Evaluation of short-term streamflow prediction methods in Urban river basins

Huang, Xinxing; Li, Yifan; Tian, Zhan; Ye, Qinghua; Ke, Qian; Fan, Dongli; Mao, Ganquan; Chen, Aifang; Liu, Junguo

DOI
10.1016/j.pce.2021.103027

Publication date
2021

Document Version
Accepted author manuscript

Published in
Physics and Chemistry of the Earth

Evaluation of short-term streamflow prediction methods in urban river basins

Xinxing Huang[a, b, #], *Yifan Li*[b, #], *Zhan Tian*[b, *], *Qinghua Ye*[c], *Qian Ke*[d], *Dongli Fan*[a], *Ganquan Mao*[b], *Aifang Chen*[b], *Junguo Liu*[b, *]

[a] *School of Chemical and Environmental Engineering, Shanghai Institute of Technology, Shanghai 201418, China*

[b] *School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China*

[c] *Deltares, Delft, the Netherlands*

[d] *Department of Hydraulic Engineering, Faculty of Civil Engineering and Geoscience, Delft University of Technology, Delft, the Netherlands*

[#] *Xinxing Huang and Yifan Li contributed equally to this work*

[*] *Corresponding Author: Zhan Tian: tianz@sustech.edu.cn,*

Junguo Liu: liujg@sustech.edu.cn

Abstract

Efficient and accurate streamflow predictions are important for urban water management. Data-driven models, especially neural network (NN) models can predict streamflow fast, while the results are uncertain in some complex river systems. Physically based models can reveal the underlying physics, but it is relatively slow and computationally costly. This work focuses on evaluating the reliability of three NN models (artificial neural networks (ANN), long short-term memory networks (LSTM), adaptive neuro-fuzzy inference system (ANFIS)) and one physically based model (SOBEK) in terms of efficiency and accuracy for average and peak streamflow simulation. All the models are applied for a tidal river and a mountainous river in Shenzhen. The results show that, the ANN model calculates fastest since the hidden layer's structure is simple. The LSTM model is reliable in average streamflow simulation in tidal river with the lowest bias while the ANFIS model has the best accuracy for peak streamflow simulation. Furthermore, the SOBEK model shows reliability in simulating average and peak streamflow in mountainous river due to its ability to capture uneven spatial rainfall in the area. Overall, the results indicate that the LSTM model can be a helpful supplementary to physically based models in streamflow simulation of complex urban river systems, by giving fast streamflow predictions with usually acceptable accuracy. Our results can provide helpful information for hydrological engineers in the application of flooding early warning and emergency preparedness in the context of flooding risk management.

Keywords: streamflow simulation, neural network models, SOBEK model, urban rivers

1. Introduction

Urban flooding has become a threat to urban water security and has increased in frequency in recent decades (Ziegler et al., 2012). China is prone to urban flooding, and many of Chinese cities have been experiencing a large increases in urban flooding in recent years (Duan et al., 2016). Between 2008 and 2010, 218 Chinese cities endured at least one urban flooding event while more than 100 cities experienced three urban flooding events, including major cities like Beijing, Shanghai, Guangzhou and Shenzhen (Jiang et al., 2018; Song and Li, 2019). Tidal river basins and mountainous river basins are two typical urban river basins with large impervious areas that are vulnerable to urban flooding (Archetti et al., 2011; Davenport et al., 2004; Dawson et al., 2008). In most tidal river basins of China, urban flooding is mainly caused by large amounts of rainfall-runoff at the same time of sustained high tide at outlet, in which the rivers cannot convey large water volume during high precipitation period (Lian et al., 2013; Orton et al., 2020). Urban flooding also occurs in mountainous river basins due to its uneven slope distribution (Ballesteros-Cánovas et al., 2015). It causes the streamflow to quickly move from high-altitude areas to low-altitude areas, and impervious areas lead to slow streamflow infiltration, so a large amount of streamflow will accumulate in a short time (Chen et al., 2008). Urban flooding can result in disasters that cause enormous public and private property losses and casualties. In March 2014, Shenzhen experienced a 50-year rainfall event, paralyzing the urban sewer system and surface water flows with more than 200 inundation areas (Xu et al., 2020). In 2016, weeks of torrential rainfall during the monsoon season led to severe urban flooding, which submerged 28 provinces and impacted 60 million people in China (Jiang et al., 2018). Therefore, streamflow prediction in tidal river basins and mountainous river basins is necessary and crucial for public safety management and social development.

Urban flooding is characterized by short duration and high intensity, making it difficult to predict. The rapid forecasting and prediction of urban flooding can minimize potential losses. Therefore, efficient and accurate simulations of the streamflow caused by urban flooding is of great concern for urban water resource management and decision makers. Hydrologists have paid increasing attention and efforts to developing urban flooding models. Physically based models have been widely used for urban flooding prediction in recent decades (Anghileri et al., 2016; Botto et al., 2018; Chen et al., 2016). They take the dynamics of the hydrological cycle process into account, build a hydrodynamic equation set based on the characteristics of runoff generation in the basin, and simulate the rainfall runoff response. They can fully simulate the whole rainfall-runoff process of a river basin (Kim and Mohanty, 2017). Physically based models, however, also present several challenges. For instance, these models need large basic data for model set-up and calibration (Li et al., 2020), and the quality of the simulation results depends on the quality and availability of the input data (Yoon et al., 2011). Additionally, the calculation of physically based models is relatively slow, and the computational cost is expensive due to considerable data calibration and validation.

Due to the slow simulation process of physically based models (Yang et al., 2020), data-driven models have gained considerable attention in hydrology in recent years due to their rapid simulation capacity (Ahani et al., 2018). Data-driven models are therefore seen as alternatives to physically based models. They consider the input and output data, without using any of the physical processes (Wang and Yao, 2013). Among the various data-driven models, neural network (NN) models are the most widely used techniques for streamflow simulation and forecasting (Humphrey et al., 2016; Yang et al., 2020; Zhang et al., 2020). Artificial neural network (ANN) model, can be seen as a black-box model, has been used in river streamflow simulation because of its ability to mimic both linear, nonlinear and hydrological systems (Aichouri et al., 2015; Kashani et al., 2016; Shoaib et al., 2014). Among the data-driven models, it has a long development history, with the first studies using the ANN model for streamflow prediction dating back to the early 1990s (Daniell, 1991; Halff et al., 1993). However, a drawback of ANN's hidden layer is that any information about the sequential order of the inputs is lost. Long short-term memory (LSTM) model overcomes the problem of ANN's hidden layer through a specially designed architecture (Kratzert et al., 2018). This architecture has memory cells replacing the traditional hidden layer. The memory cells could store, write and read data via gates that open and close (Zhang et al., 2018). This can overcome the problem of the ANN model of learning long-term dependencies representing, for example, storage effects within hydrological catchments, which may play a significant role for hydrological process (Kratzert et al., 2018). With the rapid development of data-driven modelling approaches, there has been a shift from black-box models to semantic-based fuzzy systems in recent years (Ang and Quek, 2005). Adaptive network-based fuzzy inference system (ANFIS) model is one example of a semantic-based fuzzy system which conducts learning through the minimization of global error within the model. However, there are several essential limitations for NN models, such as the lack of explanation of the physical mechanism and transparency of the simulation process and difficulties in explaining the results (Elshorbagy et al., 2010).

Through the research, we find that physically based models can make up for the disadvantages of data-driven models that cannot simulate the physical process, and data-driven models can make up for the slow simulation and prediction of physically based models. However, due to the uncertainty of physical parameters, input parameters and basic data, the simulation accuracy of the physically based model is also uncertain in some cases (Hattermann et al., 2018; Her et al., 2019; Liu et al., 2017; Sikorska and Renard, 2017). We therefore propose a hypothesis: for the streamflow simulation of a complex urban river system (i.e., a tidal river basin and a mountainous river basin), physically based models cannot fully generalize the physical process and the speed of the model is limited. Can data-driven models work as the supplementary, help to better simulate the streamflow in complex urban river systems? Existing studies that compared neural network models and physically based models for streamflow simulation in urban areas have only concentrated on data at the daily, monthly or annual scale (Chang and Chen, 2018; Mernild et al., 2018; Nikpour et al., 2019; Schuol et al., 2008; Tikhamarine et al., 2020). There are very few studies on the streamflow simulation based on hourly rainfall data. As the reliable streamflow simulation and prediction plays a key role in confronting urban flooding risks, a high temporal resolution precipitation dataset could have considerable influence on model accuracy (Bruneau et al., 1995) and help deepen our understanding of the process of streamflow, especially the process of extreme flooding events. The novelty of this study is that it conducts an evaluation on the reliability of short-term streamflow prediction methods driven by hourly rainfall data, with the goal to provide more suitable streamflow simulation models for urban rivers.

This paper is organized as follows. Section 2 describes the study area, the data and method used in this study. Section 3 shows the calibration and validation results of the models, and the comparison of model performance in average and peak streamflow simulation. Section 4 and Section 5 present the discussion and conclusion, respectively.

2. Materials

2.1 Study area

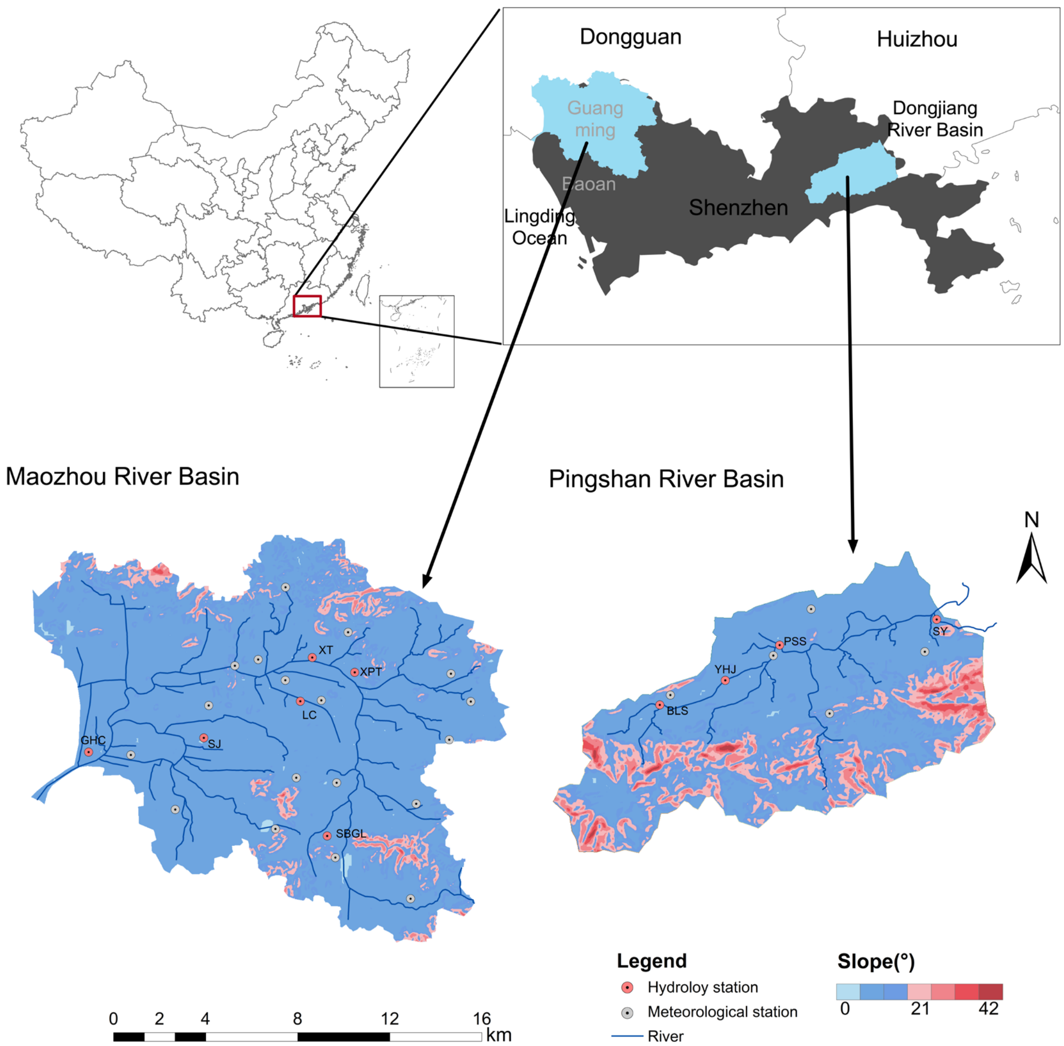

Figure 1 Geography map and slope map (coordinate system: WGS-1984) of Maozhou River Basin and Pingshan River Basin in Shenzhen city, China; the spatial distribution of hydrological stations (red dots) and meteorological stations (white dots).

Shenzhen is a coastal city in China that is experiencing rapid economic development. Urban flooding is one of the most devastating natural disasters that can occur in Shenzhen (Ke et al., 2020). Due to its rapid urbanization, the frequency of urban flooding in Shenzhen has increased in recent years (Shi et al., 2007; Yan et al., 2019). Maozhou River (MZR) is a tidal river (Cui and Guo, 2006), located in the northwest of Shenzhen close to the borders of Dongguan, which is the largest watershed in Shenzhen (see Figure 1). It flows through Baoan District and Guangming District and finally flows into the Lingding Ocean. Pingshan River (PSR), a mountainous river with an average slope of 2.76% (Xiong et al., 2010), is located in north-eastern Shenzhen and close to the borders of Huizhou. The distribution of slopes in the two river basins is different (see Figure 1), resulting in different lag time period for the start of the rain to the peak of the flooding. Maozhou River can reach its peak in one hour, while Pingshan River can reach its peak in 40 minutes (SZN, 2020). The comparison of general characteristics (including the river length, basin area, land use types, annual average rainfall, average slope and average elevation) of two rivers are shown in Table 1 (Chen et al., 2016; Cui and Guo, 2006; Peng et al., 2018; SMEEB, 2018).

Table 1 General characteristics of two river basins (MZR and PSR).

River	Length (km)	Area (km^2)	Urban land	Forest land	Green land	Annual average rainfall(mm)	Average slope	Average elevation (m)
MZR	41.61	388.23	48%	32%	20%	1800	2.2‰	<25
PSR	22.14	129.40	40%	50%	10%	2073	2.76%	<82

2.2 Data

In this study, digital elevation model (DEM) data, meteorological data, hourly streamflow observation data and river profiles were used. The DEM data is retrieved from the Shuttle Radar Topography Mission with a resolution of 30m (SRTM, 2020). Meteorological data, including hourly precipitation data, daily temperature (average, maximum and minimum), and daily average wind speed, are provided by the Meteorological Bureau of Shenzhen (SMB). Hourly streamflow observation data and the river profiles, including bed level, channel slope, width of cross section and shape of cross section, are provided by the Huadong Engineering Corporation Limited (ECIDI) and the Municipal Ecological Environment Bureau of Shenzhen (SMEEB).

2.3 Method

As shown in Figure 2, we apply three neural network models (artificial neural networks (ANN), long short-term memory networks (LSTM) and adaptive neuro-fuzzy inference system (ANFIS)) with hourly rainfall data for streamflow simulation in a tidal river basin (Maozhou River) and a mountainous river basin (Pingshan River) in Shenzhen, China, and use a physically based model (the SOBEK model) as a reference. The introduction of the four models can be found in the supplementary materials. We attempt to assess model performance in three parts: model accuracy in average and peak streamflow simulation, model accuracy. Meanwhile, we examine whether neural network models can make up for physically based models in complex urban river systems in terms of accuracy.

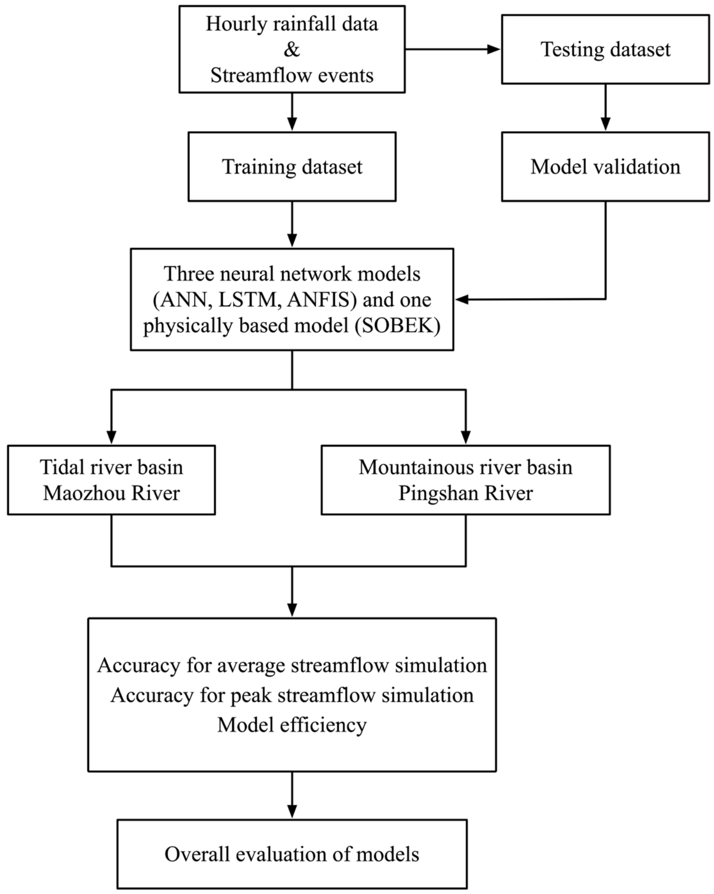

Figure 2 The general framework of this study.

2.4 The definition of streamflow event

Adams et al. (1986) found that less than 60 minutes intervals between two streamflow events, causes the division of rainfall to have greater impacts on the rainfall characteristic parameters, which will not be conducive to the statistics of rainfall characteristics. If the interval is set between 1-6 hours, the division of rainfall has a lower impact on rainfall characteristic parameters and is more reasonable and scientific. Figure 3 shows the definition of streamflow events in this paper. Considering the effects of rainfall confluence time and rainfall duration, this study selects 180 minutes as the minimum interval between two streamflow events, and the cumulative rainfall of each event is greater than 3 mm.

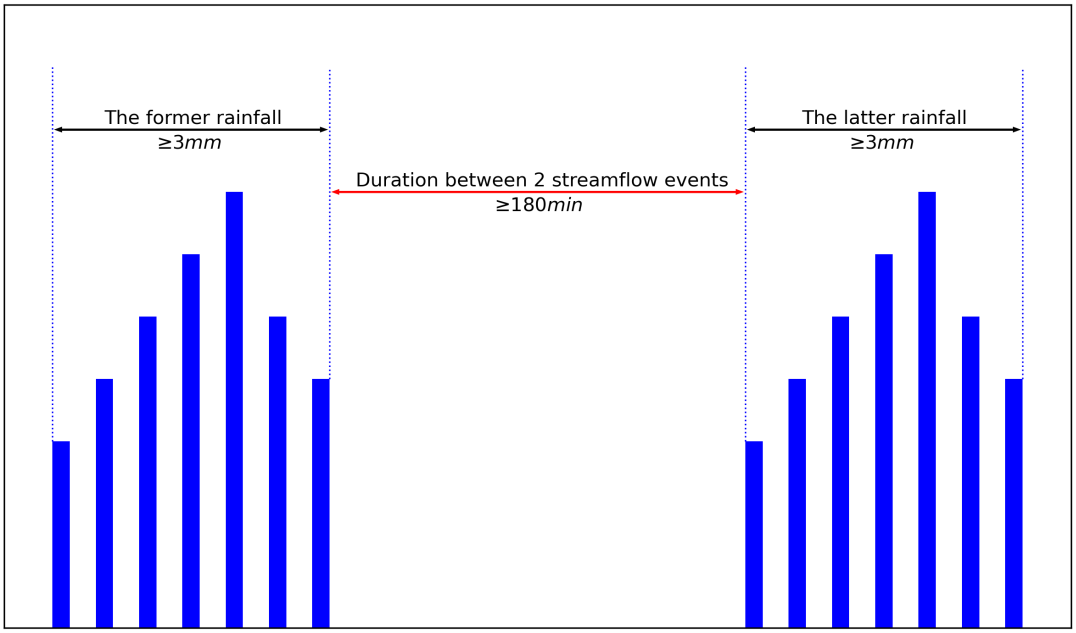

Figure 3 The definition of streamflow events in this research.

2.5 Assessment criteria

The Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970), the R square (R^2), the percent bias (PBIAS) and the root-mean-square error (RMSE) are selected as assessment criteria to evaluate the model results in calibration and validation periods. Meanwhile, the Taylor diagram is applied to visualize the model performance in average streamflow simulation (Taylor, 2001). In addition to the standard deviation and correlation coefficient (r), the center root-mean-square errors (CRMSEs) are used in the Taylor diagram. The equations for computing these objective functions are given as follows:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(X_{i,Obs}-X_{i,Sim})^2}{\sum_{i=1}^{n}(X_{i,Obs}-\bar{X}_{Obs})^2} \quad (1)$$

$$R^2 = \left(\frac{\sum_{i=1}^{n}(X_{i,Obs}-\bar{X}_{Obs})*(X_{i,Sim}-\bar{X}_{Sim})}{\sqrt{\sum_{i=1}^{n}(X_{i,Obs}-\bar{X}_{Obs})^2}*\sqrt{\sum_{i=1}^{n}(X_{i,Sim}-\bar{X}_{Sim})^2}}\right)^2 \quad (2)$$

$$PBIAS = \frac{\sum_{i=1}^{n}(X_{i,Sim} - X_{i,Obs})}{\sum_{i=1}^{n}(X_{i,Obs})} * 100 \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{i,Obs} - X_{i,Sim})^2}{n}} \quad (4)$$

$$CRMSE = \sqrt{\frac{\sum_{i=1}^{n}\left((X_{i,Obs} - \bar{X}_{Obs}) - (X_{i,Sim} - \bar{X}_{Sim})\right)^2}{n}} \quad (5)$$

where $X_{i,Obs}$ and $X_{i,Sim}$ are the i-th observation and simulation data respectively, $\bar{X}_{Obs}$ and $\bar{X}_{Sim}$ are the mean values of the observation and simulation data respectively, and n is the sample size.

The NSE can measure the goodness of fit, from 0 to 1, where a value approaching 1 means the simulations are closer to the observations. The R^2 and r value are used to express the correlation of simulation data and observation data directly, and a value approaching 1 indicates a perfect correlation. The PBIAS measures the average tendency of the simulated values to be larger or smaller than their observed ones. The optimal value of PBIAS is 0.0, positive values indicate overestimation bias, whereas negative values indicate model underestimation bias (Moriasi et al., 2007). The RMSE is used to evaluate how closely the simulation data match the observation data and the values can range from 0 to $+\infty$ based on the range of the data. The CRMSE is similar to the RMSE but has easier visualization characteristics in Taylor diagram. It ranges from 0 to 1, where a value approaching 0 means good simulation result.

3. Results

3.1 Model accuracy in average streamflow simulation

We first selected 59 streamflow events (49 events for calibration; 10 events for validation) for Maozhou River and 21 events (11 events for calibration; 10 events for validation) for Pingshan River. As seen from Table 2, four models can simulate the streamflow of two river basins satisfactorily with high NSE (larger than 0.80) and low absolute PBIAS value (less than 15%) and the ANN model is the best one among four models in calibration period in Maozhou River, the SOBEK model is the best in Pingshan River. Moreover, it can be seen that in the validation period, the LSTM model is the best in Maozhou River, and the SOBEK model is the best in Pingshan River. The details of the calibration and validation results can be found in appendices (Figure A1).

Table 2 Comparison of assessment criteria of four models for streamflow simulation in calibration and validation periods.

Basins	Models	NSE	R^2	RMSE (m^3/s)	PBIAS (%)
MZR	Calibration				
	SOBEK	0.849	0.868	10.137	-12.029
	ANN	0.954	0.956	5.601	-0.850
	LSTM	0.922	0.939	7.284	-14.317
	ANFIS	0.852	0.854	10.029	4.071
	Validation				
	SOBEK	0.948	0.950	11.387	-0.542
	ANN	0.587	0.687	32.106	-17.860
	LSTM	0.976	0.977	7.725	-1.379
	ANFIS	0.883	0.884	17.123	0.121
PSR	Calibration				
	SOBEK	0.925	0.926	1.922	-6.760
	ANN	0.874	0.879	2.498	-0.702
	LSTM	0.888	0.892	2.358	6.461
	ANFIS	0.817	0.820	3.010	3.989
	Validation				
	SOBEK	0.891	0.897	2.109	1.670

ANN	0.645	0.646	3.805	6.480
LSTM	0.830	0.848	2.633	9.791
ANFIS	0.878	0.880	2.235	4.982

Visualizing the average streamflow simulation statistic results of standard deviation, CRMSE, and the correlation coefficient in the Taylor diagram (Figure 4) verifies the distinguish performance of each model. Generally, the LSTM model in Maozhou River (yellow dot) and the SOBEK model in Pingshan River (blue star) lead the reliability of average streamflow simulation regarding the statistical performance, characterized by relatively small standard deviation and CRMSE, and relatively higher r values. The details of the comparison between the observed and simulated results for the four models in terms of average streamflow in two river basins can be found in appendices (Figure A2 and A3).

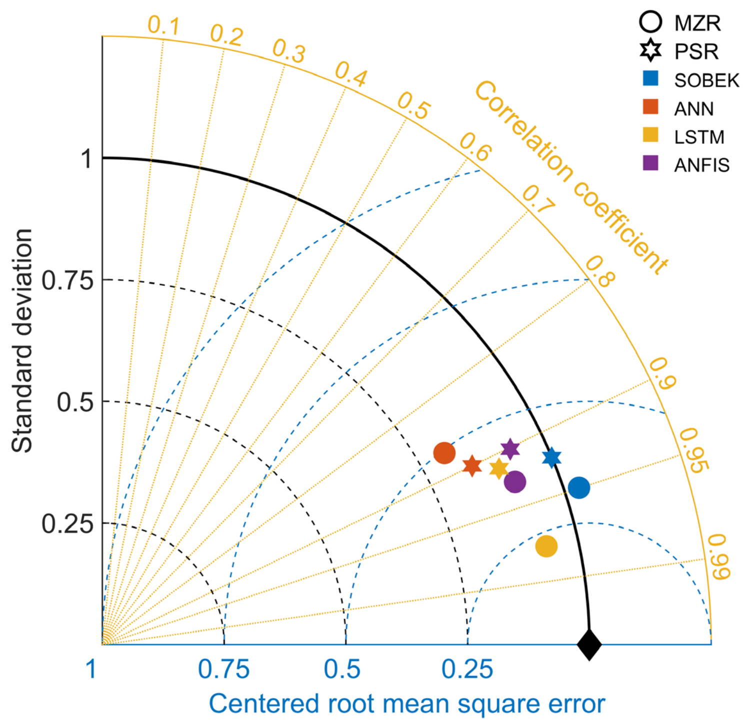

Figure 4 Taylor diagram of the four models in two basins, with shapes and colours indicating simulation and observation data.

3.2 Evaluation of model efficiency

In Section 3.1, we evaluated the model accuracy in average streamflow simulation during calibration and validation period. In this section, the model efficiency will be evaluated as the computation time of models would be of great importance for the streamflow simulation during short period.

The computation time of the models is related to the computer's CPU and memory. To eliminate the impact of different computer configurations, the computer configuration we adopted is as follows: i7-8700 CPU, 32G memory. Under the same computer configuration, the computation time of the SOBEK model depends on the complexity of the constructed river network and the amount of input rainfall data. As the river network of Maozhou River is more complicated than Pingshan River, the SOBEK model validates for almost 1 hour in Maozhou River and 30 minutes in Pingshan River. Unlike the physically based models, rainfall data is the only input required in the neural network model (after calibration) to obtain the prediction streamflow. Therefore, the computation time is much faster.

We also measured the validation time of the three neural network models. The time series of the two basins for validation is approximately 350 hours. The ANN model is the quickest one (the simulation time is 6 min for the MZR and 5 min for the PSR), and the LSTM model is the slowest among three neural network models we used (12 min for the MZR and 10 min for the PSR). The ANN model and LSTM model provided faster simulation for Pingshan River than Maozhou River as there were fewer streamflow events in Pingshan River. Fewer streamflow events mean less simulation time. The simulation time of the ANFIS model was between the ANN model and the LSTM model, and there was no difference in the simulation time of the ANFIS model in the two basins (approximately 8 min). Overall, it shows that the validation time of the neural network models is faster than that of physically based model.

For the prediction time, we performed a rough test to estimate simulation time based on the existing validation time. In Shenzhen, the streamflow from rainfall to peak value does not exceed three hours (SZN, 2020). Therefore, we used a three-hour rainfall data for testing. As seen in Table 3, under a future three-hour rainfall event, the SOBEK model has a better performance in prediction time for each basin than in validation because the amount of input rainfall data is smaller. The neural network models can increase a factor of 10-60 over the SOBEK model for 3-hours rainfall duration (the ANN model takes approximately 30 seconds, the LSTM model approximately 2 minutes and the ANFIS model approximately 1 minute).

Table 3 The prediction time of 3-hour rainfall data using four models.

Model / River	SOBEK	ANN	LSTM	ANFIS
MZR	30 min	≈30s	≈2 min	≈1 min
PSR	15min	≈30s	≈1.5 min	≈1 min

3.3 Model performance in predicting flooding events

To analyze the simulation performance of the four models during a flooding event process, we compare the observation and simulation hydrographs of four flooding events (the rainfall characteristics of these events are short duration and high intensity) using the four models in the two river basins ((a) is SJ station, (b) is LC station and (c) is XT station in MZR, (d) is PSS station in PSR see figure 1) in Figure 5. The ANFIS model is the best among the four models for the simulation of peak streamflow during a flooding event, while the ANN model is the worst, which relies in the simulated values of the ANN model fluctuate abnormally compared with the observed values in large volume streamflow (especially at LC station see Figure 5(b)). The LSTM model is not very effective in simulating small volume flow values in flooding events (the LSTM model cannot reflect the fluctuating state of small volume streamflow, see Figure 5 (d), highlighted by a grey box). The SOBEK model shows abnormal fluctuate values between the time of 40-50 hours in Figure 5(b) (see green line).

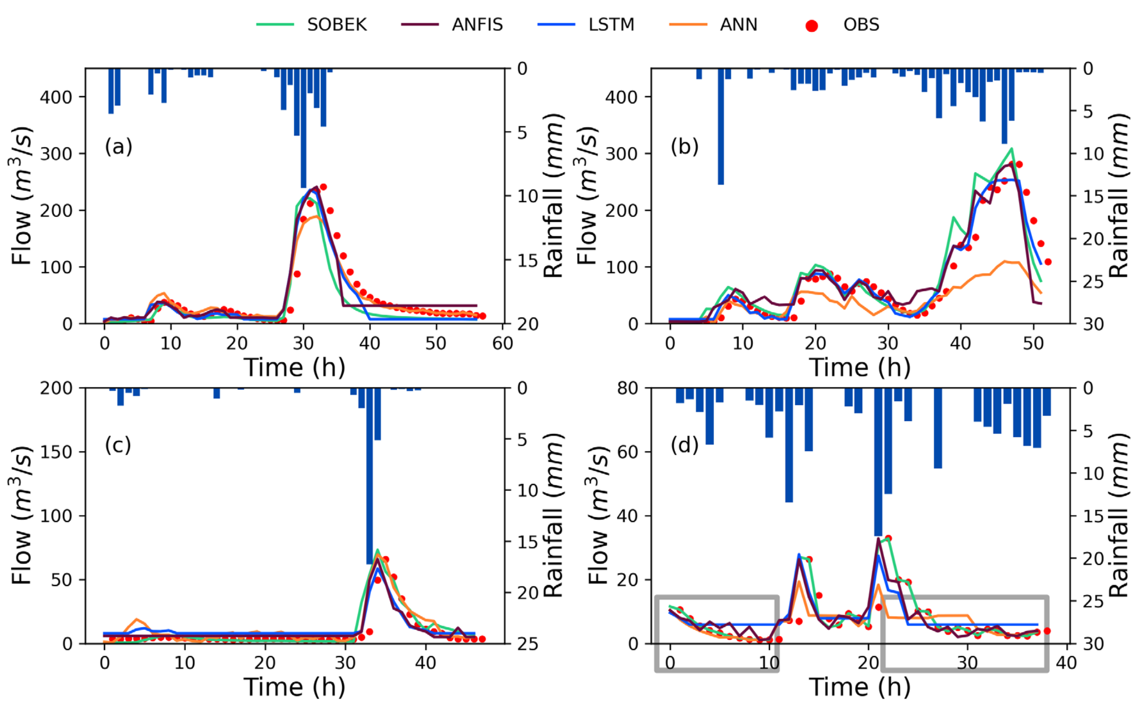

Figure 5 Observed and simulated hydrographs of streamflow in the two basins.
(a) - SJ station, (b) -LC station (c) - XT station in MZR, (d) - PSS station in PSR.

In addition, a comparison study on the correlation of observation data and simulation results is conducted, and the result is shown in Figure 6. As seen in Figure 6, the four models all have a high correlation (higher than 0.70) between the observation data and the simulation values in two river basins. Compared with other models, the distribution of the

ANFIS model (green dots) is more concentrated in Maozhou River (Figure 6(a)), and the SOBEK model (blue dots) is more concentrated in Pingshan River (Figure 6(b)), respectively. This reveals that the ANFIS and SOBEK model have great potential to simulate peak streamflow well in tidal river basin and in mountainous river basin, respectively.

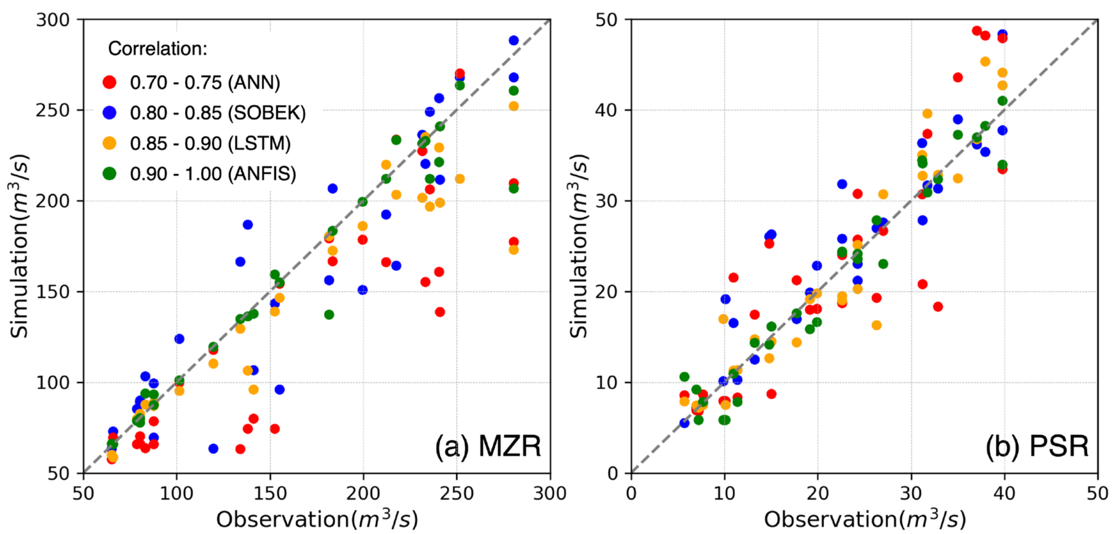

Figure 6 Observed and simulated hydrographs of streamflow in MZR (a) and PSR (b) basins.

4. Discussion

From the above research, all four models have potential to simulate the streamflow of urban river basins well. Table 4 ranks the four models' performance in terms of accuracy for average streamflow simulation (AAS), accuracy for peak streamflow simulation (APS), model efficiency and overall choice in two river basins.

Table 4 Rank performance of four models in two river basins.

	Tidal river basin				Mountainous river basin			
	SOBEK	ANN	LSTM	ANFIS	SOBEK	ANN	LSTM	ANFIS
AAS	2nd	4th	1st	3rd	1st	4th	2nd	3rd
APS	3rd	4th	2nd	1st	1st	4th	3rd	2nd
Model Efficiency	4th	1st	2nd	3rd	4th	1st	2nd	3rd
Overall Choice			√		√			

Of the four models, the LSTM model shows a good ability in simulating average streamflow well in tidal river basin. The internal memory cells of the LSTM model (‘forgotten gate’ and ‘memory gate’) have the ability to filter data and memory data features making as neural network functions to simulate the average streamflow process could be the reason for its good performance (Kratzert et al., 2018; Sahoo et al., 2019; Sudriani et al., 2019). Besides, the low streamflow has little impact on the average streamflow prediction accuracy in tidal river basin and the physical-based models sometimes cannot apply well in tidal basins due to some uncertainty sources, such as excessive rainfall data and the tidal effects (Jung et al., 2018), so the LSTM model can simulate better than the SOBEK model.

The uneven slope distribution has great influence on the average streamflow simulation accuracy in mountainous river basin. The SOBEK model can therefore, exhibit better due to the ability of physically based models to response to the rainfall-streamflow process (Noor et al., 2014). However, the ANN model does not show high reliability in validation period. According to Hu et al. (2018) and Ahmad and Simonovic (2005), the hidden layer function of the ANN model has limitations and challenges when simulate insufficient data. The input streamflow event of validation period is not enough, so the ANN model cannot validate well. Moreover, the results of the ANFIS model are similar to those of the LSTM model and more stable than those of the ANN model, which relies on the fact that the ANFIS model combines the relationship structure of neural network models with the decision-making mechanism of fuzzy logic (Amutha and Porchelvan, 2011; Vetrivel and Elangovan, 2017).

Compared to physically based model, all neural network models have the ability to simulate and predict quickly. The physically based model needs several data for model set-up and the calculation time depends on the complexity of urban river system. Therefore, physically based model has disadvantages in prediction time (Ke et al., 2020; Sun et al., 2017). The prediction speed of the ANN model is the best among the three neural network models we used. The ANN model can complete nonlinear predictions by adjusting the number and type of neurons in the hidden layer and the weights carried by each neuron (Navale and Singh, 2020). The model structure is relatively simple and can verify a large amount of data, so the prediction time is the shortest. Compared to the ANN model, the LSTM model requires more data for training and validation, so the calculation time is longer than that of the ANN model. The ANFIS model combines the characteristics of fuzzy systems and neural networks and adjusts the model by adjusting the type and number of membership functions. Due to the complex structure and gradient learning, the computational cost of ANFIS is very high, and it has more difficulty dealing with a large amount of input data (Salleh et al., 2017).

For peak streamflow prediction, the ANN model performs the largest underestimations in two river basins. However, it exhibits good performance in low volume streamflow simulation. The poor performance of the ANN model on peak streamflow simulation is in line with Sudheer et al. (2003) who highlighted that the ANN model tends to underestimate the peak streamflow even after data transformation and the learning process of an ANN model will reward a correct response of the system to input by increasing the strength of the current matrix of nodal weights. Likewise, the LSTM model is unable to simulate the small volume streamflow in a flood event well in mountainous river basin (see Figure 5 (d)), as having a continual value of streamflow for a high quantile of training data seems to pose difficulty for the LSTM model to learn and calibrate (Kratzert et al., 2018). The ANFIS model has the best accuracy in tidal river basin and the SOBEK model has the best accuracy in mountainous river basin, respectively. The peak streamflow value in tidal river basin is relatively large, the ANFIS model has a greater ability to train large data, as it combines the characteristics of fuzzy systems and neural networks, making it most appropriate for tidal river basins. The complex mechanisms of river system and lack of data are two main challenges in peak streamflow simulation of mountainous river basins (Zhang et al., 2013). The physically based model can reflect the complex mechanisms between rainfall and streamflow, and neural network models may have insufficient training data due to the low frequency of peak streamflow in mountainous river basins (Sudheer et al., 2003; Yang et al., 2019).

5. Conclusion

An evaluation on the performance of three neural network models and one physically based model for the streamflow simulation driven by hourly precipitation data in one tidal river basin and one mountainous river basin of Shenzhen has been conducted in this study. The following major findings are drawn:

(1) The four models we used are able to capture streamflow simulation well in two river basins. Specifically, the LSTM model is reliable in terms of average streamflow simulation in tidal river basin with the lowest bias but underestimates the small volume streamflow. It is suitable for tidal river basins that low streamflow has little impact on the simulation accuracy. The SOBEK model shows reliability in simulating average streamflow in mountainous river basin, while needs large basic data for model calibration and validation.

(2) All neural network models used in this research present high simulation speed. The three neural network models can predict a three-hour rainfall event in less than 2 minutes. The ANN model shows great reliability in prediction speed, is suitable for scenarios that require high forecasting speed such as emergency flooding management.

(3) The ANFIS model has best accuracy for peak streamflow simulation in tidal river basin, is suitable for extreme flooding prediction. For example, in urban flooding management, decision makers need to obtain extreme value of a flooding event in order to facilitate flooding management and decision-making. The SOBEK model has best accuracy of peak streamflow simulation in mountainous river basin, is suitable for river basins with complex mechanism systems.

(4) Overall, the LSTM model can compensate for physically based models in streamflow simulation in complex urban river systems, by giving fast streamflow predictions with usually acceptable accuracy.

Our research provides scientific support for the application in flood early warning and emergency preparedness in the context of flood risk management in the urban area.

Declaration of Competing 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.

Acknowledgements

This work was supported by the National Key R& D Program of China (2018YFE0206200), the National Natural Science Foundation of China (Grant no. 41671113 and 51761135024) and the High-level Special Funding of the Southern University of Science and Technology (Grant no. G02296302, G02296402). We would like to show our gratitude to Honglong Yang in Meteorological Bureau of Shenzhen. We also would like to show our gratitude to Meteorological Bureau of Shenzhen, Huadong Engineering Corporation Limited and Ecological Environment Bureau of Shenzhen.

Appendices

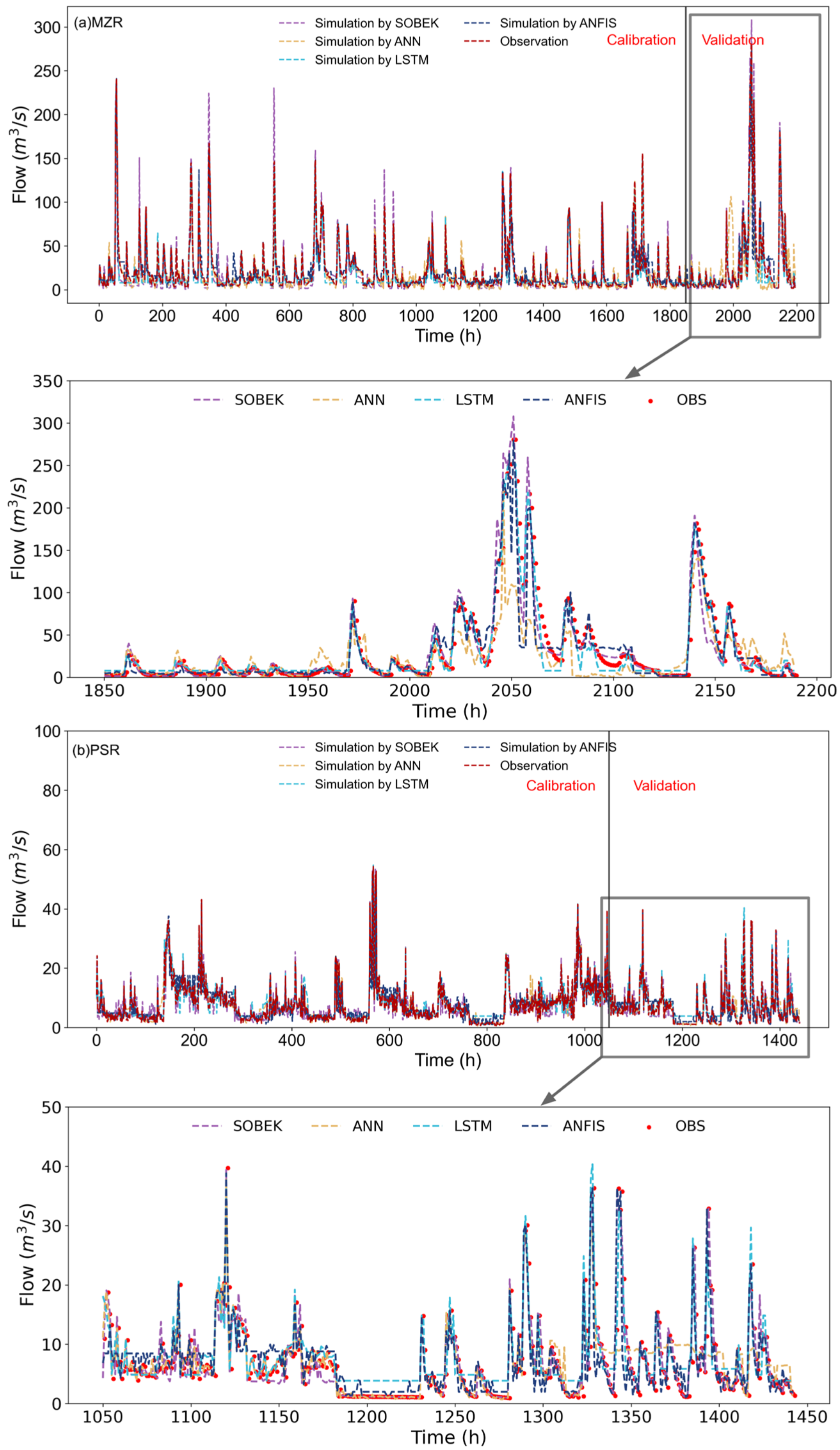

Figure A1 Details of calibration and validation results in two basins (a-MZR, b-PSR).

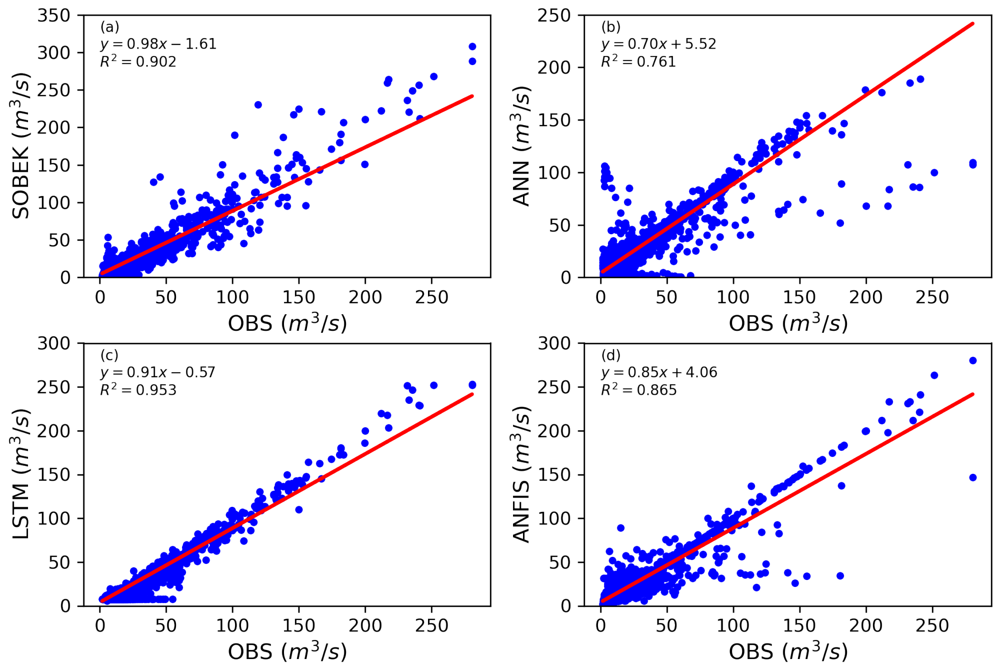

Figure A2 Comparison of observation (x-axis) and simulated results (y-axis) for the four models in terms of streamflow in MZR.

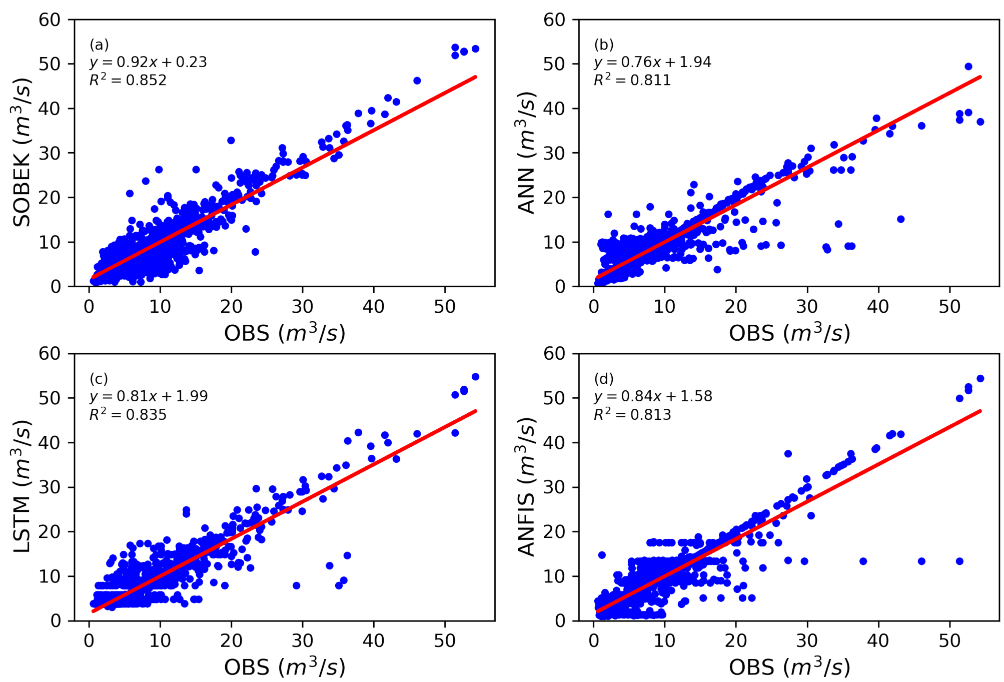

Figure A3 Comparison of observation (x-axis) and simulated results (y-axis) for the four models in terms of streamflow in PSR.

Reference

Adams, B.J., Fraser, H.G., Howard, C.D.D., Sami Hanafy, M., 1986. Meteorological data analysis for drainage system design. J. Environ. Eng. 112, 827–848.

Ahani, A., Shourian, M., Rahimi Rad, P., 2018. Performance Assessment of the Linear, Nonlinear and Nonparametric Data Driven Models in River Flow Forecasting. Water Resour. Manag. 32, 383–399. https://doi.org/10.1007/s11269-017-1792-5

Ahmad, S., Simonovic, S.P., 2005. An artificial neural network model for generating hydrograph from hydro-meteorological parameters. J. Hydrol. 315, 236–251. https://doi.org/10.1016/j.jhydrol.2005.03.032

Aichouri, I., Hani, A., Bougherira, N., Djabri, L., Chaffai, H., Lallahem, S., 2015. River Flow Model Using Artificial Neural Networks. Energy Procedia 74, 1007–1014. https://doi.org/10.1016/j.egypro.2015.07.832

Amutha, R., Porchelvan, P., 2011. Seasonal Prediction of Groundwater Levels Using Anfis 1, 98–108.

Ang, K.K., Quek, C., 2005. Rspop: Rough set–based pseudo outer-product fuzzy rule identification algorithm. Neural Comput. 17, 205–243.

Anghileri, D., Giudici, F., Castelletti, A., Burlando, P., 2016. Advancing reservoir operation description in physically based hydrological models, in: EGUGA. pp. EPSC2016-10097.

Archetti, R., Bolognesi, A., Casadio, A., Maglionico, M., 2011. Development of flood probability charts for urban drainage network in coastal areas through a simplified joint assessment approach. Hydrol. Earth Syst. Sci. 15, 3115–3122. https://doi.org/10.5194/hess-15-3115-2011

Ballesteros-Cánovas, J.A., Rodríguez-Morata, C., Garófano-Gómez, V., Rubiales, J.M., Sánchez-Salguero, R., Stoffel, M., 2015. Unravelling past flash flood activity in a forested mountain catchment of the Spanish Central System. J. Hydrol. 529, 468–479. https://doi.org/10.1016/j.jhydrol.2014.11.027

Botto, A., Belluco, E., Camporese, M., 2018. Multi-source data assimilation for physically based hydrological modeling of an experimental hillslope. Hydrol. Earth Syst. Sci. 22, 4251.

Bruneau, P., Gascuel-Odoux, C., Robin, P., Merot, P., Beven, K., 1995. Sensitivity to space and time resolution of a hydrological model using digital elevation data. Hydrol. Process. 9, 69–81. https://doi.org/10.1002/hyp.3360090107

Chang, W., Chen, X., 2018. Monthly rainfall-runoffmodeling at watershed scale: A comparative study of data-driven and theory-driven approaches. Water (Switzerland) 10, 1–21. https://doi.org/10.3390/w10091116

Chen, R.S., Lu, S.H., Kang, E.S., Ji, X. Bin, Zhang, Z., Yang, Y., Qing, W., 2008. A distributed water-heat coupled model for mountainous watershed of an inland river basin of Northwest China (I) model structure and equations. Environ. Geol. 53, 1299–1309. https://doi.org/10.1007/s00254-007-0738-2

Chen, Y., Li, J., Xu, H., 2016. Improving flood forecasting capability of physically based distributed hydrological models by parameter optimization. Hydrol. Earth Syst. Sci. 20, 375.

Cui, X., Guo, R., 2006. Hydrological characteristics of Maozhou River Basin. China Rural Water Hydropower 9, 57–60.

Daniell, T.M., 1991. Neural networks. Applications in hydrology and water resources engineering, in: National Conference Publication- Institute of Engineers. Australia.

Davenport, A.J., Gurnell, A.M., Armitage, P.D., 2004. Habitat survey and classification of urban rivers. River Res. Appl. 20, 687–704. https://doi.org/10.1002/rra.785

Dawson, R.J., Speight, L., Hall, J.W., Djordjevic, S., Savic, D., Leandro, J., 2008. Attribution of flood risk in urban areas. J. Hydroinformatics 10, 275–288. https://doi.org/10.2166/hydro.2008.054

Duan, W., He, B., Nover, D., Fan, J., Yang, G., Chen, W., Meng, H., Liu, C., 2016. Floods and associated socioeconomic damages in China over the last century. Nat. Hazards 82, 401–413. https://doi.org/10.1007/s11069-016-2207-2

Elshorbagy, A., Corzo, G., Srinivasulu, S., Solomatine, D.P., 2010. Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - Part 1: Concepts and methodology. Hydrol. Earth Syst. Sci. 14, 1931–1941. https://doi.org/10.5194/hess-14-1931-2010

Halff, A.H., Halff, H.M., Azmoodeh, M., 1993. Predicting runoff from rainfall using neural networks, in: Engineering Hydrology. ASCE, pp. 760–765.

Hattermann, F.F., Vetter, T., Breuer, L., Su, B., Daggupati, P., Donnelly, C., Fekete, B., Florke, F., Gosling, S.N., Hoffmann, P., Liersch, S., Masaki, Y., Motovilov, Y., Muller, C., Samaniego, L., Stacke, T., Wada, Y., Yang, T., Krysnaova, V., 2018. Sources of uncertainty in hydrological climate impact assessment: A cross-scale study. Environ. Res. Lett. 13. https://doi.org/10.1088/1748-9326/aa9938

Her, Y., Yoo, S.H., Cho, J., Hwang, S., Jeong, J., Seong, C., 2019. Uncertainty in hydrological analysis of climate change: multi-parameter vs. multi-GCM ensemble predictions. Sci. Rep. 9, 1–22. https://doi.org/10.1038/s41598-019-41334-7

Hu, C., Wu, Q., Li, H., Jian, S., Li, N., Lou, Z., 2018. Deep learning with a long short-term memory networks approach for rainfall-runoff simulation. Water (Switzerland) 10, 1–16. https://doi.org/10.3390/w10111543

Humphrey, G.B., Gibbs, M.S., Dandy, G.C., Maier, H.R., 2016. A hybrid approach to monthly streamflow forecasting: Integrating hydrological model outputs into a Bayesian artificial neural network. J. Hydrol. 540, 623–640. https://doi.org/10.1016/j.jhydrol.2016.06.026

Jiang, Y., Zevenbergen, C., Ma, Y., 2018. Urban pluvial flooding and stormwater management: A contemporary review of China's challenges and "sponge cities" strategy. Environ. Sci. Policy 80, 132–143. https://doi.org/10.1016/j.envsci.2017.11.016

Jung, Sungho ・ Cho, Hyoseob ・ Kim, Jeongyup ・ Lee, Giha aD, D.P.E.E., bW R H R F C O, K.N.U., E, M., 2018. Prediction of water level in a tidal river using a deep-learning based LSTM model 51, 1207–1216. https://doi.org/10.3741/JKWRA.2018.51.12.1207

Kashani, M., Ghorbani, M.A., Dinpashoh, Y., Shahmorad, S., 2016. Integration of Volterra model with artificial neural networks for rainfall-runoff simulation in forested catchment of northern Iran. J. Hydrol. 540, 340–354. https://doi.org/10.1016/j.jhydrol.2016.06.028

Ke, Q., Tian, X., Bricker, J., Tian, Z., Guan, G., Cai, H., Huang, X., Yang, H., Liu, J., 2020. Urban pluvial flooding prediction by machine learning approaches – a case study of Shenzhen city, China. Adv. Water Resour. 145, 103719. https://doi.org/10.1016/j.advwatres.2020.103719

Kim, J., Mohanty, B.P., 2017. A physically based hydrological connectivity algorithm for describing spatial patterns of soil moisture in the unsaturated zone. J. Geophys. Res. 122, 2096–2114. https://doi.org/10.1002/2016JD025591

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., Herrnegger, M., 2018. Rainfall – runoff modelling using Long Short-Term Memory (LSTM) networks 6005–6022.

Li, P., Zha, Y., Shi, L., Tso, C.H.M., Zhang, Y., Zeng, W., 2020. Comparison of the use of a physical-based model with data assimilation and machine learning methods for simulating soil water dynamics. J. Hydrol. 584, 124692. https://doi.org/10.1016/j.jhydrol.2020.124692

Lian, J.J., Xu, K., Ma, C., 2013. Joint impact of rainfall and tidal level on flood risk in a coastal city with a complex river network: A case study of Fuzhou City, China. Hydrol. Earth Syst. Sci. 17, 679–689. https://doi.org/10.5194/hess-17-679-2013

Liu, Y.R., Li, Y.P., Huang, G.H., Zhang, J.L., Fan, Y.R., 2017. A Bayesian-based multilevel factorial analysis method for analyzing parameter uncertainty of hydrological model. J. Hydrol. 553, 750–762. https://doi.org/10.1016/j.jhydrol.2017.08.048

Mernild, S.H., Liston, G.E., Hiemstra, C.A., Yde, J.C., Casassa, G., 2018. Annual river runoff variations and trends for the Andes Cordillera. J. Hydrometeorol. 19, 1167–1189. https://doi.org/10.1175/JHM-D-17-0094.1

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50, 885–900.

Nash, J.E., Sutcliffe, J. V., 1970. River flow forecasting through conceptual models part I - A discussion of principles. J. Hydrol. 10, 282–290. https://doi.org/10.1016/0022-

1694(70)90255-6

Navale, A., Singh, C., 2020. Topographic sensitivity of WRF-simulated rainfall patterns over the North West Himalayan region. Atmos. Res. 242, 105003. https://doi.org/10.1016/j.atmosres.2020.105003

Nikpour, M.R., Sanikhani, H., Babelan, S.M., Amuqin, S.N., 2019. Archive of SID Daily Rainfall – Runoff Modeling of Darreh-Rud River in Ardabil Province , Iran Archive of SID. pp. 144–146.

Noor, H., Vafakhah, M., Taheriyoun, M., Moghadasi, M., 2014. Hydrology modelling in Taleghan mountainous watershed using SWAT. J. Water L. Dev. 20, 11–18. https://doi.org/10.2478/jwld-2014-0003

Orton, P.M., Conticello, F.R., Cioffi, F., Hall, T.M., Georgas, N., Lall, U., Blumberg, A.F., MacManus, K., 2020. Flood hazard assessment from storm tides, rain and sea level rise for a tidal river estuary. Nat. Hazards 102, 729–757. https://doi.org/10.1007/s11069-018-3251-x

Peng, J., Wei, H., Wu, W., Liu, Y., Wang, Y., 2018. Storm flood disaster risk assessment in urban area based on the simulation of land use scenarios: A case of Maozhou Watershed in Shenzhen City. ACTA Ecol. Sin. 38, 3741–3755.

Sahoo, B.B., Jha, R., Singh, A., Kumar, D., 2019. Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting. Acta Geophys. 67, 1471–1481. https://doi.org/10.1007/s11600-019-00330-1

Salleh, M.N.M., Talpur, N., Hussain, K., 2017. Adaptive neuro-fuzzy inference system: Overview, strengths, limitations, and solutions, in: International Conference on Data Mining and Big Data. Springer, pp. 527–535.

Schuol, J., Abbaspour, K.C., Yang, H., Srinivasan, R., Zehnder, A.J.B., 2008. Modeling blue and green water availability in Africa. Water Resour. Res. 44, 1–18. https://doi.org/10.1029/2007WR006609

Shi, P.J., Yuan, Y., Zheng, J., Wang, J.A., Ge, Y., Qiu, G.Y., 2007. The effect of land use/cover change on surface runoff in Shenzhen region, China. Catena 69, 31–35. https://doi.org/10.1016/j.catena.2006.04.015

Shoaib, M., Shamseldin, A.Y., Melville, B.W., 2014. Comparative study of different wavelet based neural network models for rainfall-runoff modeling. J. Hydrol. 515, 47–58. https://doi.org/10.1016/j.jhydrol.2014.04.055

Sikorska, A.E., Renard, B., 2017. Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges. Adv. Water Resour. 105, 51–66. https://doi.org/10.1016/j.advwatres.2017.04.011

SMEEB, (Municipal Ecological Environment Bureau of Shenzhen), 2018. Shenzhen Pingshan River water body compliance plan. http://meeb.sz.gov.cn/ydmh/zwgk/hjxw/content/post_2090474.html

Song, J., Li, W., 2019. Linkage between the environment and individual resilience to urban flooding: A case study of shenzhen, china. Int. J. Environ. Res. Public Health 16. https://doi.org/10.3390/ijerph16142559

SRTM, (Shuttle Radar Topography Mission), 2020. The DEM data source. https://srtm.csi.cgiar.org

Sudheer, K.P., Nayak, P.C., Ramasastri, K.S., 2003. Improving peak flow estimates in artificial neural network river flow models. Hydrol. Process. 17, 677–686. https://doi.org/10.1002/hyp.5103

Sudriani, Y., Ridwansyah, I., A Rustini, H., 2019. Long short term memory (LSTM) recurrent neural network (RNN) for discharge level prediction and forecast in Cimandiri river, Indonesia. IOP Conf. Ser. Earth Environ. Sci. 299. https://doi.org/10.1088/1755-1315/299/1/012037

Sun, W., Wang, Y., Wang, G., Cui, X., Yu, J., Zuo, D., Xu, Z., 2017. Physically based distributed hydrological model calibration based on a short period of streamflow data: Case studies in four Chinese basins. Hydrol. Earth Syst. Sci. 21, 251–265. https://doi.org/10.5194/hess-21-251-2017

SZN, (Shenzhen news), 2020. A heavy rain caused water logging in Baoan. https://mp.weixin.qq.com/s/AegwOMY7HTl66pCgZ9RDkw

Taylor, K.E., 2001. in a Single Diagram. J. Geophys. Res. 106, 7183–7192. https://doi.org/10.1029/2000JD900719

Tikhamarine, Y., Souag-Gamane, D., Ahmed, A.N., Sammen, S.S., Kisi, O., Huang, Y.F., El-Shafie, A., 2020. Rainfall-runoff modelling using improved machine learning methods: Harris hawks optimizer vs. particle swarm optimization. J. Hydrol. 589, 125133. https://doi.org/10.1016/j.jhydrol.2020.125133

Vetrivel, N., Elangovan, K., 2017. Application of ANN and ANFIS model on monthly groundwater level fluctuation in lower Bhavani river basin. Indian J. Geo-Marine Sci. 46, 2114–2121.

Wang, S., Yao, X., 2013. Using class imbalance learning for software defect prediction. IEEE Trans. Reliab. 62, 434–443.

Xiong, Y., Su, Z., Zhang, Y., Liang, H., 2010. Evaluation of Water Pollution of the Pingshan River in Shenzhen. Environ. Sci. Surv. 29, 79–81.

Xu, D., Ouyang, Z., Wu, T., Han, B., 2020. Dynamic trends of urban flooding mitigation services in Shenzhen, China. Sustain. 12, 1–11. https://doi.org/10.3390/su12114799

Yan, H., He, X., Lei, Y., Wang, Y., Su, H., Jiang, S., 2019. Land use-induced change in trophic state of Shenzhen Bay (South China) over the past half-century. Mar. Pollut. Bull. 145, 208–213. https://doi.org/10.1016/j.marpolbul.2019.05.046

Yang, S., Yang, D., Chen, J., Santisirisomboon, J., Lu, W., Zhao, B., 2020. A physical process and machine learning combined hydrological model for daily streamflow simulations of large watersheds with limited observation data. J. Hydrol. 590, 125206. https://doi.org/10.1016/j.jhydrol.2020.125206

Yang, S., Yang, D., Chen, J., Zhao, B., 2019. Real-time reservoir operation using recurrent neural networks and inflow forecast from a distributed hydrological model. J. Hydrol. 579, 124229. https://doi.org/10.1016/j.jhydrol.2019.124229

Yoon, H., Jun, S.C., Hyun, Y., Bae, G.O., Lee, K.K., 2011. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. J. Hydrol. 396, 128–138. https://doi.org/10.1016/j.jhydrol.2010.11.002

Zhang, D., Martinez, N., Lindholm, G., Ratnaweera, H., 2018. Manage Sewer In-Line Storage Control Using Hydraulic Model and Recurrent Neural Network. Water Resour. Manag. 32, 2079–2098. https://doi.org/10.1007/s11269-018-1919-3

Zhang, J., Zhang, Y., Song, J., Cheng, L., Kumar Paul, P., Gan, R., Shi, X., Luo, Z., Zhao, P., 2020. Large-scale baseflow index prediction using hydrological modelling, linear and multilevel regression approaches. J. Hydrol. 585, 124780. https://doi.org/10.1016/j.jhydrol.2020.124780

Zhang, W., Li, A., Jiang, X., 2013. Preliminary study on computing the area ofmountain regions in China based on DEM (in Chinese). Geogr. Geo-information Sci. 5, 58–63.

Ziegler, A.D., Lim, H.S., Jachowski, N.R., Wasson, R.J., 2012. Reduce urban flood vulnerability. Nature 481, 145.