, Tagert ML (2017) The use of NARX neural networks to forecast daily groundwater levels. Water Resour Manag 31(5):1591–603

Hochreiter S, Schmidhuber J (1997) Ltsm Neural Comput 9:1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

IPCC (2014) Climate change 2014. Synthesis report. Versióninglés, Climate Change Synthesis Report. Contribution of Working Groups i, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. https://doi.org/10.1017/CBO9781107415324

Jeong J, Park E (2019) Comparative applications of data-driven models representing water table fluctuations. J Hydrol 572:261–273

KumarS HS, Singhal DC (2011) Groundwater resources management through flow modeling in lower part of Bhagirathi - Jalangi interfluve, Nadia. West Bengal J Geol Soc India 78:587–598. https://doi.org/10.1007/s12594-011-0118-0

Lallahem S, Mania J, Hani A, Najjar Y (2005) On the use of neural networks to evaluate groundwater levels in fractured media. J Hydrol 307(1–4):92–111

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444

Lipton ZC, Berkowitz J, Elkan CA (2015) Critical review of recurrent neural networks for sequence learning. arXiv preprint arXiv:1506.00019

Mall R K, Gupta A, Singh R, Singh R, Rathore L S (2006) Water resources and climate change: An Indian perspective. Current science pp. 1610–1626

Mirzavand MGR (2015) A stochastic modelling technique for groundwater level forecasting in an arid environment using time series methods. Water Resour Manag 29(4):1315–1328

Natural Resources Management and Environment Department (NR) under Food and Agriculture Organization (FAO) of the United Nations (1998) Crop Evapotranspiration – Guidelines for Computing Crop Water Requirements. FAO Irrigation and Drainage Papers – 56.

Rakhshandehroo G, Akbari H, AfshariIgder M, Ostadzadeh E (2017) Long-Term Groundwater-Level Forecasting in Shallow and Deep Wells Using Wavelet Neural Networks Trained by an Improved Harmony Search Algorithm. J Hydrol Eng 23:04017058. https://doi.org/10.1061/(asce)he.1943-5584.0001591

Style G (2009) Application of artificial neural network in the field of Geohydrology, university of the free State, South Africa

Sudheer KP, Gosain AK, Ramasastri KS (2002) Adata-driven algorithm for constructing artificial neural network rainfall runoff models. Hydrol Process 16(6):1325–1330

Suryanarayana C, Sudheer C, Mahammood V, Panigrahi BK (2014) An integrated wavelet-support vector machine for groundwater level prediction in Visakhapatnam, India. Neurocomputing 145:324–335. https://doi.org/10.1016/j.neucom.2014.05.026

Wunsch A, Liesch T, Broda S (2020) Groundwater Level Forecasting with Artificial Neural Networks: A Comparison of LSTM, CNN and NARX. Hydrol Earth Syst Sci Discuss 23:1–23

Zhang J, Zhu Y, Zhang X, Ye M, Yang J (2018) Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. J Hydrol 561:918–929. https://doi.org/10.1016/j.jhydrol.2018.04.065

Zhu YY, Zhou HC (2009) Rough fuzzy inference model and its application in multi-factor medium and long-term hydrological forecast. Water Resour Manag 23(3):493–507