Using LSTM GRU and Hybrid Models for Streamflow Forecasting

Abdullahi Uwaisu Muhammad, Xiaodong Li(✉), and Jun Feng

College of Computer and Information, Hohai University, Nanjing, China
uwaisabdullahi87@yahoo.com,
{xiaodong.li, fengjun}@hhu.edu.cn

Abstract. Forecasting streamflow discharge have economic impact as well as reducing the effects of floods in flood prone regimes by presenting early warning. To minimize it's effects in these regimes, a powerful class of machine learning algorithms called long short-term memory (LSTM) and gated recurrent units (GRU) models, which have become popular in time series forecasting, because they are explicitly designed to avoid the long-term dependency problems is applied. LSTM and GRU models have also demonstrated their capacity in sequence modelling, speech recognition and streamflow forecasting. In this paper we proposed a hybrid model for streamflow forecasting using 35 consecutive years Model Parameter Estimation Experiment (MOPEX) data set of ten basins having different basin characteristics from different climatic regions in United States. The proposed hybrid model's performance is compared to the conventional LSTM and GRU models. Our experiments on the 10 MOPEX's river basins demonstrate that, although the proposed hybrid model outperforms conventional LSTM with respect to streamflow forecasting, but the performance is almost same with GRU and is therefore highly recommended as an efficient and reliable approach in hydrological fields.

Keywords: LSTM · MOPEX · Recurrent neural networks · Streamflow forecasting

1 Introduction

Forecasting daily streamflow is one of the tools applied by water authorities to allocate scarce water resources among competitive users, as well as in flood prone regions where early warning can reduce the effects of flood. Floods are one of the most frequently occurring natural disasters which affect many regions of the world resulting in loss of lives and billions worth of properties especially in flood prone regions. Information at any stage of streams/rivers are very important in the analysis, design and construction of water resources projects such as reservoirs flow, dams, channels for flood controls and in streamflow forecasting. In deep learning algorithms, a model class of artificial neural networks which are inspired by biological nervous system called the LSTM are employed to solve the problems of vanishing gradient by controlling information flow using input, forget and output gates. Another default behaviour of an LSTM model is there ability in remembering past information for a long period of time.

X. B. Zhai et al. (Eds.): MLICOM 2019, LNICST 294, pp. 510–524, 2019.
https://doi.org/10.1007/978-3-030-32388-2_44

They also have the capability to process as well as for predicting time series data. Short or long term forecasting of streamflow events helps to optimize and plan for future expansion or reduction. Time series models are models that have its data collected at constant interval of time. These data are analyzed to determine the long-term trends so as to forecast the future or perform some other forms of analysis. Water resources planning are very sensitive in many regions around the globe, and as such it has to be managed in a very sustainable manner. The main purpose of this research is to proposed a hybrid model and compare with the conventional LSTM and GRU models for streamflow forecasting using 35 consecutive years daily streamflow data set from 10 MOPEX basins that have different basin characteristics from different climatic regions in United States.

More recently deep learning LSTM RNN have been applied in streamflow forecasting by [16] for rainfall-runoff modeling, in air pollution [10], compared a long short-term memory neural network extended (LSTME) model to other statistical-based models proving LSTME to be superior. For traffic speed prediction [11], provided an insightful information for transportation professionals to reduce congestion, improve traffic safety, route preplanning as well as rescheduling for the benefit of travellers, although they suggested adding multiple layers in the architecture which might enhance the learning capabilities of the neural networks. LSTM models are also applied for information retrieval and in gesture recognition [19], designed a model which exhibited good performance in sequence level classification, although the model faces the limitation of frame classification while executing gesture recognition. [1] presented an LSTM-based model that can jointly reason across multiple individuals to predict human trajectories in a scene. They suggest future research to extend multi-class settings sharing the same space, thus allowing jointly modeling of human-human and human-space interactions in the same framework.

In this study, whose main objectives is to propose a hybrid model while comparing it's relative performance with an LSTM and GRU models for streamflow forecasting using 10 MOPEX basins data set as a case study. MOPEX's basins contains historical hydrometeorlogical data and river basin characteristics from a range of climates throughout the world for long lead-streamflow forecasting. These models are applied on ten MOPEX's basins; Sandy river, Nezinscot river, Royal river, Saco river, Pemigewasset river, Quinebaug river, Ammonoosuc river, Housatonic river, Tenmile river, and Sacandaga river basins in United States. We also analysed streamflow discharge history for 35 consecutive years in the river basins and applied it to predict future streamflow. The overall performance of the models on the different climatic regimes are evaluated using root mean square error (RMSE) and mean absolute error (MAE). A very important event in water resources management that can affect flood control when designing various hydraulic structures such as dams is accurate streamflow forecast. Although there are many parameters affecting streamflow discharge such as amount of rainfall, the rate of snow pack and glacier melt due to temperature variations etc., the discharge serves only as the input to all the models in the present study.

The rest of the paper is organized as follows; Previous works related to the study of LSTM and GRU models are presented in Sect. 2. We described LSTM, GRU and our proposed hybrid models in details in Sect. 3. Conducting of our experiment including the study areas, methodology and the experimental results are presented in Sect. 4. Finally we conclude the paper and indicate some future work directions in Sect. 5.

2 Related Works

Numerous studies have presented the advantages and applications of recurrent neural network (RNN) including: [16] applies Streamflow Hydrology Estimate using Machine Learning (SHEM) for providing accurate and timely proxy streamflow data set for in-operative streamgages whose result's can be used by first responders and decision makers responding to flood events. [8] applies same data set to compare the relative performance of artificial neural networks and auto-regression models for river flow forecasting. The results proves that neural networks were able to produce better results than auto-regressive models.

Significant attention has been given lately to compare multiple climatic models for streamflow forecasting using hydrological data variables. [9] compares feedforward networks (FFNs) and RNNs models. The results proves RNNs to perform better than feedforward networks for both single step and multi-step ahead forecasting. [3] compares static and dynamic feedback neural network. From the results obtained the dynamic neural network generally produce better and are more stable in streamflow forecasting. [4] employed RNNs and were able to forecast the streamflow where meteorological and hydrological data is rarely available for advanced models.

[1] presented an LSTM-based model that can jointly reason across multiple individuals to predict human trajectories in a scene. They qualitatively prove that social-LSTM successfully predicts various non-linear behaviors arising from social interactions, such as a group of individuals moving together. [21] proposed a model to forecast off-line customer flow for over two thousand shops by considering both online and off-line periodic customer behaviors. The promising experimental results demonstrated that, the proposed approach is superior to the state-of-the-art algorithms such as lasso regression and gradient boosting regression tree. This indicated the wider applicability of the proposed forecast approach.

Another research conducted by [14] proves that, an ANNs model trained with pseudo-inverse rule was capable of performing prediction of combined sewer overflow depth with less than 0.05 error for prediction 5 times ahead for unseen data set. [24] applied Internet of Things (IoT) monitoring combined sewer overflow structures, and compares four different neural networks; multilayer perceptron (MLP), wavelet neural networks (WNN), gated recurrent unit (GRU) and LSTM. LSTM and GRU performed superior performance for multistep ahead prediction with GRU achieving quicker learning curves. [23] designed two different models per city for forecasting weather. The result proves that, LSTM can be considered a better alternative to the traditional methods.

Following the methodology of [11], who proposes a novel LSTM neural networks which is desirable for traffic prediction problem where future traffic condition is relevant to the previous events with long time span. They suggest adding multiple layers which might enhance the learning capability of the neural networks. [2] proposed a multi-step ahead reinforced real-time RNN. The proposed model achieves superior performance while improving the precision of multi-step ahead forecast when compared to two dynamic and one static neural networks. [13] proves that RNN LSTM gives satisfactory improvement with significance of 0.16 correlation and 0.68 in mean squared error over perfect prognosis statistical downscaling techniques.

[7] compares four different artificial intelligence models; ANNs, support vector regression (SVR), wavelet-ANN and wavelet-SVR. From the research conducted, non of the models outperformed the others in more than one watershed, suggesting that some models may be more suitable for certain types of data set. [17] applied timed lagged RNNs and the results out-performed general RNNs in predicting short term flood flow. [22] proposed a method for tweets classification which utilizes weakly-labeled tweets and can significantly improve the accuracy of tweets classification, although they did not include discrimination detection in multiple categories. [15] applied Levenberg Marquardt algorithm to develop an artificial neural network. The results prove's a relatively good agreement between predicted and observed values.

Research conducted by [12], demonstrated the advantages of LSTM model for analyzing the complex non-linear variations of traffic speeds as well as its promising prediction accuracy. [18] proposes a hybrid model of wavelet transform and LSTM. The hybrid model provided better results than the LSTM, Elman and Jordan recurrent neural networks. [10] compares spatio-temporal deep learning model, time delay neural network, auto regressive moving average, support vector regression, and traditional LSTM neural network models. The results demonstrated the long short-term memory neural network extended (LSTME) as superior to the other statistical-based models.

While [5] proposed a field-programmable gate array (FPGA) based accelerator for long short term memory recurrent neural networks, which optimizes both computational performance and communication requirements. The results of design achieves significant speedup over software implementations and it outperforms previous long short term memory recurrent neural network accelerators. [6] designed a model that integrates time delay and RNN. The results obtained proves it to predict better forecast than the statistical autoregressive-moving average with exogenous terms (ARMAX) model. [25] hybrid Ensemble Empirical Mode Composition (EEMD-LSTM) model performed better than the recurrent neural network, long short term memory, EMD-recurrent neural network, EMD-long short term memory and EEMD-recurrent neural network model for daily land surface temperature data series forecasting.

While [20] compares short and long term forecast. Results proves short term forecast can clearly improve real-time operation, however long term forecast still requires improvement in the forecast. From the previous researches conducted, it is clearly proven that promising results have been observed when applying LSTM and GRU models.

This paper presents a hybrid model for streamflow forecasting and compares it's relative performance with LSTM and GRU models using United States Geological Survey (USGS) National Water Information System (NWIS) data set. 35 consecutive years (35 water years) data set from 10 MOPEX river basins having different basins characteristic from different climatic regions were applied as a case study.

3 Modelling

In this study, we applied three artificial intelligence approaches for streamflow forecasting. The first approach is the conventional LSTM model, followed by GRU model and we compares it with the hybrid model (LSTM and GRU). We applied LSTM and GRU models on streamflow data set because they are equipped with the forget and update gates respectively. These gates enables the artificial intelligence models to memorize long term dependencies of specific features from the input data without being erased as well as solving the problems of vanishing gradient. Although in LSTM model, the amount of memory content is controlled by the output gate, the reverse is the case in GRU as they have no full control over it memory content. Thirty five water years historical daily streamflow data set of 10 MOPEX basins in USA obtained from the United States Geological Survey (USGS) National Water Information System (NWIS) are applied as input to these artificial intelligence models.

3.1 LSTM Model

Daily streamflow discharge of each river basins over the period of thirty five consecutive years (35 Water years) were applied on the LSTM model. To validate the effectiveness of the LSTM model. The historical data set were divided into two selected sub-groups, corresponding to 8:2 for training and testing respectively for each of the basins. The first step in the training section involves normalization of the data set between (0 and 1) using Min Max Scaler, which is a simple technique for fitting the data set in a pre-defined boundary. The second step is segmentation of the time series data input using sliding window to determine the prediction accuracy and it was set as 3. Finally we designed the LSTM model consisting of 5 LSTM blocks which is fully connected with epoch, batch size and verbose set as 31, 1 and 2 respectively. RMSE of the train score and test score as well as MAE were applied to evaluate the performance of the model at both training and testing phase respectively.

3.2 Hybrid Model

The proposed hybrid model in this paper is an integration of both LSTM and GRU models. This hybrid model is thought to exploit not only the characteristics and learning capabilities, but also the strength of both GRU and LSTM models, so as to produce a more accurate and reliable forecast on the streamflow data set. The proportion between train and test data sets is 8:2. To validate the effectiveness of the proposed hybrid model, the data set was first divided into two sets; training and testing data set, each of which contains several processes. The first step is the training section, as it involves normalization of the data set between (0 and 1) using Min Max Scaler which is a simple technique for fitting data in a pre-defined boundary. The second step is segmentation of the time series data input, using sliding window to determine the prediction accuracy and it was set as 3. The output of the LSTM model is fed into the GRU model in order to produce a single, final output as it's being concatenated and formed a fully-connected layer. The network is trained and tested with the hybrid model each of which are fully connected with epoch, batch size and verbose set as 31, 1 and 2 respectively. The configuration chosen for the hybrid model in this study is 1-2-1, namely; one input layer, two hidden layers with the first hidden layer having 5 LSTM neurons and the second hidden layer having 5 GRU neurons and an output layer, as this is the multi layer perceptron adopted throughout the study. RMSE of the training and testing data set as well as MAE were applied to evaluate the performance of the model at both training and testing phase respectively. The flowchart of the proposed hybrid model indicating the flow of data from the input to the output state consisting hybrid blocks with a fully connected hidden layer is presented below (Fig. 1).

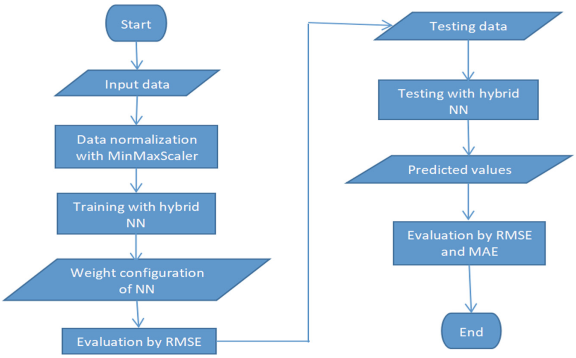

Fig. 1. Flowchart of hybrid model

4 Experiments

This section mainly describes the experimental setup in details including; the study area, the methodology applied in conducting the experiments and the experimental results.

4.1 Study Areas

The hydrological streamflow data set operated by USGS of 10 MOPEX river basins retrieved from the National Water Information System (NWIS, http://waterdata.usgs.gov/nwis, accessed March 2018) for thirty five consecutive years (35 water years) are applied in this study. Since continuous streamflow data are common to all the 10 studied basins and are therefore used for this study in calibrating, evaluating and testing the model. These 10 basins are chosen as study areas because they have minimal or no regulation streamflow. The rationale for choosing these time intervals pertained to the availability of data in each river basin, as each have atleast 50 years of continuous streamgage data records available. The study basins as represented in both geographic and climatic variability and they are; Sandy river, Saco river, Royal river, Quinebaug river, Sacandaga river, Ammonoosuc river, Pemigewasset river, Nezinscot river, Housatonic river and Tenmile river. Streamflow varies depending on the amount of rainfall, geology, rate of snow pack and glacier melt due to temperature variations, seasonal weather conditions, and land cover. Additional physical attributes of the basins were also described; mean areal precipitation (mm), climatic potential evaporation (mm), daily streamflow discharge (mm), daily maximum air temperature (Celsius), daily minimum air temperature (Celsius) by the USGS. The change of stage in a river results from variation of discharge. The gaging stations are established mainly for knowing the flow regime of the river. Details of the case study areas are summarized in the tables below (Tables 1, 2 and 3).

Table 1. Drainage and length of the 10 MOPEX basins.

Station ID	Station name	Drainage (m^2)	Length (mile)
01048000	Sandy river	516	56
01064500	Saco river	385	136
01060000	Royal river	141	39
01127000	Quinebaug river	850	69
01321000	Sacandaga river	491	64
01138000	Ammonoosuc river	850	55
01076500	Pemigewasset river	622	70
01055500	Nezinscot river	169	30
01197000	Housatonic river	1950	149
01200000	Tenmile river	203	8.6

Table 2. Location of the 10 MOPEX basins.

Station ID	Latitude	Longitude
01048000	44.71	−69.94
01064500	43.99	−71.09
01060000	43.79	−70.18
01127000	41.59	−71.99
01321000	43.31	−73.84
01138000	44.22	−71.91
01076500	43.75	−71.69
01055500	44.27	−70.23
01197000	42.23	−73.36
01200000	41.66	−73.53

Table 3. Streamflow discharge history of the 10 MOPEX basins.

Station ID	Max.	Avg.	Min.
01048000	2340(1996)	232	71(1993)
01064500	2470(1996)	297	125(1991)
01060000	11500(1977)	2500	1120(1957)
01127000	3890(1938)	810	73(1946)
01321000	32,000(1913)	12000	16(1913)
01138000	1030(1991)	74	30(2008)
01076500	3270(1996)	352	112(1919)
01055500	874(1996)	62	24.2(1985)
01197000	372(1938)	33	6.90(1962)
01200000	1980(1938)	84	12(1957)

4.2 Methodology

Daily streamflow discharge of each of the 10 MOPEX river basins for the period of thirty five consecutive years (35 Water years) was used in the study. The streamflow data set were obtained from the United States Geological Survey (USGS) National Water Information System (NWIS). It is thought that, the hybrid model to exploit not only the characteristics and learning capabilities but also the strength of both GRU and LSTM models so as to provide a more accurate and reliable forecast on the streamflow data set. The data set is divided into ten subsets of chosen size (35 water year) each. The subsets are first normalized between the range (0 to 1) using Minimum Maximum Scaler, for preserving zero entries in sparse data and including robustness to very small standard deviations of features. For model training and testing, the historical data set were divided into two selected groups, corresponding to 0.8 and 0.2 for training and testing respectively for each of the subset. The training data set were used for training the LSTM, GRU and hybrid models, followed by testing for evaluating the accuracy of the trained networks. The sliding window, which was designed to maximize the retention of the forecast sharpness for forecast systems associated with higher skills was set as 3. After testing the models on different neurons, a model consisting of 5 neurons each of LSTM, GRU and hybrid model respectively with fully connected hidden layer with batch size and verbose set as 1 and 2 respectively was adopted in this study.

In order to estimate the forecasting performance and evaluate the accuracy of the hybrid model. RMSE and MAE are applied on both the LSTM, GRU and hybrid models. The RMSE is basically applied to measure the differences between the predicted and true values of the model, while MAE measures accuracy for continuous variables without considering their directions. Both formulas were applied because they have a good performance to distribution error and could be used to measure the error rate of the models. The mathematical expressions of RMSE and MAE are presented respectively below:

$$RMSE = \sqrt{\frac{\sum (y - x)^2}{n}}$$

$$MAE = \frac{\sum |y - x|}{n}$$

Where y and x are the forecast and observed streamflow data set respectively, n is the total number of samples in each of the data set. The results of the hybrid models are compared with our reference LSTM and GRU models. Tables 4, 5 and 6 presents the statistical test results obtained from the calculated train score RMSE, test score RMSE and MAE of the LSTM, GRU and hybrid models, when the same data set of each basin are applied. The best model is normally the one having least values of MAE and RMSE.

4.3 Experimental Results

In this section, the selective results of all 10 testing phases of the river basins containing the observed and predicted values in the study are presented. The results and discussion are presented in details as shown below.

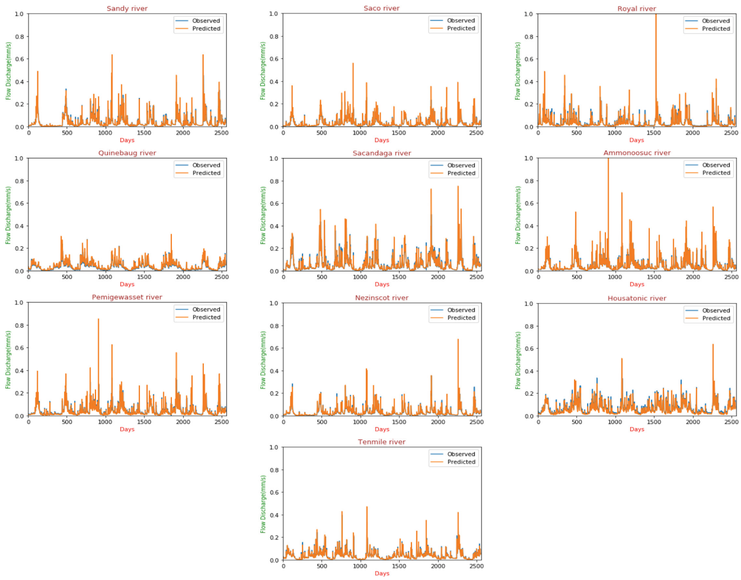

Fig. 2. Testing phase of the river basins using LSTM model.

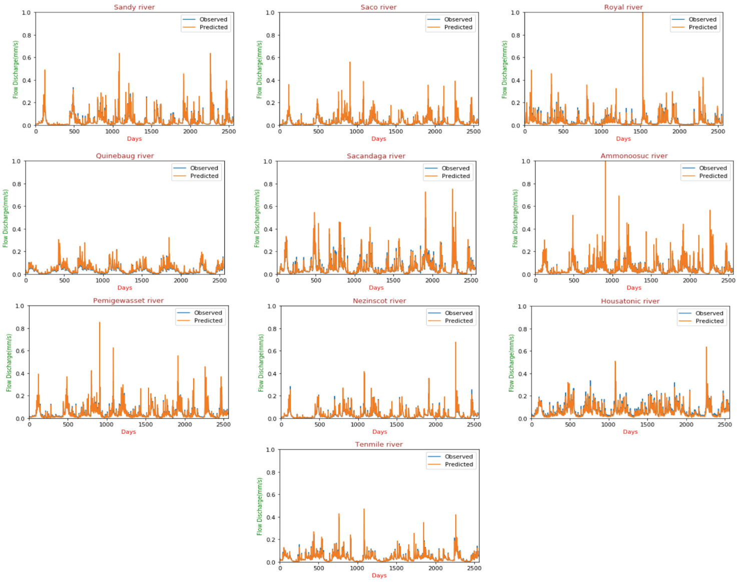

Fig. 3. Testing phase of the river basins using GRU model.

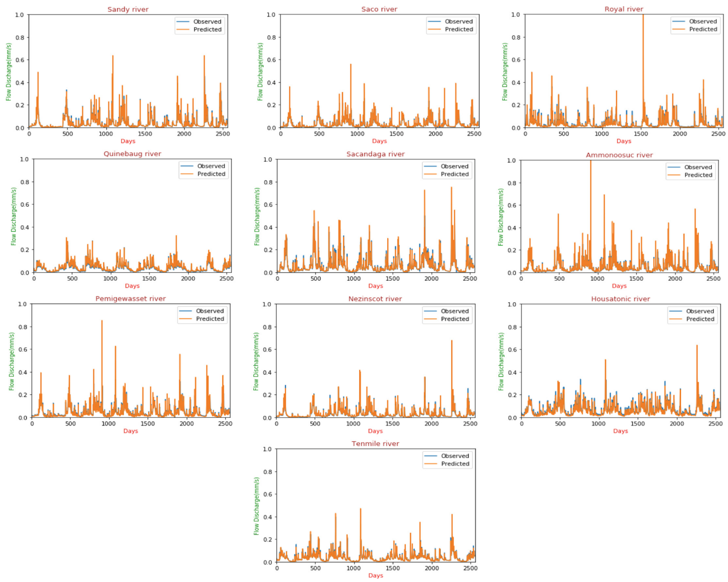

Fig. 4. Testing phase of the river basins using hybrid model.

4.4 Discussions

In this study we applied train score RMSE, test score RMSE and MAE to measure the performances of the models using 35 consecutive years (35 Water years) data set splitted into training and testing for forecasting streamflow discharge on 10 MOPEX basins. Figures 2, 3 and 4 presents the extent match between the observed and predicted values of the testing phase of the LSTM, GRU and hybrid models respectively. While Figs. 5, 6 and 7 shows the histogram distribution of the train score RMSE, test score RMSE and MAE respectively. Tables 4, 5 and 6 summarizes the results of train score RMSE, test score RMSE and MAE of LSTM, GRU and Hybrid (LSTM and GRU) models respectively. These graphs and tables compares the RMSE and MAE of the 3 models. The best performers on the tables are highlighted in bold font, while the second best performers are underlined.

The summarized forecasting performance of the LSTM, GRU and hybrid models in terms of train score RMSE and test score RMSE of all 10 MOPEX basins and the histogram distribution of train score RMSE and test score RMSE are presented in Tables 4, 5 and Figs. 5, 6 respectively. As seen from the tables, GRU and hybrid models performed better than LSTM model in streamflow forecasting because they have the least values of RMSE. Although the GRU outperformed the hybrid model in 6 of the 10 basins, but the performance of the GRU and hybrid model seems to be very close.

Table 6 presents the results of the calculated MAE, while Fig. 7 shows the plot of the histogram distribution of MAE of the 10 MOPEX basins when same data set are applied on LSTM, GRU and hybrid models. As seen from the table, the GRU model exhibited it's best performances in 6 MOPEX basins, while hybrid model exhibited it's best performances in 4 MOPEX basins with both models slightly different to each other, but never the less they outperformed the LSTM model in streamflow forecasting.

It is obvious from the Tables 4, 5 and 6 presented that, the GRU are most efficient for streamflow forecasting on 01048000, 01064500, 01060000, 01127000, 01321000, 01138000 basins with the hybrid model having best performance in 01076500, 01055500, 01197000, 01200000 basins. The performance of the GRU and hybrid models does not differ much, although the hybrid model have both the characteristics and properties of both the LSTM and GRU models.

Table 4. Train score RMSE of LSTM, GRU and hybrid models of the 10 MOPEX basins,

Rivers	LSTM RMSE	GRU RMSE	Hybrid RMSE
Sandy	4.9782	**1.5376**	<u>1.5393</u>
Saco	7.4384	**2.2672**	<u>2.3089</u>
Royal	5.4146	**1.5047**	<u>1.5122</u>
Quinebaug	3.7090	**0.6765**	<u>0.7797</u>
Sacandaga	4.1191	**1.2369**	<u>1.3485</u>
Ammonoosuc	3.8715	**1.2414**	<u>1.2488</u>
Pemigewasset	5.4605	<u>1.8866</u>	**1.8151**
Nezinscot	4.5402	<u>1.0741</u>	**1.0463**
Housatonic	3.2846	<u>0.8739</u>	**0.7079**
Tenmile	3.7099	<u>1.8853</u>	**0.9426**

Table 5. Test score RMSE of LSTM, GRU and hybrid models of the 10 MOPEX basins.

Rivers	LSTM RMSE	GRU RMSE	Hybrid RMSE
Sandy	5.8595	**1.8161**	1.8209
Saco	7.9457	**2.2349**	2.2530
Royal	6.1550	**2.0571**	2.0745
Quinebaug	3.8878	**0.6481**	0.7333
Sacandaga	4.8287	**1.4959**	1.5465
Ammonoosuc	4.2415	**1.5033**	1.5351
Pemigewasset	6.0844	2.1386	**2.0704**
Nezinscot	5.1345	1.2112	**1.2064**
Housatonic	3.5813	0.9273	**0.7877**
Tenmile	4.2478	1.8923	**0.9839**

Table 6. MAE of LSTM, GRU and hybrid models of the 10 MOPEX basins.

Rivers	LSTM MAE	GRU MAE	Hybrid MAE
Sandy	0.6454	**0.1193**	0.1196
Saco	0.8904	**0.2150**	0.2159
Royal	0.6494	**0.1647**	0.1657
Quinebaug	0.3367	**0.0325**	0.0349
Sacandaga	0.5647	**0.0943**	0.0977
Ammonoosuc	0.5685	**0.1400**	0.1425
Pemigewasset	0.7577	0.1954	**0.1906**
Nezinscot	0.4595	0.0556	**0.0554**
Housatonic	0.3975	0.0911	**0.0853**
Tenmile	0.3837	0.1014	**0.0510**

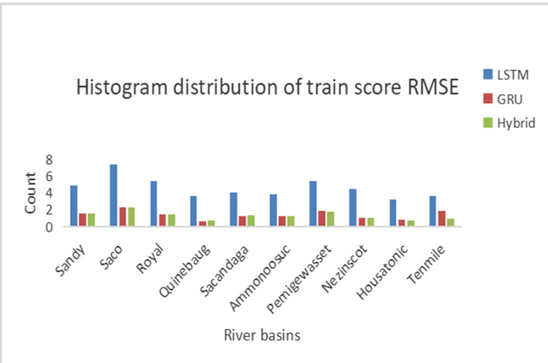

Fig. 5. Histogram on distribution of train score RMSE.

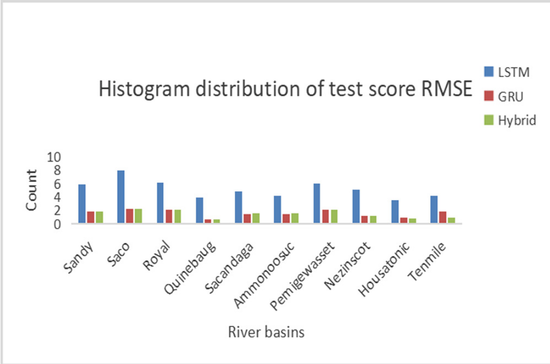

Fig. 6. Histogram on distribution of test score RMSE.

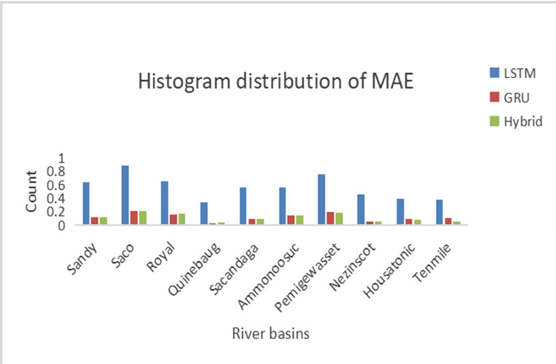

Fig. 7. Histogram on distribution of MAE.

5 Conclusion

In this research we proposed a hybrid model which is an integration of LSTM and GRU models while applying it for streamflow forecasting. The hybrid model's performance is compared with our reference LSTM and GRU models. These models are trained and tested with 35 consecutive years streamflow discharge data set of 10 MOPEX basins having different basin characteristics from different climatic regions in United States. MOPEX basins were selected because they have atleast 50 years continuous streamflow data set available that have minimal or no regulation. The performance of each model in terms of train score RMSE, test score RMSE and MAE are evaluated. The proposed hybrid and GRU models outperformed the LSTM model for streamflow forecasting. Although the performance of the hybrid model is almost the same with GRU in streamflow forecasting, but the GRU model outperformed the hybrid model slightly in 6 of the 10 MOPEX basins suggesting that some models may be more suitable for certain types of data sets.

Acknowledgement. This work was supported in part by the National Key R&D Program of China under Grant 2018YFC0407901, in part by the National Natural Science Foundation of China under Grant 61602149, and in part by the Fundamental Research Funds for the Central Universities under Grant 2019B15514.

References

1. Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., Savarese, S.: Social LSTM: human trajectory prediction in crowded spaces. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 961–971 (2016)
2. Chen, P.A., Chang, L.C., Chang, F.J.: Reinforced recurrent neural networks for multi-step-ahead flood forecasts. J. Hydrol. **497**, 71–79 (2013)
3. Chiang, Y.M., Chang, L.C., Chang, F.J.: Comparison of static-feedforward and dynamic-feedback neural networks for rainfall runoff modeling. J. Hydrol. **290**(3–4), 297–311 (2004)
4. Daniyal, H., et al.: Application of neural network for river flow forecasting. Int. J. Sci. Eng. **8**, 1365–1371 (2004)
5. Guan, Y., Yuan, Z., Sun, G., Cong, J.: FPGA-based accelerator for long short-term memory recurrent neural networks. In: 2017 22nd Asia and South Pacific Design Automation Conference (ASP-DAC), pp. 629–634. IEEE (2017)
6. Karamouz, M., Razavi, S., Araghinejad, S.: Long-lead seasonal rainfall forecasting using time-delay recurrent neural networks: a case study. Hydrol. Process. Int. J. **22**(2), 229–241 (2008)
7. Karran, D.J., Morin, E., Adamowski, J.: Multi-step streamflow forecasting using data-driven non-linear methods in contrasting climate regimes. J. Hydroinformatics **16**(3), 671–689 (2014)
8. Kisi, O.: Daily river flow forecasting using artificial neural networks and auto-regressive models. Turk. J. Eng. Environ. Sci. **29**(1), 9–20 (2005)
9. Kumar, D.N., Raju, K.S., Sathish, T.: River flow forecasting using recurrent neural networks. Water Resour. Manag. **18**(2), 143–161 (2004)
10. Li, X., et al.: Long short-term memory neural network for air pollutant concentration predictions: method development and evaluation. Environ. Pollut. **231**, 997–1004 (2017)
11. Ma, X., Tao, Z., Wang, Y., Yu, H., Wang, Y.: Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transp. Res. Part C Emerg. Technol. **54**, 187–197 (2015)
12. Chen, M., Yu, G., Chen, P., Wang, Y.: Traffic congestion prediction based on long-short term memory neural network models. In: CICTP 2017: Transportation Reform and Change—Equity, Inclusiveness, Sharing, and Innovation, pp. 673–681 (2018)
13. Misra, S., Sarkar, S., Mitra, P.: Statistical downscaling of precipitation using long short-term memory recurrent neural networks. Theor. Appl. Climatol. **134**(3–4), 1179–1196 (2018)
14. Mounce, S.R., Shepherd, W., Sailor, G., Shucksmith, J., Saul, A.J.: Predicting combined sewer overflows chamber depth using artificial neural networks with rainfall radar data. Water Sci. Technol. **69**(6), 1326–1333 (2014)
15. Othman, F., Naseri, M.: Reservoir inflow forecasting using artificial neural network. Int. J. Phys. Sci. **6**(3), 434–440 (2011)
16. Petty, T.R., Dhingra, P.: Streamflow hydrology estimate using machine learning (SHEM). JAWRA J. Am. Water Resour. Assoc. **54**(1), 55–68 (2018)
17. Deshmukh, R.P., Ghatol, A.A.: Short term flood forecasting using general recurrent neural network modeling a comparative study. Int. J. Comput. Appl. **8**(12), 5–9 (2010)

18. Sugiartawan, P., Pulungan, R., Sari, A.K.: Prediction by a hybrid of wavelet transform and long-short-term-memory neural network. Int. J. Adv. Comput. Sci. Appl. **8**(2), 326–332 (2017)
19. Tsironi, E., Barros, P., Wermter, S.: Gesture recognition with a convolutional long short-term memory recurrent neural network. In: Proceedings of the European Symposium on Artificial Neural Networks Computational Intelligence and Machine Learning (ESANN), pp. 213–218 (2016)
20. Tucci, C.E., Collischonn, W., Clarke, R.T., Paz, A.R., Allasia, D.: Short-and long-term flow forecasting in the Rio Grande watershed (Brazil). Atmos. Sci. Lett. **9**(2), 53–56 (2008)
21. Yin, Z., Zhu, J., Zhang, X.: Forecast customer flow using long short-term memory networks. In: 2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), pp. 61–66. IEEE (2017)
22. Yuan, S., Wu, X., Xiang, Y.: Incorporating pre-training in long short-term memory networks for tweet classification. Social network analysis and mining. In: Proceedings of the IEEE 16th International Conference on Data Mining, pp. 1329–1334 (2018). https://doi.org/10.1109/icdm.2016.91
23. Zaytar, M.A., El Amrani, C.: Sequence to sequence weather forecasting with long short-term memory recurrent neural networks. Int. J. Comput. Appl. **143**(11), 7–11 (2016)
24. Zhang, D., Lindholm, G., Ratnaweera, H.: Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring. J. Hydrol. **556**, 409–418 (2018)
25. Zhang, X., Zhang, Q., Zhang, G., Nie, Z., Gui, Z., Que, H.: A novel hybrid data-driven model for daily land surface temperature forecasting using long short-term memory neural network based on ensemble empirical mode decomposition. Int. J. Environ. Res. Public Health **15**(5), 1032 (2018)