Modeling Earth Systems and Environment (2024) 10:579–594
https://doi.org/10.1007/s40808-023-01803-5

ORIGINAL ARTICLE

# Analysis of reservoir outflow using deep learning model

**Surendra Singh Choudhary[1] · S. K. Ghosh[1]**

Received: 23 November 2022 / Accepted: 28 May 2023 / Published online: 3 June 2023


## Abstract

Predicting reservoir outflow is crucial for managing water resources in under extreme flood and drought conditions. Time series study of reservoir outflow relies heavily on previous information on climate and reservoir factors. The Long Short Term Memory model of Deep Learning is applied using rainfall, rainfall intensity, runoff rate, temperature, surface water area, and reservoir outflow to predict reservoir outflow. This study summarizes the parameter setting effect on model performance and analyzes the main factors that affect reservoir outflow prediction. Monthly rainfall, rainfall intensity, runoff rate, temperature, outflow, and surface water area data are used in the multipurpose reservoir prediction model to analyze monthly and yearly water outflow of the reservoir. This system help in water management to reduce the risk of flooding downstream while ensuring sufficient water storage for monthly utilization, i.e., an outflow of a reservoir to the city. This method determines the appropriate amount of water released from the reservoir during the dry season and helps set a relationship with other input variables and outflow. The model has been trained and tested using the obtained data. The result analyzes that combined iterations and neurons of a hidden layer mainly impact manipulating the model precision; computation speed is primarily affected by the batch size of the model. The proposed model can simultaneously predict entire parameters in an accurate and efficient way.

**Keywords** Deep learning · Long short-term memory · Climatic parameters · Outflow-prediction

### Abbreviations

| | |
|---|---|
| ANN | Artificial neural network |
| RNN | Recurrent neural network |
| MLP | Multilayer perceptron |
| LSTM | Long short term memory |
| BPTT | Back propagation through time |
| GRU | Gated recurrent unit |

## Introduction

Reservoir variable prediction is a crucial task and challenge in hydrology. The outflow of reservoirs is high fuzziness, complex, and random. These variables are influenced by many uncertain variables such as rainfall, rainfall intensity, runoff, temperature, surface water area, water level, and characteristic topography (Sun 2020). Statistical methods, machine learning, and deep learning methods are available for predicting hydrological parameters (Matheussen, et al. 2019). Reservoir operation is based on operating functions. (Like linear, polynomial, sigmoid, neural network, fuzzy logic method, etc.)

The traditional way provides a range of reservoir outflow but does not provide an exact value of outflow. So at this stage, this method typically fails to interact with the influencing factors (Chaki et al. 2020; Paul et al. 2021). The influence factors consider natural variables, human needs and reservoir regulation capacity. Natural variables classified as flow, precipitation, temperature and evaporation. Human needs are categrazed in water demand, power generation and flood control. The reservoir regulation capacity are considered as total and regulation storage (Ren et al. 2020; Zhu et al. 2020; Qiu et al. 2021).

Science and Technology have recently regularly collected data acquisition, pre-processing, and the mass of reservoir operating data (Latif et al. 2022). New data mining methods, like artificial intelligence and neural network, provide a unique solution for reservoir operating decisions (He et al.

✉ Surendra Singh Choudhary
schoudhary1@ce.iitr.ac.in

S. K. Ghosh
sanjay.ghosh@ce.iitr.ac.in

[1] Civil Engineering, Indian Institute of Technology, Roorkee 247667, India

2021). The historical data of reservoir operation have more information and knowledge for reservoir managers, which supports the help of decisions to operating water inflow and outflow situation of reservoirs (Herbert et al. 2021; Gangrade et al. 2022). Therefore, using historical data of reservoir extracted reservoir operating rules using Artificial Intelligence (AI) algorithms provides a fast and effective operating scheme to deal with various flow scenarios under different hydrological periods (Chaves et al. 2004; Zhang et al. 2018). ANN contains adequate highly nonlinear complex system and is widely used in hydrological fields to handle complex modeling systems (Kao et al. 2020; García-Feal et al. 2022).

A problem with the RNN to handle extensive time interval data; Long Short-Term Memory has overcome data limitation (LSTM extension form to RNN (Al-Shabandar et al. 2020). Long sequences and vanishing gradients problem of RNN solve by LSTM using constant error flow within special memory cells. Its performs better time series variation data characteristics (Kratzert et al. 2018). Many broad areas have used Multilayer Perceptron (MLP), RNN, and solved many non-linear problems (Shi et al. 2015; Shen and Lawson 2021).

In the last decade, many machine learning and deep learning models have been generated and used in various water resources articles (Zhu et al. 2020; Huang et al. 2022). Previous studies have represented that machine learning is a powerful prediction tool for reservoir studies. However, the creation of machine learning models is hampered by the absence of full original data. (Wang et al. 2022) demonstrate that how process-based models can offer machine learning models with training data.

An impressive advanced ANN, primarily for temporal dynamics memorizing precious information using feedback connections settled in structure (Elman 1990). RNN structure and data training have been used in tasks with sequential inputs for different range, like time series modeling, machine translation, natural language processing, and reservoir operations. Thereupon, for modeling hydrological resources and time-series data of meteorological variables, RNNs are considered. Learning long-range dependencies is challenging with vanilla neural networks. The LSTM model resolves the vanishing and exploding gradients problem in the back propagation of the vanilla RNN model (Ni et al. 2020; Sharifi et al. 2021).

The most accepted form of sequential data is time-series data. Many studies explored the potentiality of LSTM on a sequence of time-series data forecasting and found better results with machine transcription (Zheng et al. 2022). Multiple layers are assigned for data processing that displays an abstraction of numerous levels. Multiple processing layers are composed in computation models with simulation with the human brain to understand numerous abstraction levels in the LSTM model of deep learning; these advanced techniques can systematically capture persistent data using advanced hidden layer units (Zhang et al. 2019; Fang et al. 2020). Traditional methods are inappropriate for conserving the correspondence of time-series information with a Fully Connected Neural Network (FCNN) to provide a model for predicting the reservoir outflow using the LSTM model. This model predicts the reservoir outflow from the relationship between input variables and parameters of model inputs (Hongliang et al. 2020). Its structure has a memory cell; information regulates into and out of the cell by non-linear gating units with the state over time (Greff et al. 2017). Conventional RNNs have less effective than LSTM networks. It was employed in the conjecture water table in agriculture and Feed-Forward Neural Network (FFNN), which has provided better results than FFNN. The results of multi-step advance prediction of time series, LSTM, and GRU were admirable to other models. (Kratzert et al. 2018) Investigate the potential of LSTM with many catchments and meteorological data for runoff modeling. The results show that, it has been provided more accurately than a well-settled benchmark model. They developed a conventional LSTM for casting precipitation adopting recognized maps of radar (Ni et al. 2020; Song et al. 2020). It is so useful for prediction to time-series analysis. Wind speed can be measured within a 5% error using LSTM. It also provides for medium to long-time prediction generation of power of photovoltaic. It gives impressive results for different applications; it can capture trend variation of data and the credence relationship characteristic of time-series information. The novelty of the LSTM neural network's perspective layout architecture enhances forecast to production (Yang et al. 2016). Complexity is only one drawback of LSTM. Significant research opportunities apply to this model in Civil, Computer science, and Electrical with simple structures (Zhang et al. 2019).

Several deep learning methods (Xu et al. 2021; Zarei et al. 2021) are available, which can achieve better prediction performance to represent reservoir variables. The study area is selected as Kaylana reservoir, which is located near Jodhpur city, Rajasthan. The reservoir provides water supply to the larger population of the city. Figure 1 shows the reservoir's location and base map. It has been generated using shape map region and the latest False Color Composite (FCC) satellite data image. Table 1 represents all details of the reservoir, as the capacity of the study area, surface water area, and catchment of the study area have been calculated from the toposheet.

Table 2 represents data used for model analysis. This study used the LSTM of RNN to evaluate reservoir operations using model parameters and predicted reservoir outflow. LSTM of RNN is proposed to predict the reservoir outflow using meteorological and reservoir variables. This study aims to analyze the performance of the model by assigning parameters to a model, predicting outflow

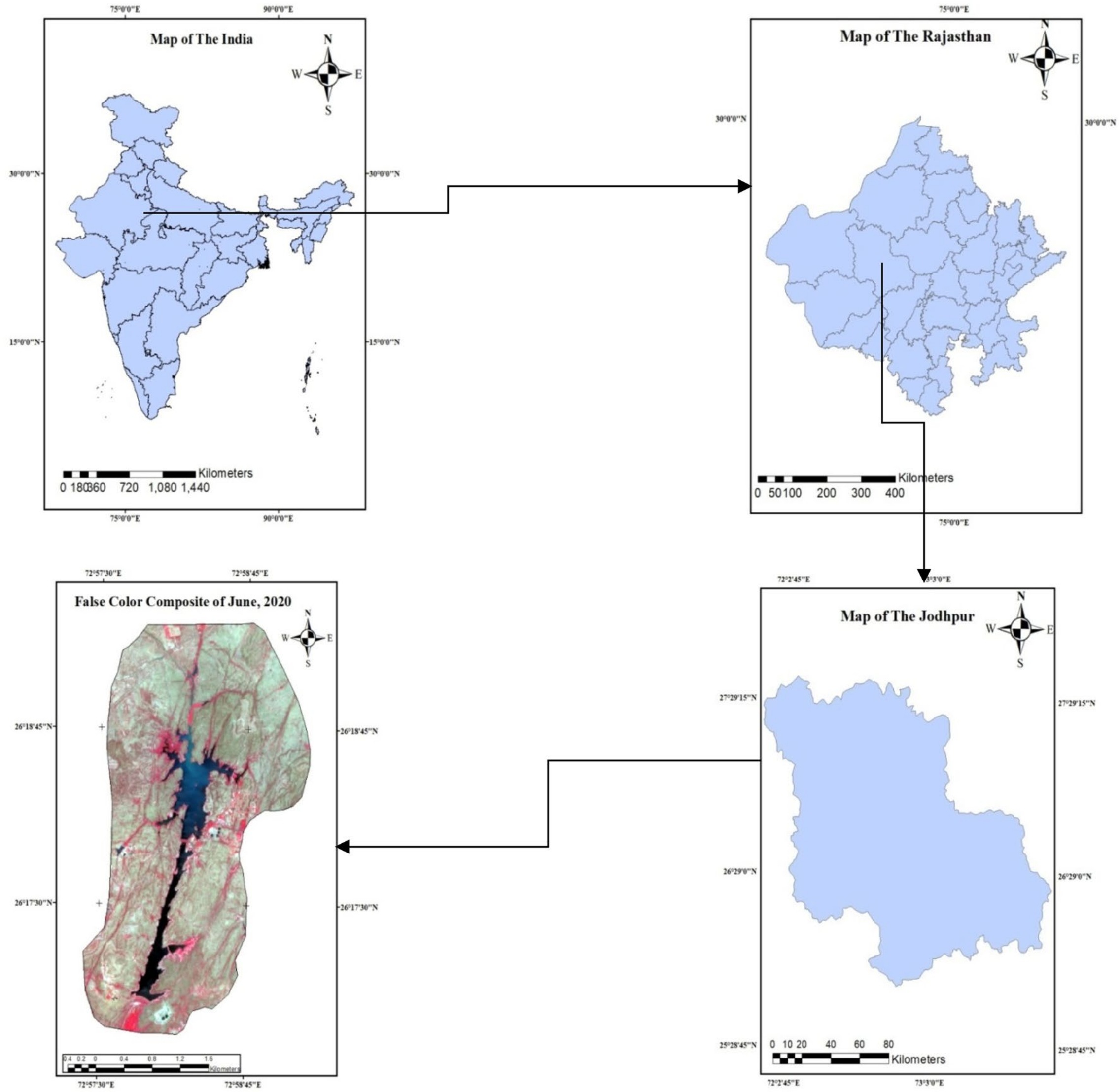



**Fig. 1** Base map of the study area (Source: Survey of India)

**Table 1** Specification of the Kaylana reservoir (Water Resource Department, Rajasthan)

| Capacity (MCFT) | Surface area ($Km^2$) | Catchment area ($Km^2$) |
|---|---|---|
| 324.75 | 1.00 | 11.791 |

using different variables of the reservoir, and comparative analysis of reservoir inputs variable to effect prediction of monthly discharge of a reservoir.

This study developed a model that has a new initial application of LSTM of deep learning to predict reservoir outflow. This created model also predicts other reservoir variables; it means to use to predict multiple variables simultaneously.

**Table 2** Details of input factors of the model

| Category | Factors | Output unit |
|---|---|---|
| Time | Months of year | – |
| Runoff, water level, and a surface water area | Runoff rate | $m^3/s$ |
| | Reservoir outflow | $Mm^3$ |
| | Surface water area (SWA) | $km^2$ |
| | Maximum control level (5ft) | ft |
| | Minimum control level (68ft) | ft |
| Meteorological information | Precipitation | mm |
| | Air temperature | $^{\circ}C$ |

## Methodologies

Many different machine learning methods and deep learning (Tut Haklidir and Haklidir 2020; Idrees et al. 2021; Jiang et al. 2022) have been used in previous studies to predict other monthly and yearly time series parameters. This study focuses on predicting water outflow of the reservoir using the latest LSTM. The LSTM deep learning neural network method is mainly used to predict time series variables in current decades. The RNN and LSTM Method have been described in this section.

### Recurrent neural network

The RNN model has three layers, an input layer, hidden neuron layers, and an output layer shown in Fig. 2. A hidden layer of RNN is used to continuously recurrent the model based on the input time sequence. For this study, three layers are used in a neural network architecture. The hidden layer with inputs and output functions are shown in Fig. 3

Neural networks have three layers: input layers, output layers, and a third hidden layer provided internally connected to the input and output layers. Input variables have been represented by ($x_1, x_2, \ldots, x_i, \ldots, x_{n1}$) outputs variable represented by ($y_1, y_2, \ldots, y_k, \ldots, y_{n3}$) and hidden node is represented by ($h_1, h_2, \ldots, h_j, \ldots, h_{n2}$). The model was trained with a trial and error method that gives better performance. Here weight coefficient matrix and offset are represented by $b$, and the activation function is shown by $f()$, learned and trained of a model by $h_t$ and $y_t$ functions. Hence RNN can better handle nonlinear time series; some problems still suffer from vanishing gradient problems with a large time-series dataset. Network errors are reproduced from the output to the input layer of the training phase (Al-Shabandar et al. 2020). Gradient loss and gradient explosion, an RNN may not take some desirable features, especially long-term dependencies (Zhu et al. 2020).

$$h_t = \mathrm{f}\left(\mathrm{W}_{\mathrm{xh}}\mathrm{x}_{\mathrm{t}} + \mathrm{W}_{\mathrm{hh}}\mathrm{h}_{\mathrm{t-1}} + \mathrm{b}_{\mathrm{h}}\right) \quad (1)$$

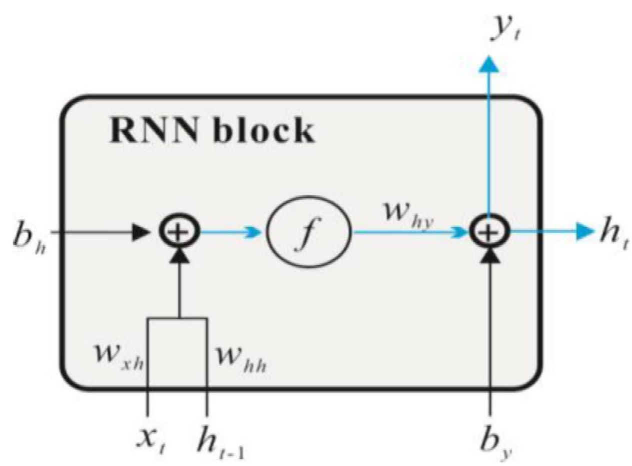


Fig. 3 Simple RNN of hidden layer (Zhang et al. 2019)

$$y_t = \mathrm{W}_{\mathrm{hy}}\mathrm{h}_{\mathrm{t}} + \mathrm{b}_{\mathrm{y}} \quad (2)$$

### Long short-term memory neural network

LSTM is one form of RNN that could handle long-term dependencies and resolve the vanishing gradient problem (Al-Shabandar et al. 2020). The special units of memory blocks of the hidden layers are involved in the LSTM structure. The self-connected and multiplicative cell units are parts of each memory block. The sensual state of networks is stored in memory cells of hidden layers. Multiplicative units of hidden layers have input, output, forget gates, and information regulation between the cells.The flow of inputs to the memory cell is control by the input gate, output flow of the activation cell conducts by the output gate (Fig. 4). Constitutional state of the cell is scaled by forget gate (Cheng et al. 2020).

The first step of LSTM decides by the data removal process from the cell state. The sigmoid layer (forget gate) is proposed to conclude the data, evacuating against the memory cell state. Value of $\mathrm{f_t}$ varies 0–1 generated using

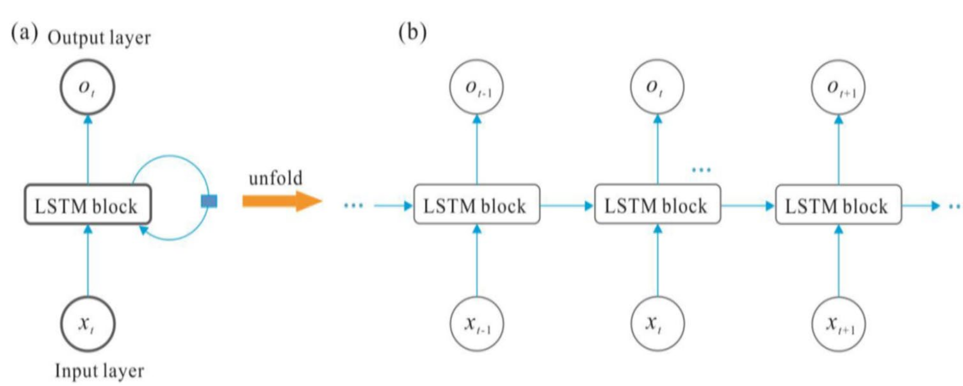


Fig. 2 RNN with folded computation (Zhang et al. 2019)

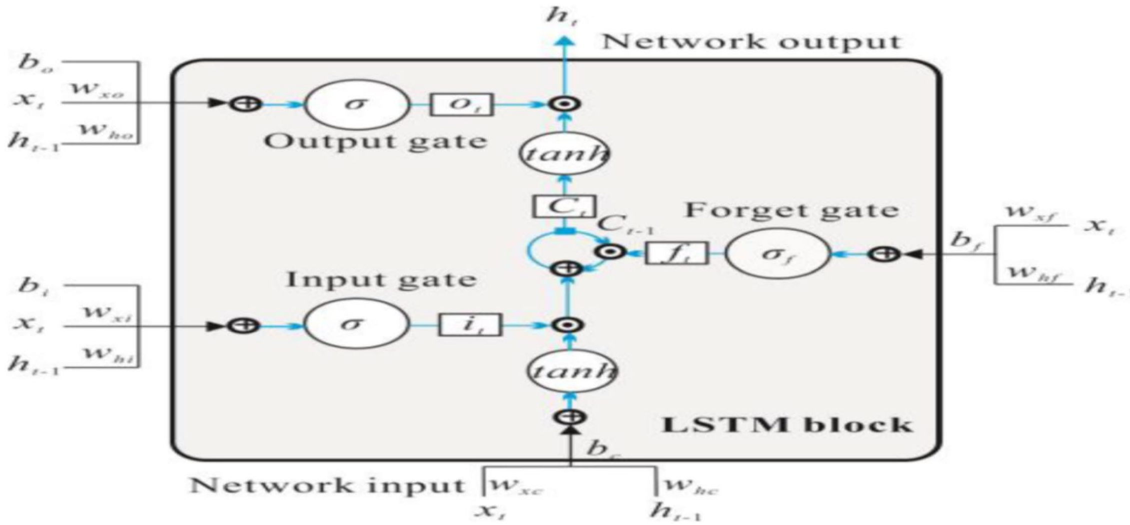



**Fig. 4** LSTM block structure (Zhang et al. 2019)

forward propagation equation in forget gate presented in Eq. 3 depends on output $h_{t-1}$ of the previous moment and current input $x_t$ to decide to pass or partially pass information $C_{t-1}$ generated during the last moment pass.

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + b_f\right) \tag{3}$$

The second step decides which information desire to store in the memory cell state. It can be seen that one part handles value updating, and the second part creates a vector of new contestant value $C_t$ so that it can be added in the cell state. The value of both amounts will be added for the update.

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{4}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tanh\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right) \tag{5}$$

In the last stage, the model output is computed. This output was initially calculated by the sigmoid function for the outcome, then resized the $C_t$ value by tanh and multiplied with sigmoid gate output.

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{6}$$

$$h_t = o_t \times \tanh\left(C_t\right) \tag{7}$$

Equations (6) and (7) portrays that f, i, C, o, and hare the forget gate, input gate, cell, output gate, and hidden output, respectively. σ and tanh indicates the gate activation function and the hyperbolic tangent activation function, respectively, while the weight coefficient matrix is represented by W.

Output coefficient $o_t$ controls to final output ($h_t$) indicate by output gate of the network. Long-term and short-term dependencies of time-series input information preclude the gradient depreciation or eruption of information-carrying. The central part of LSTM recognizes long-term memory use to preserve input data from every memory unit stage. All input data in the current moment represents in hidden layer states before it is stored. The network progressively squeezes all the information as time passes, so the hidden layer state consistently shows a vector of an assured length. Moreover, such indiscriminate compression may be a weak variation in time between inputs to an extent and maybe not retain crucial historical information. Hence, remarkable improvement is required to enhance the inequity of LSTM (Ding et al. 2020).

The training progress of the model utilized the Back Propagation Through Time (BPTT). The necessary training process adopts the following steps initially hidden layer output computed by forwarding the computation process. In the forward step, hidden layer errors are computed using backpropagation of the time step and network architecture sequence. In the last Adaptive Moment Estimation (Adam) was used to update the weight coefficient. In LSTM, the iterations are fundamental parameters that affect the model's performance as alike in BPNN. Batch size is also a supporting parameter that affects the performance of the model. After that, the model updates the network's parameters for error computation between observed and expected output. Overall, LSTM is a part of RNN used to generate a model for the prediction of time series data of reservoir operation.

## Proposed model

In this study, monthly operating variables of Kaylana reservoir from January 2014 to May 2020 (77 records of data) were collected from the Water Resource Department, Rajasthan.

The operational data includes the outflow and surface water area, and the meteorological data consists of rainfall, rainfall intensity, runoff rate, and atmospheric temperature. The overall flow chart of the proposed model has shown in Fig. 5.

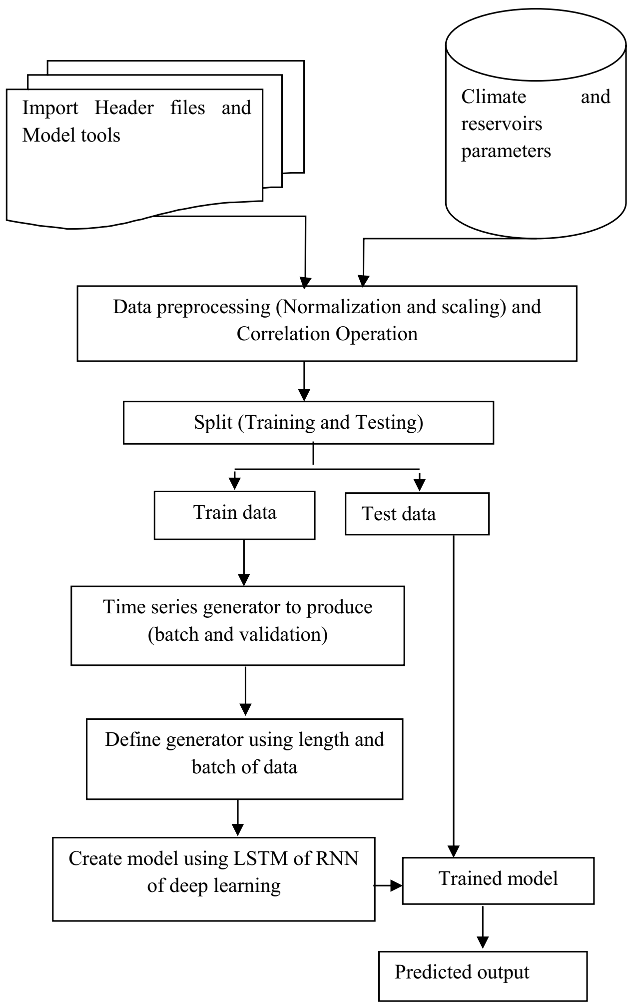



**Fig. 5** Proposed Methodology Chart

The regular split training and testing rules are used to test the model for the data collected from June 2019 to May 2020 and are used for training. The monthly reservoir variables are used as model input (decision variable). Input factors of the model are represented in Table 2. The output of the model is performance, predicted and actual outflow, and accuracy assessment using different error analyses of multi variance features. The various combinations of input parameters of a model for finding out accuracy, performance, speed, and predicted output of the model are shown in Table 3. Model performance has been affected by the batch size, iterations, and neurons of hidden nodes, but calculation speed is primarily influenced by the batch size (Zhu et al. 2020).

First, the batch size is fixed and then executed with the iteration and neurons of hidden layers for simulation completion of the model. The number of iterations and hidden neurons range from 20 to 500 with a frequency of 20 for 25 combinations of steps. Due to the previous step's practice, affect the parameters setting of the model on the performance set batch size range is from 1 to 38 with an interval of 4 for 11 parameters combinations.

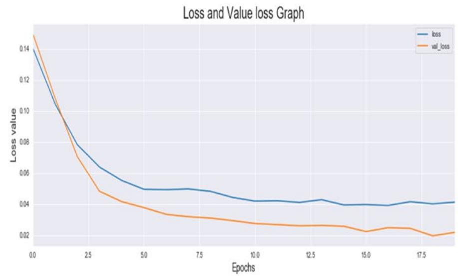
From the obtained results of previous steps, batch factors are adjusted to approval for the batch size's effect on model achievement. The range of batch size is set 1–38, a difference of 4, for 11 combinations of parameters. At the same time, the training frequency is set to 10 to reduce the random error.

## Results and discussion

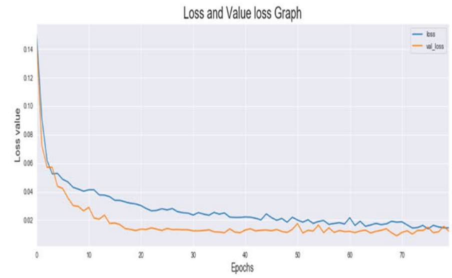
The deep learning model working function mainly depends on hidden nodes and activation functions used in a model to forecast and predict time series data. Input–output layers and hidden layer neurons depend on the type of problem (Navale and Mhaske 2022). This study used LSTM model to predict the reservoir outflow. It is helpful in removing the vanishing gradient problem in backpropagation in the recurrent process

**Table 3** Different combinations of input parameters of the model for accuracy and performance

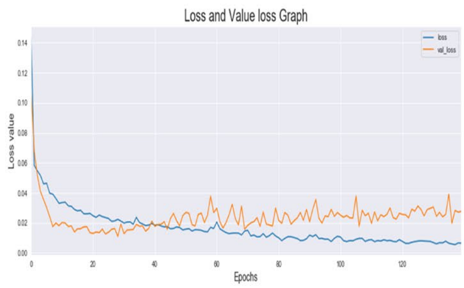
| Data length | Batch size | Hidden layer neurons | No. of iterations (epochs) |
|---|---|---|---|
| 6 | 1, 4, 8, 14, 18, 22, 26, 30, 34, 38 | 100 | 300 |
| 6 | 2 | 20, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900 | 300 |
| 6 | 2 | 100 | 50, 100, 300, 500, 1000, 2000, 3000 |
| 6 | 4 | 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 420, 440, 460, 480, 500 | |



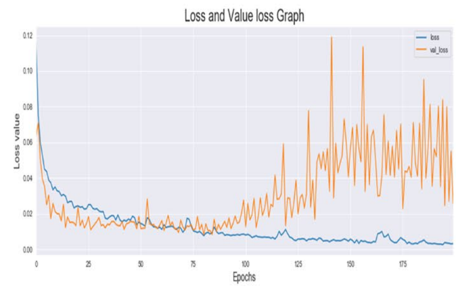

(a). {Inputs=6, Batch size=4, Hidden neurons=20, Epochs=20}



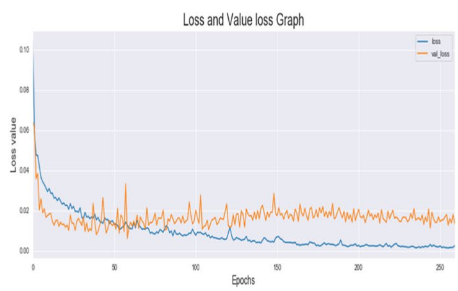

(b). {Inputs=6, Batch size=4, Hidden Neurons=80, Epochs=80}



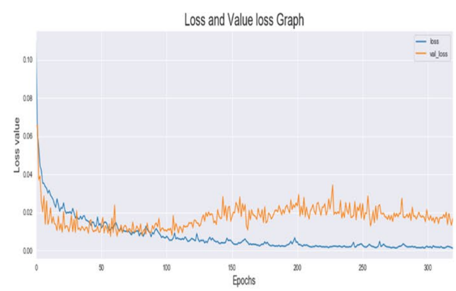

(c). {Inputs=6, Batch size=4,Hidden neurons=140, Epochs=140}



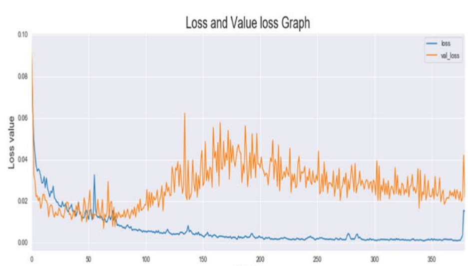

(d). {Inputs=6, Batch size=4,Hidden neurons=200, Epochs=200}



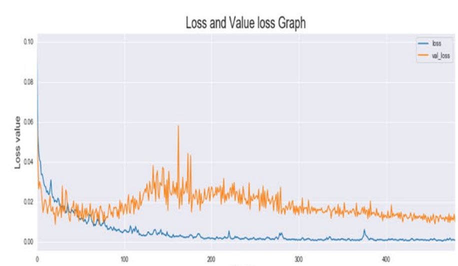

(e). {Inputs=6, Batch size=4,Hidden neurons=260, Epochs=260}




(f). {Inputs=6, Batch size=4,Hidden neurons=320, Epochs=320}



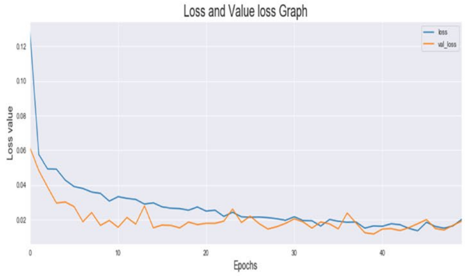

(g). {Inputs=6, Batch size=4,Hidden neurons=380, Epochs=380}




(h). {Inputs=6, Batch size=4,Hidden neurons=480, Epochs=480}

**Fig. 6** Progress of the model varies with the combination of hidden nodes and epochs

(Yang et al. 2017). This study analyzes the accuracy and execution time of the LSTM model with input parameters in the parameter sensitivity section of the result. Subsequently, the model performance index was also evaluated using the calculated mean square error of input parameter combinations, i.e., iteration, hidden neurons, and batch size of the model. Further analyzed outflow prediction of a reservoir with input factors of model means whenever an increased



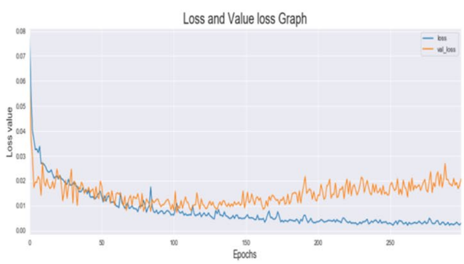

(a). {Inputs=6, Batch size=2,Hidden neurons=100, Epochs=50}


(b). {Inputs=6, Batch size=2,Hidden neurons=100, Epochs=500}


(c). {Inputs=6, Batch size=2, Hidden neurons=100, Epochs=1500}

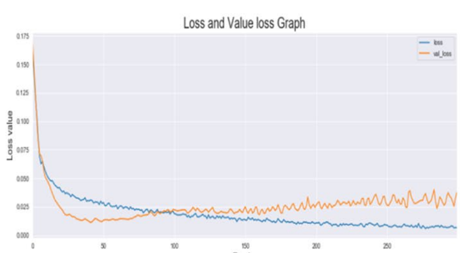

(d). {Inputs=6, Batch size=2,Hidden neurons=100, Epochs=3000}

**Fig. 7** Performance of the model varies with epochs

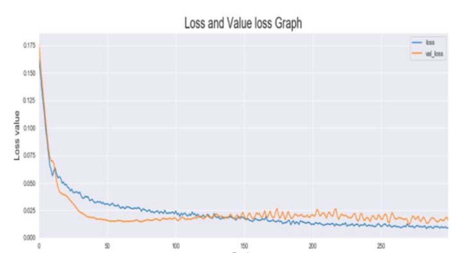



(a). {Inputs=6, Batch size=1,Hidden neurons=100, Epochs=300}


(b). {Inputs=6, Batch size=8, Hidden neurons=100, Epochs=300}




(c). {Inputs=6, Batch size=26, Hidden neurons=100, Epochs=300}

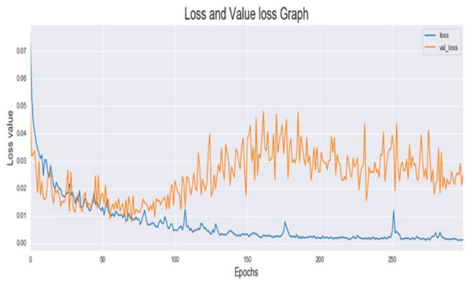



(d). {Inputs=6, Batch size=34, Hidden neurons=100, Epochs=300}

**Fig. 8** Performance of the model varies with batch size

(a). {Inputs=6, Batch size=2, Hidden neurons=20, Epochs=300}

(b). {Inputs=6, Batch size=2, Hidden neurons=100, Epochs=300}

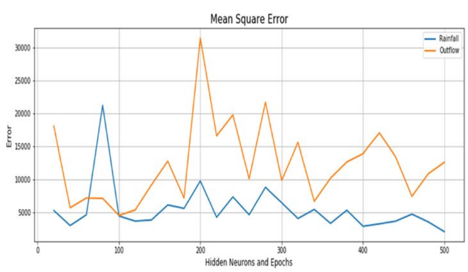


(c). {Inputs=6, Batch size=2, Hidden neurons=400, Epochs=300}

(d). {Inputs=6, Batch size=2, Hidden neurons=800, Epochs=300}

**Fig. 9** Performance of the model varies with hidden nodes

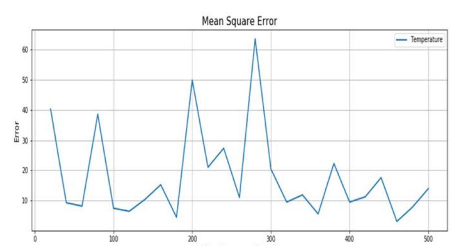
Input parameters (hidden neurons and epochs)



(a) Rainfall and outflow of the reservoir

(b) Rainfall intensity and runoff

(c). Surface water area



(d). temperature

**Fig. 10** Mean square error (MSE) with input parameters (hidden neurons and epochs)

Input parameters (hidden neurons)


(a) Rainfall and outflow of the reservoir


(b). Rainfall intensity and runoff

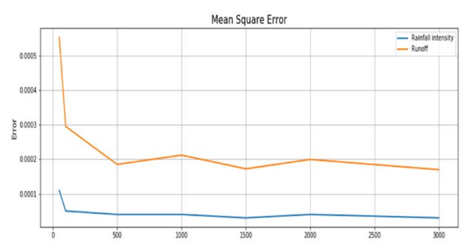

(c). Surface water area


(d). temperature

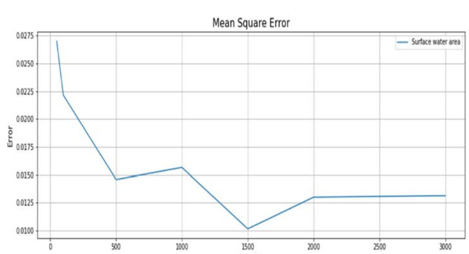
**Fig. 11** Mean square error (MSE) with input parameters (hidden neurons)

Input parameters (epochs)

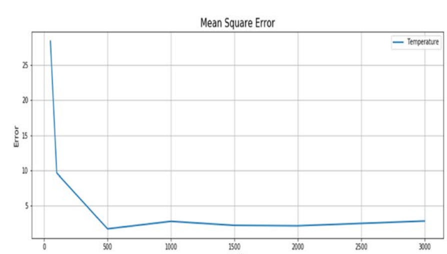

(a) Rainfall and outflow of the reservoir




(b). Rainfall intensity and runoff



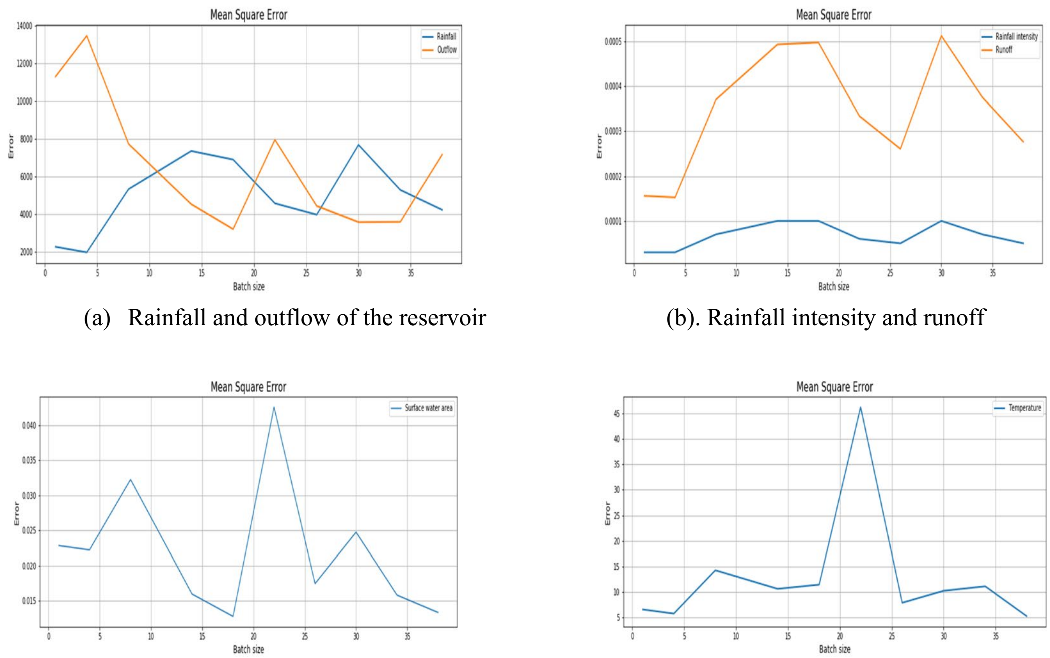

(c). Surface water area




(d). temperature

**Fig. 12** Mean square error (MSE) with input parameters (epochs)

Input parameters (batch size)




(a) Rainfall and outflow of the reservoir

(b). Rainfall intensity and runoff

(c). Surface water area

(d). temperature

**Fig. 13** Changes in mean square error (MSE) with input parameters (batch size)

combination of hidden neurons, epochs, and batch size of the model, the outflow prediction of the reservoir will vary with input factors.

## Parameter sensitivity

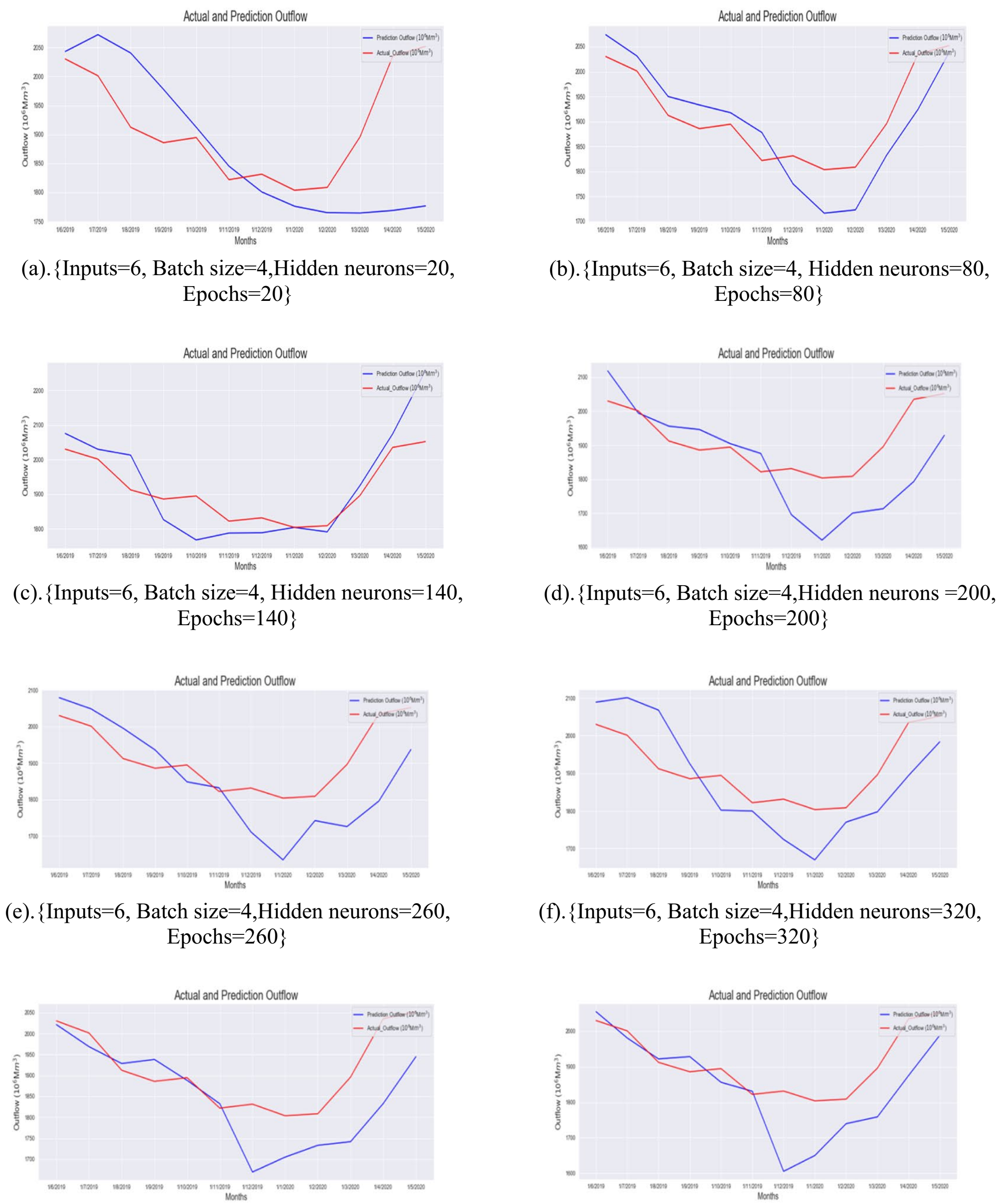
Initially, the model's progress is evaluated using the influence of iterations and hidden nodes. The optimization performance of different parameters of the LSTM model is shown in (Fig. 6, 7, 8, and 9). As a result, when the increase of iterations and hidden layer neurons, the accuracy of the model increases at a certain level, and the model's execution time also increases (Fig. 6). Further, suppose an increase in the number of epochs, there is no improved model accuracy and precision, but the calculation time has been increased. In that case, i.e., the model works efficiently on a lesser number of iterations (Fig. 7).

When increases the combined epochs and hidden layer neurons, the accuracy of the model decreases. The model is accurate when epochs and neurons are more than 200 (Fig. 6). The need for epochs is less to a model convergence when the hidden neuron is less than 20 (Fig. 6).

If the iterations are less than 50, the model gets more precision (Fig. 7). The accuracy of the model and computational speed are analyzed with the batch size (Fig. 8). Epochs are set to 300, and the hidden nodes are set to 100 for the model to exact prediction output. The model's computation speed increases and is mainly affected by the batch sizes (Fig. 8). Previous research studies (Zhang et al. 2019; Zhu et al. 2020) have proved that the model's computational speed is affected by the batch size.

The effect of batch size on model performance is continuously increasing (Fig. 8). Model performance, precision, and computation speed have more variation with an increased number of hidden neurons on the model (Fig. 9). Further investigation shows that forwarding algorithms calculate output, and each hidden cell's error is computed using backward algorithms with no further significant improvement after reaching training at a particular limit. Problems were faced in the effect of hidden nodes on model precision in the research of ANN (Yao 1999; Lv et al. 2020). Hidden nodes are essential parameters whenever accuracy is considered. The selection criterion of the number of hidden layer nodes mostly depends on nodes of input and output layers. For a specific application, optimal hidden nodes selection is based on the trial and error approach (Zhang et al. 2018).

For deep learning models, changes in epochs (iterations) analyze the model precision, as hidden neurons are weakly affected. The accuracy of the model is increased with the increase of the batch in a combination of parameters, but batch size mostly affect the model's computational speed when batch size results in a faster computational rate.




(a). {Inputs=6, Batch size=4,Hidden neurons=20, Epochs=20}

(b). {Inputs=6, Batch size=4, Hidden neurons=80, Epochs=80}

(c). {Inputs=6, Batch size=4, Hidden neurons=140, Epochs=140}

(d). {Inputs=6, Batch size=4,Hidden neurons =200, Epochs=200}

(e). {Inputs=6, Batch size=4,Hidden neurons=260, Epochs=260}

(f). {Inputs=6, Batch size=4,Hidden neurons=320, Epochs=320}

(g). {Inputs=6, Batch size=4,Hidden neurons=380, Epochs=380}

(h). {Inputs=6, Batch size=4,Hidden neurons=440, Epochs=440}

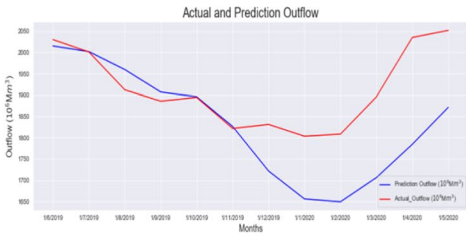
**Fig. 14** Comparison of predicted and observed outflow using the invariance of combined hidden neurons and epochs of theLSTM algorithm

Although, large batch size may be a source of a native best solution for the model but affects the model's accuracy.

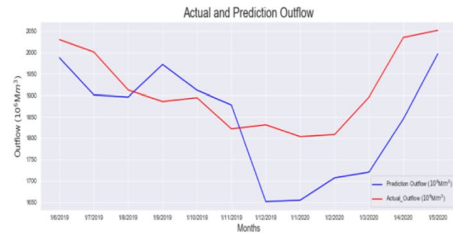
A combination of parameter values is prescribed to a convergence of model and need for better prediction of time series input variables.

## Model performance index

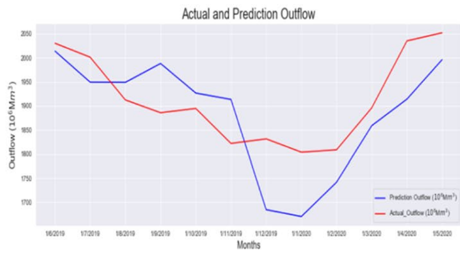
The model progress was assessed by assigning parameters of the model with the Mean Square Error (MSE) method. Model input is given to 6, the number of epochs is set to fixed with 300, the batch size is kept to 4, and hidden nodes are considered 100.



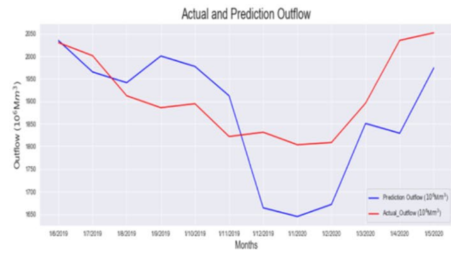

(a). {Inputs=6, Batch size=2,Hidden neurons=100, Epochs=50}




(b). {Inputs=6, Batch size=2, Hidden neurons=100, Epochs=500}

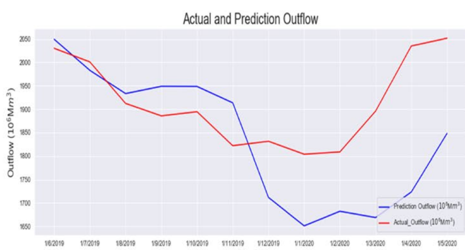



(c). {Inputs=6, Batch size=2, Hidden neurons=100, Epochs=1500}

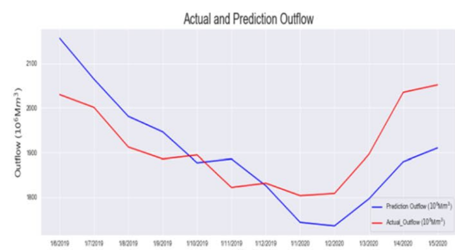



(d). {Inputs=6, Batch size=2,Hidden neurons=100, Epochs=3000}

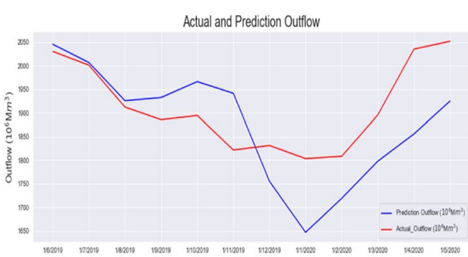
**Fig. 15** Comparison of predicted and observed outflow using invariance of epochs of an LSTM algorithm



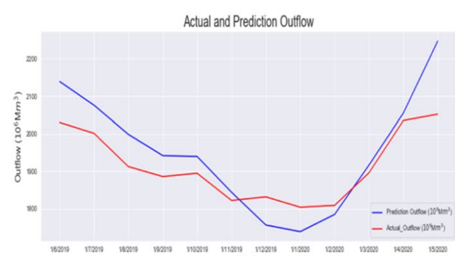

(a) {Inputs=6, Batch size=1, Hidden neurons=100, Epochs=300}




(b). {Inputs=6, Batch size=8, Hidden neurons=100, Epochs=300}




(c). {Inputs=6, Batch size=26, Hidden neurons=100, Epochs=300}

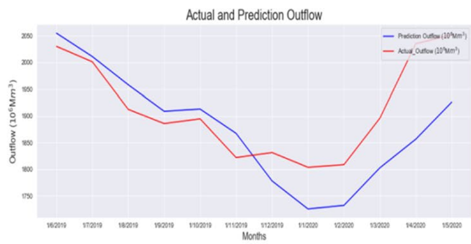



(d). {Inputs=6, Batch size=34, Hidden neurons=100, Epochs=300}

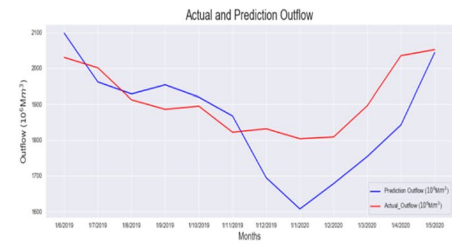
**Fig. 16** Correlations of predicted and observed outflow applying invariance of a batch size of an LSTM algorithm

The analyzes show that the decreased MSE with an increased combination of iterations and hidden neurons in predicting the reservoir's outflow and other parameters raises to a certain level. It remains constant or decreases, as seen in Fig. 10.

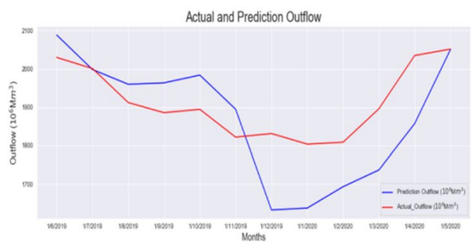
Subsequently, MSE decreases or is less variant when increasing hidden neurons Fig. 11. When epochs increase



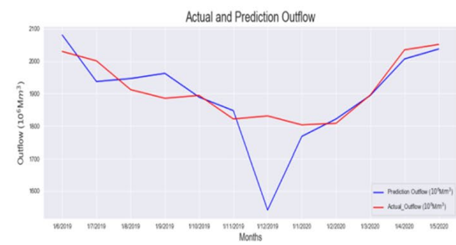

(a). {Inputs=6, Batch size=2, Hidden neurons=20, Epochs=300}




(b). {Inputs=6, Batch size=2, Hidden neurons=100, Epochs=300}




(c). {Inputs=6, Batch size=2, Hidden neurons=400, Epochs=300}




(d). {Inputs=6, Batch size=2, Hidden neurons=800, Epochs=300}

**Fig. 17** Comparison of predicted and observed outflow using invariance of neurons of the LSTM algorithm

up to 500, both errors will continuously decrease, but both are constant, as seen in Fig. 12. MSE is constant or decreases with increased batch size in all input parameters cases Fig. 13.

## Analysis of prediction of outflow with input factors

Reservoir operations are affected by many factors. The proposed study uses deep learning concepts to develop a reservoir operation model using the reservoir and meteorological data variables.

The results obtained show the performance optimization of prediction of the reservoir's outflow using variables and parameters of network models is shown in (Fig. 14, 15, 16, and 17). The results show the combined increased iterations and hidden neurons, the accuracy of outflow prediction increases at a certain level with a respective combination of iteration (epochs) and hidden neurons, i.e., from 20 to 200, the time consumption also increases after that model frequently predicted outflow of a reservoir (Fig. 14). Figure 15 shows that increased iterations improve the precision and accuracy of the model and time consumption also increase, i.e., model efficiently predicted outflow on all iterations from low to high. The batch size effect on expected outflow performance continuously grows with a batch size of 38 (Fig. 16). Subsequently analyzed that model precision and computation speed varied irregularly with growing hidden neurons (Fig. 17). Now it's a fact that changes and prediction of outflow of the reservoir mostly depend on the model's selection process.

A sensitivity analysis of the relationship was also carried out for the rest of the influence of decision variables on reservoir outflow prediction. The sensitivity of model performance with a change of input variables is analyzed by sensitivity, after modeling applied procedure of sensitive analysis of the model with changes of the model's input parameters. The influencing factors are mostly the time information, water level, and meteorological variables for climatic details (Table 1).

This study evaluates model progress and performance using the influence of input parameters iterations and hidden neurons, and mean square error (MSE). The accuracy variation of a model is also evaluated using combinations of input parameters.

The computation speed of a model is affected by the batch size and the convergence speed of the model described by a combination of input parameters (Zhang et al. 2019). Further analyzed performance optimization of the reservoir outflow using variables and parameters of model inputs. Accuracy and performance of outflow prediction increased overall with batch size compared to other parameters. In this study, more information has been given by runoff, water level, and meteorological conditions of data to predict the outflow of the reservoir. Reservoir outflow mostly changes by selected inputs decision variables. Consideration of the degree of selection of input factors depends on the seasons of the year. However, the main sensitive input factors for predicting reservoir output are runoff, water level, and meteorological parameters. Models are quickly understood and give a fast response to the prediction of input dependent variables. In neural network models, a better understanding of training and updating, deep learning models easily get acquainted with the training and testing of new input variables and continuously enhance model performance.

## Conclusions

Prediction reservoir outflow is very crucial for management of water demand of city. In this study, the LSTM model of deep learning is identified, trained and tested to prediction outflow of reservoir using climate and reservoir variables. The proposed model is analyzes the prediction of outflow of reservoirs using the LSTM model of deep learning. Many input variables and combinations of input parameters of the model are considered in this study to predict reservoir outflow. In result section it has been analyzed from the experimentation that model performance mostly depends on input variables and combinations of the selection of parameters. Model accuracy is very much improved with an optimized combination of the number of neurons and epochs.

Computation speed also is improved using a combination of epochs and hidden layer neurons. The computational speed of the model is mostly affected by the batch size inputs of the model. If batch size increase more than limit then also increases precision and decreases the accuracy of the model. Further, analyses reveal that the model's accuracy increases when increased epochs (iterations), but precision is randomized after specific periods.

This study proved that the LSTM of the RNN model has been used to predict the reservoir's outflow and may be used to predict many variables simultaneously, like rainfall, temperature, surface water area. It further proved that only LSTM of RNN is sufficient and better predict reservoir variables using input variables of the model and other independent reservoir variables. It is also proved that when a number of variables are more output of prediction is better than to fewer variables. The benefit of this model generated using LSTM is that it may require less training data set to train the model to predict input variables. It is also helpful to the operators of a reservoir to understand the relationship between reservoir and climate variables. The study gives a brief idea to select combinations of key input parameters to build a new model using a base model of LSTM for better predictions of input variables. It is also provide concepts of how to predict multiple variables simultaneously.

**Acknowledgements** I am sincerely thankful to the Department of Water Resources, Jodhpur, Rajasthan, for providing related data on the Kaylana reservoir for this study.

**Data availability statement** Some or all data, models, or code generated or used during the study are proprietary or confidential and may only be provided with restrictions.

## Declarations

**Conflict 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

Al-Shabandar R, Jaddoa A, Liatsis P, Hussain AJ (2020) A deep gated recurrent neural network for petroleum production forecasting. Mach Learn Appl. https://doi.org/10.1016/j.mlwa.2020.100013

Chaki S, Zagayevskiy Y, Shi X, Wong T, Noor Z (2020) Machine learning for proxy modeling of dynamic reservoir systems: deep neural network dnn and recurrent neural network RNN applications. Int Pet Technol Conf. https://doi.org/10.2523/IPTC-20118-MS

Chaves P, Tsukatani T, Kojiri T (2004) Operation of storage reservoir for water quality by using optimization and artificial intelligence techniques. Math Comput Simul 67:419–432

Cheng M, Fang F, Kinouchi T, Navon IM, Pain CC (2020) Long lead-time daily and monthly streamflow forecasting using machine learning methods. J Hydrol 590:125376. https://doi.org/10.1016/j.jhydrol.2020.125376

Ding Y, Li Z, Zhang C, Ma J (2020) Prediction of ambient PM2. 5 concentrations using a correlation filtered spatial-temporal long short-term memory model. Appl Sci 10(1):14

Elman JL (1990) Finding structure in time. Cogn Sci 14(2):179–211. https://doi.org/10.1207/s15516709cog1402_1

Fang Z, Wang Y, Peng L, Hong H (2020) Predicting flood susceptibility using long short-term memory (LSTM) neural network model. J Hydrol. https://doi.org/10.1016/j.jhydrol.2020.125734

Gangrade S, Lu D, Kao SC, Painter SL (2022) Machine learning assisted reservoir operation model for long-term water management simulation. JAWRA J Am Water Resour Assoc. https://doi.org/10.1111/1752-1688.13060

García-Feal O, González-Cao J, Fernández-Nóvoa D, Astray Dopazo G, Gómez-Gesteira M (2022) Comparison of machine learning techniques for reservoir outflow forecasting. Nat Hazard 22(12):3859–3874. https://doi.org/10.5194/nhess-22-3859-2022

Greff K, Srivastava RK, Koutník J, Steunebrink BR, Schmidhuber J (2017) LSTM: a search space odyssey. IEEE Trans Neural Networks Learn Syst 28(10):2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924

He S, Gu L, Tian J, Deng L, Yin J, Liao Z, Hui Y (2021) Machine learning improvement of streamflow simulation by utilizing remote sensing data and potential application in guiding reservoir operation. Sustainability 13(7):3645. https://doi.org/10.3390/su13073645

Herbert ZC, Asghar Z, Oroza CA (2021) Long-term reservoir inflow forecasts: enhanced water supply and inflow volume accuracy using deep learning. J Hydrol 601:126676. https://doi.org/10.1016/j.jhydrol.2021.126676

Hongliang WANG, Longxin MU, Fugeng SHI, Hongen DOU (2020) Production prediction at ultra-high water cut stage via recurrent neural network. Pet Explor Dev 47(5):1084–1090. https://doi.org/10.1016/S1876-3804(20)60119-7

Huang I, Chang MJ, Lin GF (2022) An optimal integration of multiple machine learning techniques to real-time reservoir inflow

forecasting. Stoch Env Res Risk Assess 36(6):1541–1561. https://doi.org/10.1007/s00477-021-02085-y
Idrees MB, Jehanzaib M, Kim D et al (2021) Comprehensive evaluation of machine learning models for suspended sediment load inflow prediction in a reservoir. Stoch Environ Res Risk Assess 35:1805–1823. https://doi.org/10.1007/s00477-021-01982-6
Jiang D, Xu Y, Lu Y, Gao J, Wang K (2022) Forecasting water temperature in cascade reservoir operation-influenced river with machine learning models. Water 14(14):2146. https://doi.org/10.3390/w14142146
Kao IF, Zhou Y, Chang LC, Chang FJ (2020) Exploring a long short-term memory based encoder-decoder framework for multi-step-ahead flood forecasting. J Hydrol. https://doi.org/10.1016/j.jhydrol.2020.124631
Kratzert F, Klotz D, Brenner C, Schulz K, Herrnegger M (2018) Rainfall-runoff modelling using long-short-term-memory (LSTM) networks. Hydrol Earth Syst Sci 22(11):6006–6022. https://doi.org/10.5194/hess-22-6005-2018
Latif SD, Birima AH, Ahmed AN, Hatem DM, Al-Ansari N, Fai CM, El-Shafie A (2022) Development of prediction model for phosphate in reservoir water system based machine learning algorithms. Ain Shams Eng J 13(1):101523. https://doi.org/10.1016/j.asej.2021.06.009
Lv N, Liang X, Chen C, Zhou Y, Li J, Wei H, Wang H (2020) A long short-term memory cyclic model with mutual information for hydrology forecasting: a case study in the Xixian basin. Adv Water Resour. https://doi.org/10.1016/j.advwatres.2020.103622
Matheussen BV, Granmo OC, Sharma J (2019) Hydropower optimization using deep learning. International conference on industrial, engineering and other applications of applied intelligent systems. Springer, Cham, pp 110–122. https://doi.org/10.1007/978-3-030-22999-3_11
Navale V, Mhaske S (2022) Artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) model for Forecasting groundwater level in the Pravara River Basin, India. Model Earth Syst Environ. https://doi.org/10.1007/s40808-022-01639-5
Ni L, Wang D, Singh VP, Wu J, Wang Y, Tao Y, Zhang J (2020) Streamflow and rainfall forecasting by two long short-term memory-based models. J Hydrol 583:124296. https://doi.org/10.1016/j.jhydrol.2019.124296
Paul T, Raghavendra S, Ueno K, Ni F, Shin H, Nishino K, Shingaki R (2021) Forecasting of reservoir inflow by the combination of deep learning and conventional machine learning. 2021 international conference on data mining workshops (ICDMW). IEEE, pp 558–565. https://doi.org/10.1109/ICDMW53433.2021.00074
Qiu R, Wang Y, Rhoads B, Wang D, Qiu W, Tao Y, Wu J (2021) River water temperature forecasting using a deep learning method. J Hydrol 595:126016. https://doi.org/10.1016/j.jhydrol.2021.126016
Ren T, Liu X, Niu J, Lei X, Zhang Z (2020) Real-time water level prediction of cascaded channels based on multilayer perception and recurrent neural network. J Hydrol. https://doi.org/10.1016/j.jhydrol.2020.124783
Sharifi MR, Akbarifard S, Qaderi K et al (2021) Comparative analysis of some evolutionary-based models in optimization of dam reservoirs operation. Sci Rep 11:15611. https://doi.org/10.1038/s41598-021-95159-4
Shen C, Lawson K (2021) Applications of deep learning in hydrology. Deep Learn Earth Sci Compr Approach Remote Sens Clim Sci Geosci. https://doi.org/10.1002/9781119646181.ch19
Shi X et al (2015) Convolutional LSTM network: a machine learning approach for precipitation nowcasting. Adv Neural Inf Process Syst 28:802–810
Song X, Liu Y, Xue L, Wang J, Zhang J, Wang J, Cheng Z (2020) Time-series well performance prediction based on long short-term memory (LSTM) neural network model. J Pet Sci Eng 186:106682. https://doi.org/10.1016/j.petrol.2019.106682
Sun AY (2020) Optimal carbon storage reservoir management through deep reinforcement learning. Appl Energy 278:115660. https://doi.org/10.1016/j.apenergy.2020.115660
Tut Haklidir FS, Haklidir M (2020) Prediction of reservoir temperatures using hydrogeochemical data, Western Anatolia geothermal systems (Turkey): a machine learning approach. Nat Resour Res 29(4):2333–2346. https://doi.org/10.1007/s11053-019-09596-0
Wang L, Xu B, Zhang C, Fu G, Chen X, Zheng Y, Zhang J (2022) Surface water temperature prediction in large-deep reservoirs using a long short-term memory model. Ecol Indic 134:108491. https://doi.org/10.1016/j.ecolind.2021.108491
Xu W, Meng F, Guo W, Li X, Fu G (2021) Deep reinforcement learning for optimal hydropower reservoir operation. J Water Resour Plan Manag. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001409
Yang T, Gao X, Sorooshian S, Li X (2016) Simulating California reservoir operation using the classification and regression-tree algorithm combined with a shuffled cross-validation scheme. Water Resour Res 2016(52):1626–1651
Yang Y, Dong J, Sun X, Lima E, Mu Q, Wang X (2017) A CFCC-LSTM model for sea surface temperature prediction. IEEE Geosci Remote Sens Lett 15(2):207–211. https://doi.org/10.1109/LGRS.2017.2780843
Yao X (1999) Evolving artificial neural networks. Proc IEEE 87(9):1423–1447
Zarei M, Bozorg-Haddad O, Baghban S, Delpasand M, Goharian E, Loáiciga HA (2021) Machine-learning algorithms for forecast-informed reservoir operation (FIRO) to reduce flood damages. Sci Rep 11(1):1–21. https://doi.org/10.1038/s41598-021-03699-6
Zhang D, Lindholm G, Ratnaweera H (2018) Use long short-term memory to enhance internet of things for combined sewer overflow monitoring. J Hydrol 556:409–418. https://doi.org/10.1016/j.jhydrol.2017.11.018
Zhang D, Peng Q, Lin J, Wang D, Liu X, Zhuang J (2019) Simulating reservoir operation using a recurrent neural network algorithm. Water 11(4):865. https://doi.org/10.3390/w11040865
Zheng Y, Liu P, Cheng L, Xie K, Lou W, Li X, Zhang W (2022) Extracting operation behaviors of cascade reservoirs using physics-guided long-short term memory networks. J Hydrol Reg Stud 40:101034. https://doi.org/10.1016/j.ejrh.2022.101034
Zhu S, Hrnjica B, Ptak M, Choiński A, Sivakumar B (2020) Forecasting of water level in multiple temperate lakes using machine learning models. J Hydrol. https://doi.org/10.1016/j.jhydrol.2020.124819