Engineering 24 (2023) 230–239

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
Hydraulic Engineering—Article

Enhanced LSTM Model for Daily Runoff Prediction in the Upper Huai River Basin, China

Yuanyuan Man [a], Qinli Yang [a,b,*], Junming Shao [c], Guoqing Wang [d], Linlong Bai [e], Yunhong Xue [e]

[a] School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China
[b] Yangtze Delta Region Institute (Huzhou), University of Electronic Science and Technology of China, Huzhou 313001, China
[c] School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[d] State Key Laboratory of Hydrology–Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing 210029, China
[e] Xinyang Hydrology and Water Resources Survey Bureau, Xinyang 450003, China

ARTICLE INFO

Article history:
Received 8 June 2021
Revised 10 December 2021
Accepted 17 December 2021
Available online 28 April 2022

Keywords:
Runoff prediction
Long short-term memory
Upper Huai River Basin
Extreme runoff
Loss function

ABSTRACT

Runoff prediction is of great significance to flood defense. However, due to the complexity and randomness of the runoff process, it is hard to predict daily runoff accurately, especially for peak runoff. To address this issue, this study proposes an enhanced long short-term memory (LSTM) model for runoff prediction, where novel loss functions are introduced and feature extractors are integrated. Two loss functions (peak error tan*h* (PET), peak error swish (PES)) are designed to strengthen the importance of the peak runoff's prediction while weakening the weight of the normal runoff's prediction. The feature extractor consisting of three LSTM networks is established for each meteorological station, aiming to extract temporal features of the input data at each station. Taking the upper Huai River Basin in China as a case study, daily runoff from 1960–2016 is predicted using the enhanced LSTM model. Results indicate that the enhanced LSTM model performed well, achieving Nash–Sutcliffe efficiency (NSE) coefficient ranging from 0.917–0.924 during the validation period (November 2005–December 2016), outperforming the widely used lumped hydrological models (Australian Water Balance Model (AWBM), Sacramento, SimHyd and Tank Model) and the data-driven models (artificial neural network (ANN), support vector regression (SVR), and gated recurrent units (GRU)). The enhanced LSTM with PES as loss function performed best on extreme runoff prediction with a mean NSE for floods of 0.873. In addition, precipitation at a meteorological station with a higher altitude contributes more runoff prediction than the closest stations. This study provides an effective tool for daily runoff prediction, which will benefit the basin's flood defense and water security management.

1. Introduction

Among all the natural disasters, flood is the most frequent type, and it endangers the population and property [1–3]. Floods are also increasing in frequency and intensity due to climate change and human activities [4]. Since ancient times, human beings have made considerable efforts to combat floods, including structural and non-structural methods [5]. Structural methods are the most visible flood defense measures, such as levees, bunds, dams, and weirs. Meanwhile, the non-structural methods like flood forecasting models and systems, which facilitate disaster preparedness planning, have played prominent roles in flood risk mitigation. With the development of computer science and hydrological science in recent decades, flood prediction models have been leveraged worldwide to tackle the flood issue. However, due to its complexity and nonlinearity, flood prediction is a non-trivial task that demands advanced models and higher accuracy.

In general, runoff prediction and flood forecast models can be categorized into process-based and data-driven models [6,7]. The process-based model, which dates back to the 1960s, is a mathematical formulation that explicitly represents the hydrologic state variables and fluxes. Up to the present, numerous process-based hydrological models, including lumped and distributed types, have been proposed and widely applied in runoff prediction [8]. For instance, TOPMODEL is one of the first models to explicitly use topographic data to reflect a basin's hydrological response characteristics in the model formulation [9]. It was used for runoff

* Corresponding author.
E-mail address: qinli.yang@uestc.edu.cn (Q. Yang).

https://doi.org/10.1016/j.eng.2021.12.022
2095-8099/

prediction in different regions [10,11]. The Soil and Water Assessment Tool (SWAT) [12], a semi-distributed hydrological model, has undergone sustaining development since its establishment. Diverse modules such as radar precipitation [13], groundwater units [14], and snowmelt units [15] have been successively integrated. SWAT model performed well on runoff prediction [16,17]. Australian Water Balance Model (AWBM) is a lumped hydrological model that can use daily rainfall to estimate daily runoff [18,19]. Despite their widespread applications and advantage on interpretability, process-based models still have some drawbacks on flood prediction, such as over-parameterization, high complexity, a wealth of expert knowledge, and high requirements for data.

Alternatively, the data-driven models based on statistical theory can learn the relationships among influencing factors and runoff automatically, which is not only cost-effective but also highly efficient. A large number of data-driven models for runoff prediction have been proposed and applied in practices [20], including artificial neural network (ANN) [21], support vector machine (SVM) [22], neuro-fuzzy [23], adaptive neuro-fuzzy inference system (ANFIS) [24], wavelet neural network [25], and multilayer perceptron (MLP), and the like. Among them, ANN is the most popular data-driven model for runoff prediction with good generalization ability and relatively high accuracy among all the models mentioned above. However, it fails in modeling the time dependency of data sequences and predicting the peak value accurately. Time dependency refers to the autocorrelation relationship between the previous data and the current data in a time series, which are often difficult to express directly in equations [26]. In comparison, the long short-term memory (LSTM) model with a gated mechanism stands out due to its excellent performance, simple architecture, and superior time dependency ability.

LSTM is a deep learning model proposed by Hochreiter and Schmidhuber [27] to solve the complexity of information storing in the long sequence backpropagation process. Because of the outstanding performance on long sequence tasks, LSTM has been widely applied in various fields since its inception, especially in time series. In recent years, LSTM has attracted much attention in hydrology [28–32]. For instance, Kratzer et al. [33] studied runoff prediction using LSTM in 241 catchments in the Catchment Attributes and Meteorology for Large-sample Studies (CAMELS) data set and proved that LSTM could get good simulation results in most basins. Meanwhile, the effectiveness of LSTM in watershed-scale transformation is also verified, which shows the great potential of LSTM. However, peak flow prediction is still a challenge for LSTM [34,35].

The challenge of peak flow prediction mainly lies in two aspects. One is how to identify the important input features for runoff prediction. Another is how to optimize the model's objective function to achieve good prediction results. Although the LSTM model can capture the temporal features of the input data series (e.g., rainfall), it ignores the spatial heterogeneity of the temporal features. Actually, in practice, the impacts of different rain gauges on runoff prediction are different. Therefore, the structure of the LSTM model needs to be modified and improved. LSTM generally takes mean square error (MSE) as the loss function regarding model optimization. However, MSE treats samples with different prediction errors equally, which fails to emphasize the importance of peak flow prediction. Ding et al. [36] used extreme value theory to design a novel loss function (extreme value loss (EVL)) in the extreme events model, whose main idea is to adjust the weights on extreme events so that the model pays more attention to extreme values in the parameter optimization process. Therefore, new loss functions can be designed for LSTM to improve its ability to deal with peak flow prediction.

Accurate flood prediction for the upper Huai River Basin is significant for flood management of the whole Huai River Basin. The Huai River, one of the seven major rivers in China, is located in eastern China between the Yangtze River and Yellow River. Due to its continental monsoon climate with a complex and variable atmospheric system, the Huai River Basin is prone to flood, suffering floods about once per two years [37]. The upper Huai River Basin is the essential control basin for the upper Huai River. In the past few decades, many efforts have been made on runoff prediction in the upper Huai River Basin. For instance, Liu et al. [38] used the fully distributed model Topographic Kinematic Approximation and Integration (TOPKAPI) to predict runoff with a six-hour time step. Lv et al. [39] constructed an LSTM model for cyclic prediction and achieved good results for hourly flood forecasting. However, most existing studies are fragmented, focusing on hourly runoff prediction or flood events analysis relatively quickly but lack continuous daily runoff prediction over a long period. Furthermore, in the context of climate change, population growth, and economic development, the upper Huai River Basin has experienced significant climate change and land-use change over the past few decades [40,41]. Runoff prediction in the changing environment in the upper Huai River Basin is in high demand.

This study aims to propose an enhanced LSTM model to improve the accuracy of daily runoff prediction to facilitate flood defense. The objectives are as follows: ① to explore the trends of the hydro-meteorological variables and land use in the upper Huai River Basin in the last decades; ② to predict runoff and flood in the upper Huai River Basin by using a structurally improved LSTM with novel loss functions designed for peak flows; ③ to compare the improved LSTM model with different existing models to proof its outperformance on runoff prediction in the study area.

2. Study area and data acquisition

2.1. Study area

The upper Huai River Basin (as shown in Fig. 1) located in the east of China with a drainage area of 10 190 km^2 is selected as the study area. The basin belongs to the subtropical monsoon humid climate zone, with an average temperature of 15.43 °C. Monsoon mainly affects precipitation, with an average annual precipitation of 1043 mm, 50 % concentrated from June to September. The drainage system of the basin is distributed in a dendritic form, and the average runoff depth is about 350 mm. The basin topography is high in the west and low in the east, with an average elevation of 47 m (Fig. 1). Most of the land in the basin is cultivated, with a small number of cities and woodland. Six meteorological stations covering the spatial heterogeneity were set up in/around the basin.

2.2. Data acquisition

Climatic data, including daily precipitation, pan evaporation, and temperature from 1951–2016, were collected from the China Meteorological Data Service Center[‡]. Daily runoff ($m^3 \cdot s^{-1}$) during 1951–2016 at the Xixian hydrometric station was provided by Henan Hydrology Bureau. The SRTM1 DEM data[§] with a spatial resolution of 30 m was used in this study. Landsat 5 Thematic Mapper (TM) Collection 1 Level-1 data was used for land use classification of the study area in 1987, and Landsat 8 Operational Land Imager (OLI) Collection 1 Level-1 data was used for that in 2016.

3. Methodology

The non-parametric rank-based Mann-Kendall test (recommended by the World Meteorological Organization) [42] and linear

[‡] https://data.cma.cn.
[§] https://glovis.usgs.gov.

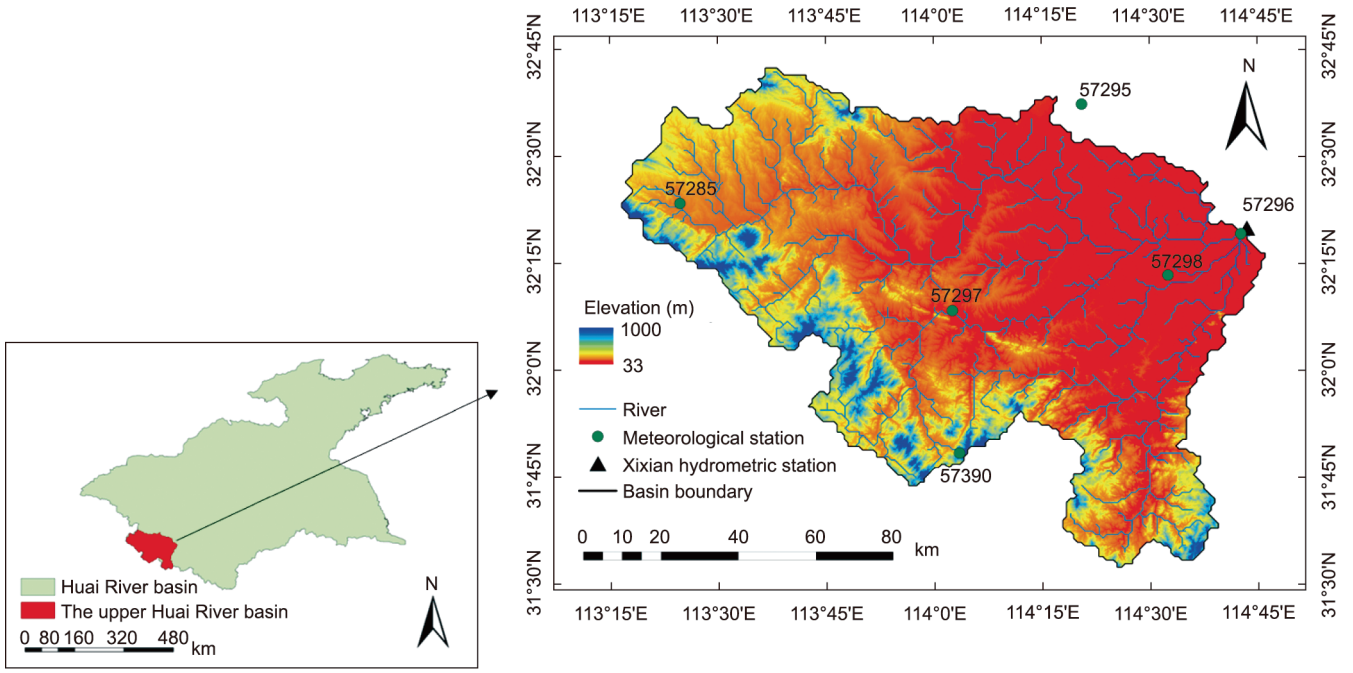

Fig. 1. River system and the distribution of hydro-meteorological stations in the upper Huai River Basin, China.

regression method are applied to identify the trends of meteorological and hydrological variables. Before presenting the enhanced LSTM model, the basic structure of the LSTM model is introduced in Section 3.1. For comparison with the existing methods, representative models are selected and briefly introduced in Section 3.2. In evaluating the performances of different models, specific evaluation metrics are finally presented at the end of this section.

3.1. LSTM network

LSTM is developed from the recurrent neural network (RNN) [43]. Compared with RNN, LSTM adds a forgetting mechanism, which can also solve the gradient explosion problem. In the structure of LSTM (as shown in Fig. 2), a particular unit called a memory cell is similar to an accumulator and a gated neuron. The next sequential step has a parallel weight and copies its state's actual value and accumulation. The LSTM has a self-connection mechanism controlled by a multiplication gate that learns and decides when to clear the memory content by another unit. For better understanding, define time as subscript t, hidden state as h, cell state as C, input as x, the output of input gate as i, the output of forget gate as f, the output of output gate as o, and the output of the reserved portion of the original loop layer as $\hat{C}$. σ is an activation function, such as sigmoid and Rectified Linear Unit (ReLU).

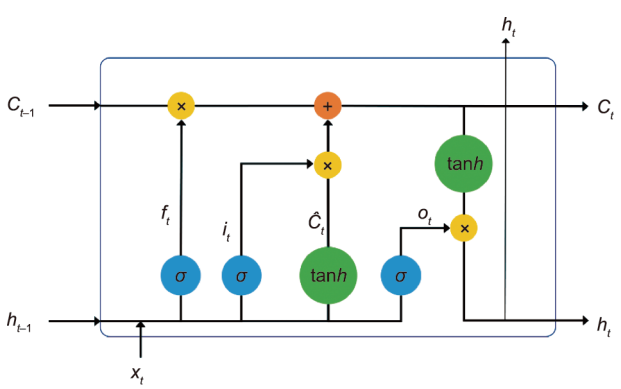

Fig. 2. The structure of the LSTM model. The subscript "t" here represents the current time, h_{t-1} and C_{t-1} represent the hidden state and cell state respectively received from the previous node. In other words, C and h represent long-term and short-term changes, respectively. The rest are intermediate variables.

LSTM comprises three gates: input gate, output gate, and forget gate. Its forward propagation process can be expressed by Eqs. (1)–(3):

Input gate:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{1}$$

Output gate:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{2}$$

Forget gate:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{3}$$

The cell information C_t and the hidden information h_t are updated by Eqs. (4)–(6):

$$\hat{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \tag{4}$$

$$C_t = f_t \times C_{t-1} + i_t \times \hat{C}_t \tag{5}$$

$$h_t = o_t \times \tanh(C_t) \tag{6}$$

where W is the weight matrix and b is the bias, which are updated and optimized in the training process.

Compared with RNN, after adding the forgetting mechanism f_t, the LSTM no longer passes all the historical information backward but selectively forgets part of the historical content, memorizes part of the historical context, and adds new input information to the backward transfer. Then it uses the backpropagation algorithm [44] to update the parameters and optimize the model.

3.2. Enhanced LSTM model for runoff prediction

3.2.1. Integrating feature extractor into the LSTM model

The flowchart of the enhanced LSTM model is illustrated in Fig. 3. It includes four layers: input layer, feature extractor,

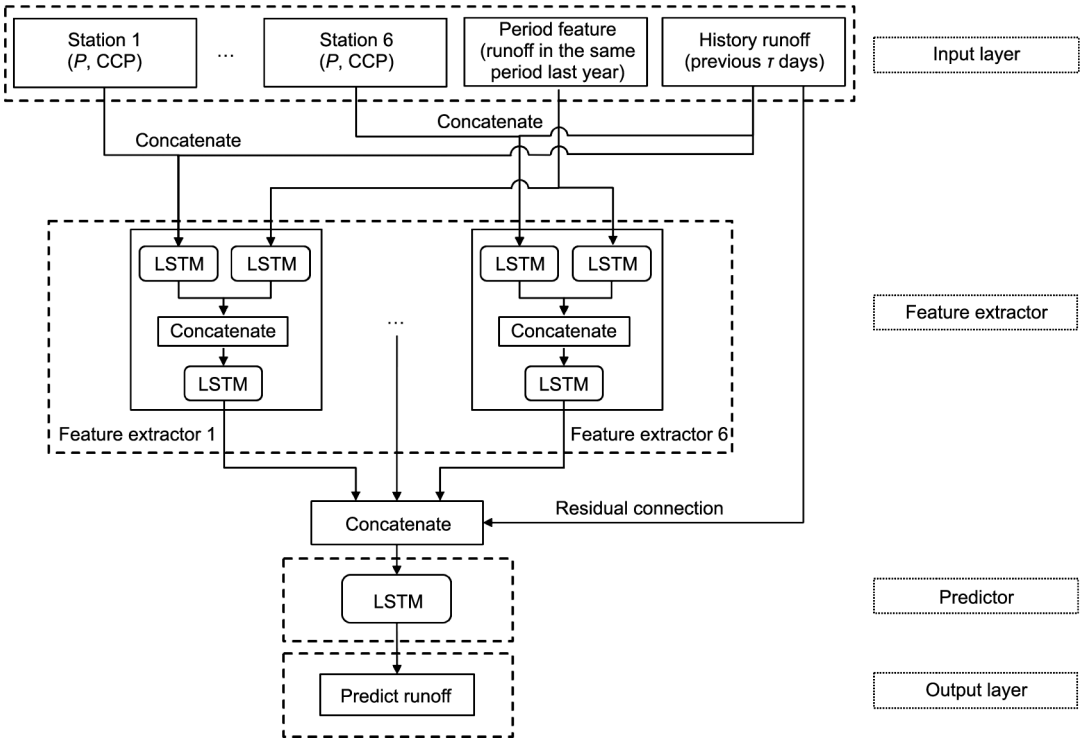

Fig. 3. The framework of the enhanced LSTM model for daily runoff prediction.

predictor, and output layer. Unlike the original LSTM model, the enhanced LSTM model has an integrated feature extractor. The motivation of this integration is to identify the critical features for the runoff prediction task. For this current work, a separate feature extractor is proposed for each meteorological station at Xixian hydrometric station, assuming that precipitation at different stations might contribute to different runoff generation.

A feature extractor is composed of three LSTMs. Specifically, it takes precipitation (P), conditional cumulative precipitation (CCP), and runoff with a step length ($\hat{o}$ days) before the time t as the input features. It also takes the historical period feature's runoff data of the same period last year. The first two LSTMs are used to extract long-term trend features and short-term change features, respectively. Then, the two parts of the extracted features are spliced together to output site features through the third LSTM.

Subsequently, the output features of each station obtained from each feature extractor and the historical runoff with a step length of $\hat{o}$ are spliced together using a residual connection technique. Finally, an LSTM is used to predict runoff at time t. Historical runoff provides constraints for boundary conditions, breaks the network's symmetry, and improves the characterization ability of the network [45,46].

3.2.2. Designing new loss functions to improve peak runoff prediction

The MSE is generally used as the loss function for typical regression problems. The formula is as follows:

$$\mathrm{MSE} = \frac{\sum_t^n \left(y_t^{\mathrm{pre}} - y_t\right)^2}{n} \tag{7}$$

where y_t^{pre} is the prediction and y_t is the observation.

The above formulation indicates that MSE treats runoff prediction errors equally no matter the runoff prediction error, either high or low. However, the peak runoff prediction error is a concern in flood forecasting, and the normal runoff prediction error is relatively unimportant. To solve this problem, two new loss functions are designed to increase the importance of peak loss to improve the accuracy of peak runoff prediction.

(1) **Peak error tan*h* (PET).** Since the error of the extreme value is more significant than that of the normal flow, the weight can be increased for the larger error in the MSE. Therefore, a tan*h* function is added alongside the MSE to amplify larger errors and reduce small errors simultaneously. The formula is as follows:

$$\mathrm{PET} = \tan h(\mathrm{MSE}) \tag{8}$$

Its function curve is shown in Fig. 4. When the independent variable is larger, the function value is also larger; when the independent variable is smaller, the function value will correspondingly become smaller. Therefore, the purpose of amplifying large errors and reducing small errors is achieved.

(2) **Peak error swish (PES).** Swish is a new activation function proposed by Google [47]. Swish is adapted to local response

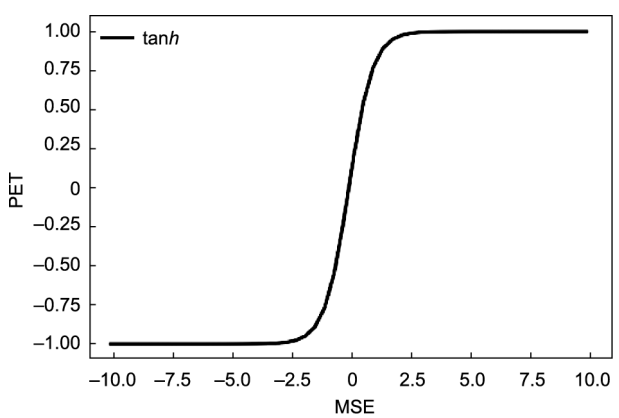

Fig. 4. tan*h* function graph.

normalization, and the effect of fully connected layers above 40 is much better than other activation functions. Furthermore, it has shown better performance than the current best activation function on different data sets. The Swish function image is shown in Fig. 5.

The expression of PES after fusing MSE is

$$\mathrm{PES} = (\mathrm{MSE}) \times \mathrm{sigmoid}(\mathrm{MSE}) \quad (9)$$

3.2.3. The training of the enhanced LSTM model

The enhanced LSTM model training adopts the mini-batch training technique. A batch size of 64, epoch size of 200, and the Adam optimizer are used for training the model. Three loss functions, MSE, PET, and PES are used. To keep the features of each input data within the same numerical range, the data is normalized by using Eq. (10) as follows:

$$X_{\mathrm{norm}} = \frac{X - X_{\mathrm{min}}}{X_{\mathrm{max}} - X_{\mathrm{min}}} \quad (10)$$

where X is the variable, X_{min} and X_{max} are the minimum and maximum values of the variable X, respectively. X_{norm} represents the normalized data.

3.3. Selected comparative models for runoff prediction

In verifying, the effectiveness of the enhanced LSTM model on daily runoff prediction, three data-driven models (support vector regression (SVR), ANN, gated recurrent units (GRU)) and four lumped hydrologic models (AWBM, Sacramento, SimHyd, and Tank Model) are selected for comparison.

SVR is an important application of SVM [22] in regression tasks. SVR works by finding a regression plane so that all the data in a set are closest to that plane. To achieve the regression task of nonlinear data, SVR can also use a nonlinear kernel to get a hyperplane to fit the data. SVR is favored for its simplicity, efficiency, and superior performance. The ANN model is an information processing system that mimics brain functions according to biological neural networks [21,48]. It consists of the input layer, hidden layers, and output layer. As one of the classical machine learning models, it is also the basis for most deep learning models. Due to the high flexibility of ANN structure, suitable network structure and loss function can be designed according to the specific applications. ANN learning is robust to errors in training data and has been successfully applied to many fields.

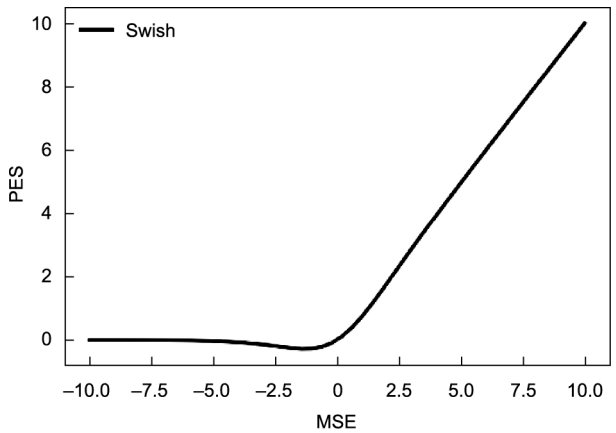

Fig. 5. Swish function graph.

GRU is a type of RNN [49]. In many cases, it performs similarly to LSTM, but it is easier to train and essentially improves training efficiency. SVR and ANN are both widely used and representative traditional data-driven models. To verify the advance of the proposed model over the traditional data-driven methods, SVR and ANN are selected for comparison. In addition, comparing the proposed model with GRU, which has a similar effect with LSTM, can indicate the structural superiority of the enhanced LSTM.

The AWBM is a catchment water balance model that links rainfall and evapotranspiration to runoff through daily or sub-daily data [18]. It calculates rainfall losses for flood hydrological models. The model contains five stores, including three surface stores, a base flow store, and a surface runoff routing store. Sacramento model is a lumped catchment water balance model with 16 parameters and performs at a daily time step [50]. The runoff production can be divided into five parts: direct runoff, surface runoff, soil flow, fast groundwater, and slow groundwater. A linear reservoir simulated medium flow, fast groundwater, and slow groundwater. SimHyd is a conceptual rainfall-runoff model with seven parameters containing three stores for interception loss, soil moisture, and groundwater [51]. Sugawara et al. [52] developed the tank model to explain a catchment's water flow phenomena. It is a straightforward model, composing four tanks placed vertically in series. The precipitation is poured into the top tank, and evaporation is subtracted from the top tank downward. As each tank is emptied, the evaporation gap begins at the next tank until all tanks have been emptied. The output of the side outlets is the calculated runoff [53]. The four classical lumped hydrological models have been successfully applied worldwide in catchment runoff simulation and prediction. Comparison with the selected lumped hydrological models intends to verify the superiority of the proposed model over the traditional physical models.

3.4. Evaluation metrics

To evaluate the performance of different (environmental) models, please refer to the literature published by Bennett et al. [54]. In this study, the Nash–Sutcliffe model efficiency coefficient (NSE),

Table 1
Statistical information of trend analysis for the hydro-meteorological variables on annual and seasonal scales during 1951 and 2016 for the upper Huai River Basin, China.

Statistic	Annual		Spring		Summer		Autumn		Winter	
	Slope	*Z*	Slope	*Z*	Slope	*Z*	Slope	*Z*	Slope	*Z*
P (mm)	−0.210	0.020	−0.190	−1.000	0.010	0.560	−0.140	−0.060	0.110	−0.130
Runoff depth (mm)	−1.770	−1.270	−0.270	−1.670	−0.950	−1.500	−0.630	−0.530	0.070	0.370
Temperature (°C)	0.018[a]	4.710[a]	0.024[a]	2.710[a]	0.021[a]	4.180[a]	0.010	−0.650	0.017[a]	3.430[a]
Evp (mm)	−3.960[a]	−4.910[a]	−0.610[b]	−2.350[b]	−0.410	−1.330	−2.100[a]	−5.150[a]	−0.840[a]	−3.460[a]
D (*P* > 0 mm)	−0.208	−2.080[b]	−0.057	−1.640	−0.086	−2.340[b]	0.059	−0.200	−0.125	−2.460[b]
D (*P* > 25 mm)	−0.029	−1.560	0.003	−4.570[a]	0.002	−0.890	−0.025	−2.500[b]	−0.010	−5.290[a]
D (*P* > 50 mm)	−0.002	−1.580	0.000	0.000	0.001	−2.810[a]	−0.005	−2.830[a]	0.000	0.000
P_{max}	−0.250	−0.960	0.035	0.390	−0.116	−0.400	0.013	0.200	−0.044	−0.980

Z means the statistic value derived from the Mann–Kendall test method, Evp: pan evaporation; *D* (*P* > 0 mm), *D* (*P* > 25 mm), and *D* (*P* > 50 mm) represent the number of days with daily precipitation large than 0 mm, 25 mm, and 50 mm, respectively; P_{max}: maximum daily precipitation.
[a,b] indicate that the variable has a significant change at the significance level of 0.01 (2.58) and 0.05 (1.96), respectively.

mean absolute error (MAE), root mean square error (RMSE), relative volume error (RE), qualification rate (QR), and NSE_{flood} are selected as the evaluation criteria. QR refers to the forecast flood QR [55]. NSE_{flood} refers to the average NSE for forecasting floods (Eqs. (11–16)):

$$NSE = 1 - \frac{\sum_{t=1}^{N}\left(Y_t^{est} - Y_t\right)^2}{\sum_{t=1}^{N}\left(Y_t - \bar{Y}_t\right)^2} \tag{11}$$

$$MAE = \frac{\sum_{t=1}^{N}\left|Y_t^{est} - Y_t\right|}{N} \tag{12}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}\left(Y_t^{est} - Y_t\right)^2}{N}} \tag{13}$$

$$RE = \frac{\sum_{t=1}^{N}\left(Y_t^{est} - Y_t\right)}{\sum_{t=1}^{N} Y_t} \tag{14}$$

$$QR = \frac{m}{n} \times 100\% \tag{15}$$

$$NSE_{flood} = \frac{\sum_{t=1}^{m} NSE_t}{m} \tag{16}$$

where Y_t^{est} is the simulated runoff by the model at the time t, Y_t represents the observed runoff at the time t, $\bar{Y}_t$ indicates the average value of the observed runoff, and N means the number of observations. m refers to the number of floods where the relative error between the predicted flood peak and the real flood peak is less than 20%, n is the total number of flood events. NSE_{flood} refers to NSE during floods.

4. Results and discussion

4.1. Trends of the hydro-meteorological variables

In exploring the trends of meteorological and hydrological elements in the study area, precipitation, runoff depth, temperature, and pan evaporation during 1951–2016 are analyzed on annual and seasonal scales using the Mann–Kendall method and linear regression method. The statistical results are shown in Table 1. In addition, Fig. 6 illustrates the inter-annual changes of these hydro-meteorological variables.

In general, it can be seen that precipitation and runoff depth have insignificant decreasing trends on the annual scale in the upper Huai River Basin. The temperature has a clear upward trend (0.18 °C per ten years), relatively lower than the average temperature increase rate (0.24 °C per ten years) in China during 1951–2018 [56]. Fig. 6(b) shows that the annual pan evaporation has a significant downward trend (3.96 mm per year). Similar to findings

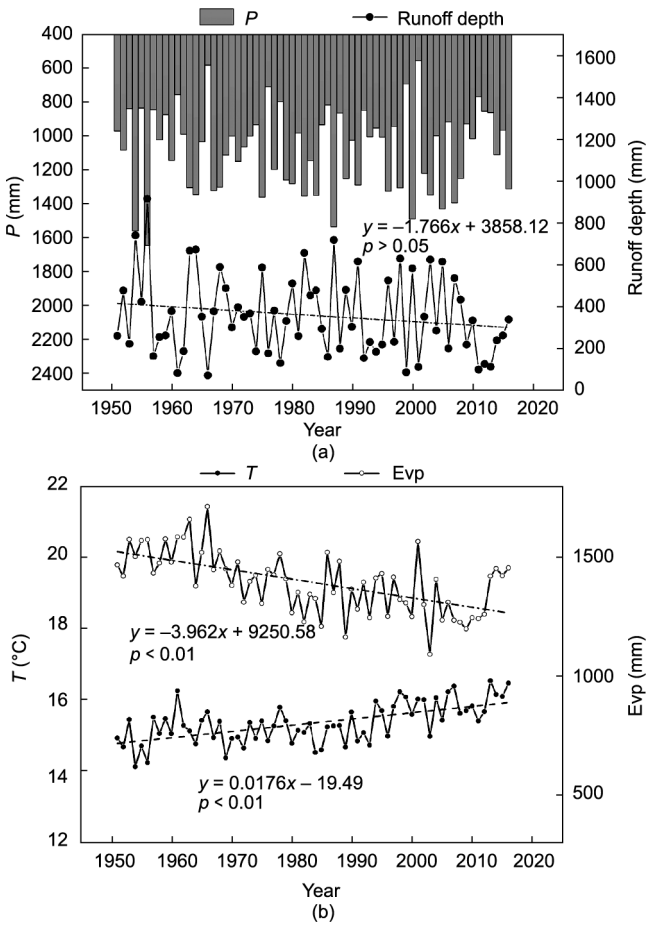

Fig. 6. The inter-annual change of hydro-meteorological variables. (a) *P* and runoff depth; (b) temperature and pan evaporation during 1951 and 2016 for the upper Huai River Basin, China.

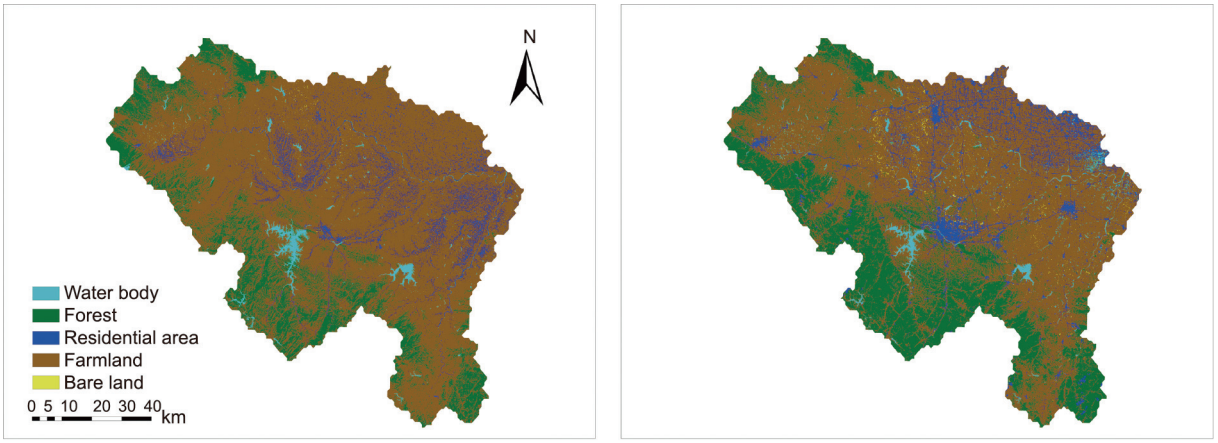

Fig. 7. Land use classification results in (a) 1987 and (b) 2016 for the upper Huai River Basin, China.

in Han et al. [57], the evaporation paradox existed in the upper Huai River Basin, which can be attributed to changes in solar radiation, relative humidity, and wind speed.

Specifically, consistent with annual precipitation, precipitation in spring and autumn shows insignificant downward trends, while precipitation in summer and winter shows insignificant upward trends. Except for runoff depth in winter (insignificant increase), the trend of seasonal runoff depth is consistent with that on the annual scale (insignificant decrease). The seasonal temperature has risen significantly except for autumn. The trend of pan evaporation on the seasonal scale is consistent with that on the annual scale, but the decreasing trend in summer (0.41 mm per year) is not significant.

In addition, the number of days for different intensities of precipitation is also studied (as shown in Table 1). The number of rainfall days ($P > 0$) shows a downward trend at the 95% confidence level and remains consistent on the seasonal scale, especially in summer and winter. The number of days with rainfall greater than 25 and 50 mm mainly showed a downward trend on the annual and seasonal scales. The annual maximum rainfall also changed insignificantly. Therefore, extreme precipitation events in the basin did not show a noticeable change in the context of climate change.

Overall, the above results reveal that from 1951 to 2016, the temperature in the upper Huai River Basin increased significantly, the pan evaporation decreased significantly, and the precipitation and runoff were stable. Although the hydro-meteorological variables change, the hydrological status remains stable in the upper Huai River in general.

4.2. Land use change in the upper Huai River Basin

Besides climate change, land use change is another critical driving factor for the hydrological cycle. For instance, land use change may influence runoff generation and formation processes via changing the characteristics of the underlying surface. Different land change patterns (e.g., afforestation, deforestation, urbanization) exert different impacts on runoff. In recent decades, land use change in the upper Huai River Basin needs to be investigated to better understand the impact of land use on runoff change. To this end, we classified land use in the upper Huai River Basin into five types (water body, forest, residential area, farmland, and bare land). We analyzed the land use status in 1987 and 2016 based on Landsat images using the random forest algorithm. The land use classification results in 1987 and 2016 are shown in Fig. 7, and

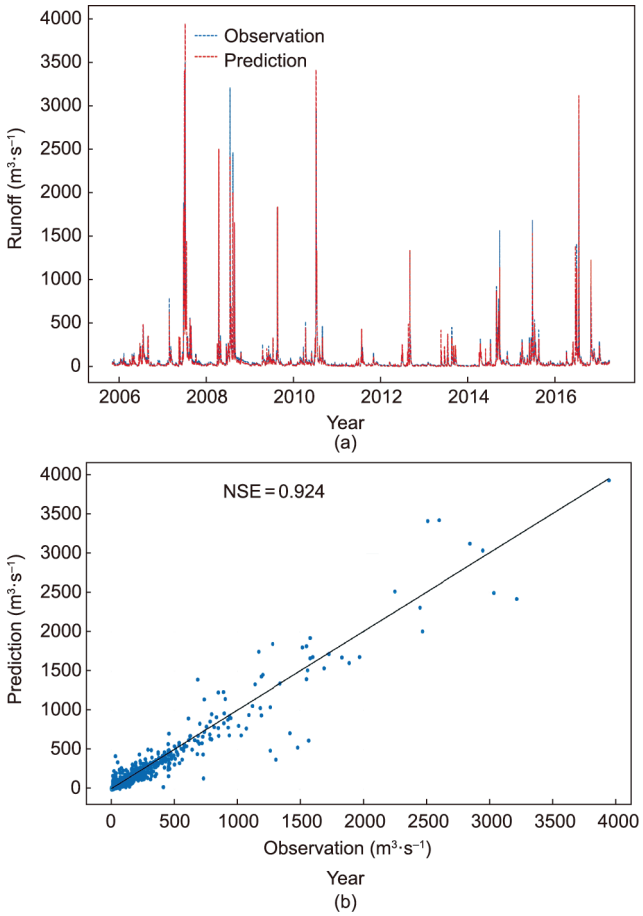

Fig. 8. The observed and predicted runoff based on the enchanted LSTM model with PET as loss function during November 2005–December 2016 for the upper Huai River Basin, China.

Table 2
Transition matrix of land use changes from 1987 to 2016 for the upper Huai River Basin, China (km^2).

2016 land use	1987 land use					
	Water body	Forest	Residential area	Farmland	Bare land	Total
Water body	110.786	2.4710	28.621	78.196	0.077	220.151
Forest	4.497	1446.7000	1.844	1091.665	0.001	2544.707
Residential area	3.656	17.74530	76.751	575.269	0.073	673.494
Farmland	48.270	294.1500	386.876	5594.016	0.329	6323.641
Bare land	0.843	1.1574	2.159	52.247	0.024	56.430
Total	168.052	1762.2240	496.251	7391.393	0.504	9865.830

Table 3
Performance comparison of different runoff prediction models.

Category	Models	NSE (%)	MAE($m^3 \cdot s^{-1}$)	RMSE($m^3 \cdot s^{-1}$)	RE	QR (%)	NSE_{flood} (%)
Data driven models	LSTM (PES)	91.7	24.80	59.29	0.18	92.3	87.3
	LSTM (PET)	92.4	21.03	56.82	−0.14	84.6	86.3
	LSTM (MSE)	91.7	17.73	52.92	0.07	77.8	74.7
	SVR	86.9	28.78	74.35	0.15	50.0	83.8
	ANN	89.8	25.33	65.74	−0.05	53.8	86.7
	GRU	90.4	24.35	63.87	0.16	69.2	85.8
Lumped hydrologic models	AWBM	55.9	48.43	136.69	0.05	41.7	28.7
	Sacramento	67.4	51.95	117.52	−0.12	66.7	68.8
	SimHyd	56.1	55.26	136.38	0.07	25	38.0
	Tank Model	57.7	48.70	133.84	0.30	25	32.0

the transfer matrix of land use change from 1987 to 2016 is presented in Table 2.

Intuitively, Fig. 7 shows that the area of farmland has decreased, while the residential area and forest cover have expanded obviously. Quantitatively, the area of farmland has been reduced by about 1100 km^2 in various forms, mainly converted to residential area and forest land. The residential area has expanded by about 180 km^2, which increases the impervious surface area and may result in urban flooding. The decrease of farmland area and the increase of forest (about 780 km^2) is consistent with the "returning cropland to the forest" initiative proposed by the Chinese government. The initiative aims to protect the ecological environment while developing the economy. As forests play important roles in regulating rainfall and reducing flood peaks, the increase of the forest area may lead to more water for water conservation while less water for flooding.

4.3. Runoff prediction based on the enhanced LSTM model

As precipitation data at some meteorological stations are missing during 1951 and 1959, daily runoff during 1960 and 2016 is selected for the upper Huai River Basin prediction framework based on the enhanced LSTM model. Specifically, data during January 1960–October 2005 are used for training and the remaining data (November 2005–December 2016) for testing. To predict daily runoff at the day (d), the input data, precipitation (P_{d-6}, P_{d-5}, ..., P_{d-1}), conditional cumulative precipitation (CCP_{d-6}, CCP_{d-5}, ..., CCP_{d-1}, cumulative days are two days), runoff (R_{d-6}, R_{d-5}, ..., R_{d-1}), and runoff data at the same time last year are used. CCP means accumulated precipitation over specified days. For instance, CCP_{d-1} with two cumulative days represents the accumulated precipitation of the day $d-1$ and the day $d-2$. The performances of the enhanced LSTM models on daily runoff prediction are listed in Table 3 It can be noticed that the NSE of overall runoff prediction based on the enhanced LSTM models all exceeds 91%. Notably, the improved LSTM model with the loss function of PET shows the best performance, achieving an NSE of 0.924, as demonstrated in Fig. 8.

To further assess the performance of the enhanced LSTM model on extreme runoff (flood) prediction, QR and NSE_{flood} are calculated and presented in Table 3. Here, only a peak flow rate more than or equal to 1000 $m^3 \cdot s^{-1}$ is regarded as a flood event. As a result, 13 flood events were identified in the test dataset. It can be found from Table 3 that the improved LSTM with PES as loss function performed best, achieving a QR of 92.3% and NSE_{flood} (average NSE during the flood period) of 0.873. Based on this model, Fig. 9 displays the prediction results of nine representative flood events.

In evaluating the contributions of precipitation at different meteorological stations to runoff prediction, the Pearson correlation coefficient (PCC) was calculated between the extracted site feature (output of feature extractor for each meteorological station in Fig. 2) and the predicted runoff during the test period (November 2005–December 2016). The results are shown in Table 4. The higher the PCC is, the more influential the feature is to runoff prediction.

Table 4 indicates that the extracted site features have negative correlations with predicted runoff when the loss functions of the model are MSE and PES. In contrast, their correlations are positive when PET is used as the loss function. The absolute value of the correlation coefficient is the most crucial information, representing

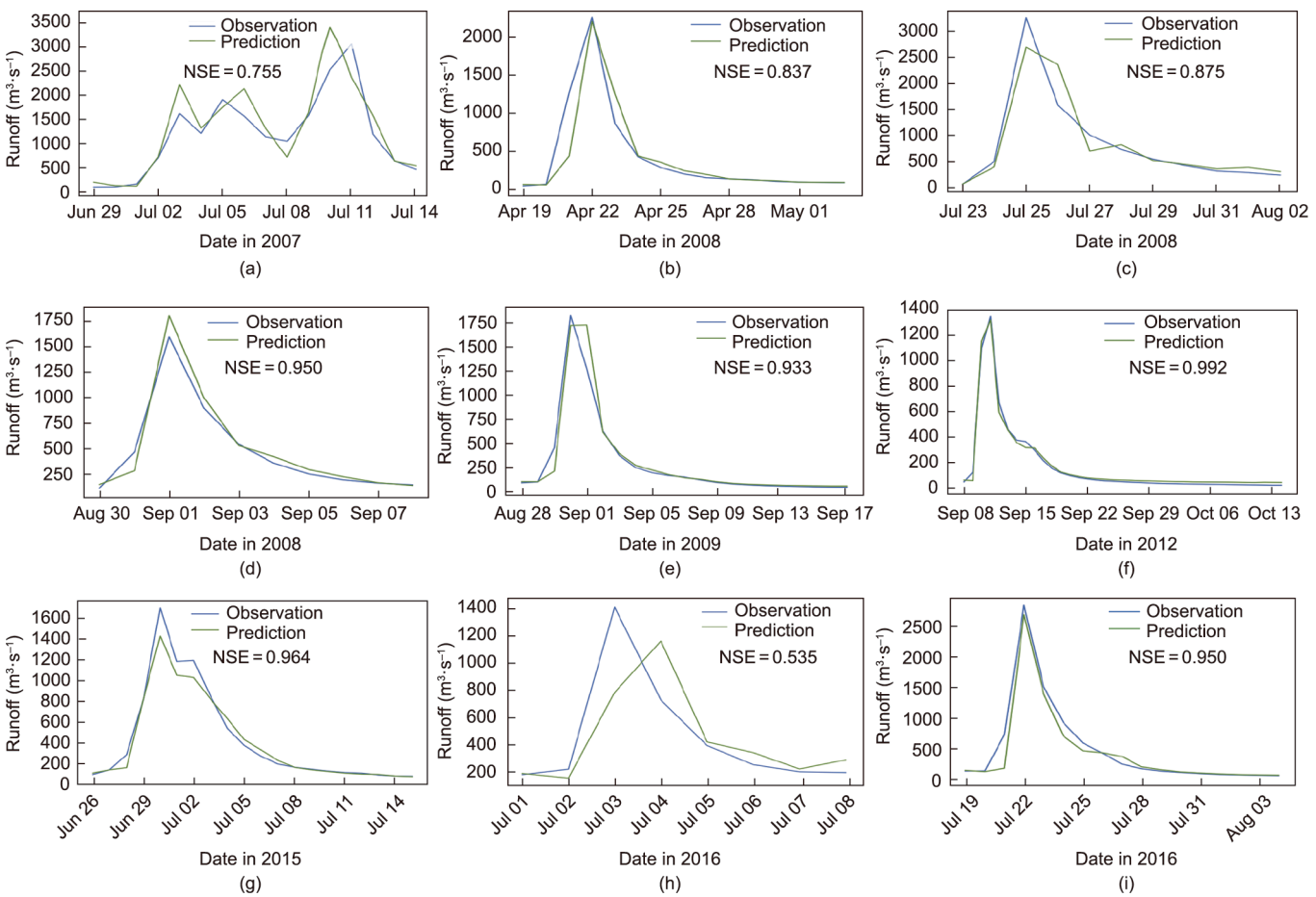

Fig. 9. Performance evaluation on peak runoff prediction based on the improved LSTM with PES as loss function for the upper Huai River Basin, China.

Table 4
PCC between the extracted site features and runoff predictions with different loss functions.

Site No.	PCC–MSE	PCC–PET	PCC–PES
57285	−0.464	0.397	**−0.672**
57295	−0.455	0.416	−0.625
57296	−0.486	0.278	−0.560
57297	−0.489	0.360	−0.560
57298	−0.458	0.347	−0.534
57390	**−0.547**	**0.430**	−0.636

Note: All the results have a significant change at the significance level of 0.01. MSE, PET, and PES represent the different loss functions used in the enhanced LSTM model. The data in bold indicate the strongest correlations.

the correlation between the features extracted from the meteorological station and the predicted runoff. The data in bold indicates the strongest correlation The positive and negative signs are determined by the parameter training process inside the model, working in coordination with the whole model and eliminating the effect of the sign within the model itself. It can be seen that PES significantly improves the correlation coefficient between site features and predicted runoff, with the highest PCC of −0.672 at station 57285. Besides, meteorological stations with high altitudes contribute to runoff prediction than those closer to the hydrometric station. For instance, stations 57390 and 57285, which are the first and second highest altitudes, have relatively higher PCC than those closer to the hydrometric station.

4.4. Comparison with the selected comparative models

The performances of different models are shown in Table 3. In general, the data-driven models outperformed the lumped hydrological models on daily runoff prediction. The enhanced LSTM model achieved better results than comparative models such as SVR, ANN, and GRU among the data-driven models. Although GRU has a simpler structure and higher training efficiency, its overall performance is slightly lower than the enhanced LSTM. Their discrepancy in performance is even more apparent in the flood forecasting results, which indicates that the enhanced LSTM model structure and the two loss functions (PET and PES) have improved the model's ability to predict runoff and flood. Specifically, for the overall runoff, the LSTM model with PET as loss function exhibits the highest NSE of 0.924. However, regarding the flood peak prediction, the enhanced LSTM model with PES performed best, achieving a QR of 92.3% and the NSE during the flood period (NSE_{flood}) of 0.873, respectively. The above results imply that the enhanced LSTM model with PES as a loss function has more significant potential for flood forecasting.

In the training process, data is firstly normalized. So MSE is usually between zero and one. There is no problem of output saturation for the PET loss function when there is a rapid rise stage, but the rising rate gradually decreases. One possible explanation is that PET magnifies MSE as a whole. The larger the MSE, the larger the PET (MSE is in the range of 0–1). Therefore, when the model optimizes the parameters, it is optimized as a whole, so the overall prediction result of runoff is better. When the horizontal axis is between zero and one, PES is approximately linear, but it is not. PES reduces the MSE, but the first derivative of PES gradually increases and approaches a constant value. Therefore, when the MSE is close to zero, the rising rate of PES is slower, and when the MSE is close to one, the increasing rate of PES is faster. The first derivative increasing monotonically may be why PES has a better effect on improving flood flow with more significant errors.

The floods in the upper reaches of the Huai River are concentrated from June to August, and most of them are caused by heavy rain, which brings enormous flood control pressure to the middle and lower reaches. To protect most social property and people's safety, the middle and lower reaches of the Huai River have operated part of the flood storage areas mainly for agriculture and industry and built many reservoirs in the upper reaches. Enhanced LSTM can timely and accurately predict the arrival time of flood peak and the flow, which is significant to the operation of reservoirs and flood storage areas. On the other hand, the reservoirs also greatly influence runoff. Therefore, some of the prediction errors may be caused by the reservoirs. Future works could consider the impact of the reservoirs on runoff prediction.

5. Conclusions

To improve the accuracy of runoff prediction, this study proposed an enhanced LSTM model. Based on the original LSTM model, a feature extractor is designed for discriminative feature identification, and two novel loss functions (PET and PES) designed for flood peak prediction are also introduced. Taking the upper Huai River Basin as a case study, the enhanced LSTM model was applied and evaluated for daily runoff prediction during 1960–2016.

During the study period, the upper Huai River Basin has experienced a warmer and drier climate but a relatively stable hydrologic status. Land use has changed between 1987 and 2016, mainly from cropland to forest and residential areas. Results indicate that the enhanced LSTM performed well on daily runoff prediction (achieving the highest NSE of 0.924), outperforming the comparative models (i.e., SVR, ANN, GRU, AWBM, Sacramento, SimHyd, and Tank Model). Regarding the flood peak prediction, the enhanced LSTM model with *PES* as loss function performed best with the QR of 92.3% and the NSE during the flood period (NSE_{flood}) of 0.873. Furthermore, there is a correlation between the meteorological station's extracted features and predicted runoff. The correlation reveals that precipitation at a station with a high elevation contributes more to runoff generation than those closer to the hydrometric station. This study provides an effective tool for daily runoff prediction, which would benefit the local basin's flood risk management and water security.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (52079026), the National Key Research and Development Program of China (2021YFC3201100), the National Natural Science Foundation of China (41830863 and 61976044), Sichuan Science and Technology Program (2020YFH0037), the Belt and Road Fund on Water and Sustainability of the State Key Laboratory of Hydrology–Water Resources and Hydraulic Engineering (2019nkzd02), the Open Research Fund of State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin (IWHR-SKL-201911), and the Fundamental Research Funds for the Central Universities (ZYGX2019Z014). Many thanks go to Mr. Cobbinah M. Bernard, who significantly contributed to the manuscript revision.

Compliance with ethics guidelines

Yuanyuan Man, Qinli Yang, Junming Shao, Guoqing Wang, Linlong Bai, and Yunhong Xue declare that they have no conflict of interest or financial conflicts to disclose.

References

[1] Barendrecht MH, Sairam N, Cumiskey L, Metin AD, Holz F, Priest SJ, et al. Needed: a systems approach to improve flood risk mitigation through private precautionary measures. Water Secur 2020;11:100080.

[2] Moore RJ, Bell VA, Jones DA. Forecasting for flood warning. C R Geosci 2005;337(1–2):203–17.
[3] Williams BS, Das A, Johnston P, Luo B, Lindenschmidt KE. Measuring the skill of an operational ice jam flood forecasting system. Int J Disast Risk Re 2021;52:102001.
[4] Mizutori M, Guha-Sapir D. Economic losses, poverty and disasters 1998–2017 Report. Brussels: the United Nations Office for Disaster Risk Reduction, Geneva & Center for Research on the Epidemiology of Disasters; 2018.
[5] Kundzewicz ZW, Takeuchi K. Flood protection and management: quo vadimus? Hydrol Sci J 1999;44(3):417–32.
[6] Wang L, Li X, Ma C, Bai Y. Improving the prediction accuracy of monthly streamflow using a data-driven model based on a double-processing strategy. J Hydrol: X 2019;573:733–45.
[7] Zhang H, Yang Q, Shao J, Wang G. Dynamic streamflow simulation via online gradient-boosted regression tree. J Hydrol Eng 2019;24(10):04019041.
[8] Clark MP, Bierkens MFP, Samaniego L, Woods RA, Uijlenhoet R, Bennett KE, et al. The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism. Hydrol Earth Syst Sci 2017;21(7):3427–40.
[9] Beven KJ, Kirkby MJ, Freer JE, Lamb R. A history of TOPMODEL. Hydrol Earth Syst Sci 2021;25(2):527–49.
[10] Wang J, Zhang J, Wang G, Song X, Yang X, Wang Y. Ensemble flood simulation for the typical catchment in humid climatic zone by using multiple hydrological models. Proc IAHS 2020;383:213–22.
[11] Beven KJ, Kirkby MJ, Schofield N, Tagg AF. Testing a physically-based flood forecasting model (TOPMODEL) for three U.K. catchments. J Hydrol: X 1984;69(1–4):119–43.
[12] Arnold JG, Srinivasan R, Muttiah RS, Williams JR. Large area hydrologic modeling and assessment part I: model development. JAWRA 1998;34(1):73–89.
[13] Jayakrishnan R, Srinivasan R, Santhi C, Arnold JG. Advances in the application of the SWAT model for water resources management. Hydrol Processes 2005;19(3):749–62.
[14] Kim NW, Chung IM, Won YS, Arnold JG. Development and application of the integrated SWAT–MODFLOW model. J Hydrol: X 2008;356(1–2):1–16.
[15] Fontaine TA, Cruickshank TS, Arnold JG, Hotchkiss RH. Development of a snowfall–snowmelt routine for mountainous terrain for the soil water assessment tool (SWAT). J Hydrol: X 2002;262(1–4):209–23.
[16] Azimi S, Dariane AB, Modanesi S, Bauer-Marschallinger B, Bindlish R, Wagner W, et al. Assimilation of Sentinel 1 and SMAP-based satellite soil moisture retrievals into SWAT hydrological model: the impact of satellite revisit time and product spatial resolution on flood simulations in small basins. J Hydrol: X 2020;581:124367.
[17] Rajib A, Liu Z, Merwade V, Tavakoly AA, Follum ML. Towards a large-scale locally relevant flood inundation modeling framework using SWAT and LISFLOOD-FP. J Hydrol: X 2020;581:124406.
[18] Boughton W. The Australian water balance model. Environ Model Softw 2004;19(10):943–56.
[19] Boughton W, Chiew F. Estimating runoff in ungauged catchments from rainfall, PET and the AWBM model. Environ Model Softw 2007;22(4):476–87.
[20] Mosavi A, Ozturk P, Chau K. Flood prediction using machine learning models: literature review. Water 2018;10(11):1536.
[21] Wang W, Gelder PH, Vrijling JK, Ma J. Forecasting daily streamflow using hybrid ANN models. J Hydrol: X 2006;324(1–4):383–99.
[22] Huang S, Chang J, Huang Q, Chen Y. Monthly streamflow prediction using modified EMD-based support vector machine. J Hydrol: X 2014;511:764–75.
[23] Mukerji A, Chatterjee C, Raghuwanshi NS. Flood Forecasting Using ANN, neuro-fuzzy, and neuro-GA models. J Hydrol Eng 2009;14(6):647–52.
[24] Arora A, Arabameri A, Pandey M, Siddiqui MA, Shukla UK, Bui DT, et al. Optimization of state-of-the-art fuzzy-metaheuristic ANFIS-based machine learning models for flood susceptibility prediction mapping in the Middle Ganga Plain, India. Sci Total Environ 2021;750:141565.
[25] Kasiviswanathan KS, He J, Sudheer KP, Tay JH. Potential application of wavelet neural network ensemble to forecast streamflow for flood management. J Hydrol: X 2016;536:161–73.
[26] Zeng C, Wang Q, Wang W, Li T, Shwartz L. Online inference for time-varying temporal dependency discovery from time series. In: Joshi J, Karypis G, Liu L, Hu X, Ak R, Xia Y, editors. 2016 IEEE International Conference on Big Data (Big Data); 2016 Dec 5–8; Washington, DC, USA. Washington: Institute of Electrical and Electronics Engineers (IEEE); 2016. p. 1281–90.
[27] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8):1735–80.
[28] Kratzert F, Klotz D, Shalev G, Klambauer G, Hochreiter S, Nearing G. Benchmarking a catchment-aware long short-term memory network (LSTM) for large-scale hydrological modeling. Hydrol Earth Syst Sci 2019;23:5089–110.
[29] Le XH, Ho HV, Lee G, Jung S. Application of long short-term memory (LSTM) neural network for flood forecasting. Water 2019;11(7):1387.
[30] Wu Y, Liu Z, Xu W, Feng J, Palaiahnakote S, Lu T. Context-aware attention LSTM network for flood prediction. In: 24th International Conference on Pattern Recognition (ICPR); 2018 Aug 20–24; Beijing, China. Washington: Institute of Electrical and Electronics Engineers (IEEE); 2018. p. 1301–6.
[31] Zhang J, Zhu Y, Zhang X, Ye M, Yang J. Developing a long short-term memory (LSTM) based model for predicting water table depth in agricultural areas. J Hydrol: X 2018;561:918–29.
[32] Kao IF, Zhou Y, Chang LC, Chang FJ. Exploring a long short-term memory based encoder-decoder framework for multi-step-ahead flood forecasting. J Hydrol: X 2020;583:124631.
[33] Kratzert F, Klotz D, Brenner C, Schulz K, Herrnegger M. Rainfall–runoff modelling using long short-term memory (LSTM) networks. Hydrol Earth Syst Sci 2018;22(11):6005–22.
[34] Davis N, Raina G, Jagannathan K. LSTM-based anomaly detection: detection rules from extreme value theory. In: Oliveira PM, Novais P, Reis LP, editors. Proceedings, Part I of 19th EPIA Conference on Artificial Intelligence; 2019 Sep 3–6; Vila Real, Portugal. Cham: Springer Nature Switzerland; 2019. p. 572–83.
[35] Chen Z, Yu H, Geng Ya, Li Q, Zhang Y. EvaNet: an extreme value attention network for long-term air quality prediction. In: 2020 IEEE International Conference on Big Data (Big Data); 2020 Dec 10–13. Atlanta, GA, USA, Washington: Institute of Electrical and Electronics Engineers (IEEE); 2020. p. 4545–52.
[36] Ding D, Zhang M, Pan X, Yang M, He X. Modeling extreme events in time series prediction. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; 2019 Aug 4–8; Anchorage, AK, USA. New York: Association for Computing Machinery (ACM); 2019. p. 1114–22.
[37] Zhang YL, You WJ. Social vulnerability to floods: a case study of Huaihe River Basin. Nat Hazards 2014;71(3):2113–25.
[38] Liu Z, Martina MLV, Todini E. Flood forecasting using a fully distributed model: application of the TOPKAPI model to the Upper Xixian Catchment. Hydrol Earth Syst Sci 2005;9:347–64.
[39] Lv N, Liang X, Chen C, Zhou Y, Li J, Wei H, et al. A long short-term memory cyclic model with mutual information for hydrology forecasting: a case study in the Xixian Basin. Adv Water Resour 2020;141:103622.
[40] Li M, Chu R, Islam ARMT, Shen S. Characteristics of surface evapotranspiration and its response to climate and land use and land cover in the Huai River Basin of Eastern China. Environ Sci Pollut R 2021;28(1):683–99.
[41] Shi P, Ma X, Hou Y, Li Q, Zhang Z, Qu S, et al. Effects of land-use and climate change on hydrological processes in the Upstream of Huai River, China. Water Resour Manage 2013;27(5):1263–78.
[42] Kendall MG. Rank correlation methods. Oxford: Griffin; 1948.
[43] Liu M, Huang Y, Li Z, Tong B, Liu Z, Sun M, et al. The applicability of LSTM-KNN model for real-time flood forecasting in different climate zones in China. Water 2020;12(2):440.
[44] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323(6088):533–6.
[45] Emin Orhan A, Pitkow X. Skip connections eliminate singularities. 2017. arXiv:1701.09175.
[46] Wenling S, Kihyuk S, Diogo A, Honglak L. Understanding and improving convolutional neural networks via concatenated rectified linear units. In: Balcan MF, Weinberger KQ, editors. Proceedings of the 33rd International Conference on International Conference on Machine Learning; 2016 June 19–24; New York, NY, USA. Cambridge: MIT Press; 2016. p. 2217–25.
[47] Ramachandran P, Zoph B, Le QV. Searching for activation functions. 2017. arXiv:1710.05941.
[48] Zhang G, Eddy PB, Hu MY. Forecasting with artificial neural networks: the state of the art. Int J Forecast 1998;14(1):35–62.
[49] Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. 2014. arXiv:1412.3555.
[50] Burnash RJ, Ferral RL, McGuire RA. A generalized streamflow simulation system: conceptual modeling for digital computers. Sacramento: US Department of Commerce, National Weather Service, and State of California, Department of Water Resour; 1973.
[51] Chiew FHS, Peel MC, Western AW. Application and testing of the simple rainfall-runoff model SIMHYD. In: Singh VP, Frevert DK, editors. Mathematical models of small watershed hydrology and applications. Colorado: Water Resources Publications, LLC; 2002. p. 335–67.
[52] Sugawara M, Sentā KBKG. Tank model with snow component Report. Tsukuba: National Research Center for Disaster Prevention, Science and Technology Agency; 1984.
[53] Song JH, Her Y, Park J, Lee KD, Kang MS. Simulink implementation of a hydrologic model: a tank model case study. Water 2017;9(9):639.
[54] Bennett ND, Croke BFW, Guariso G, Guillaume JHA, Hamilton SH, Jakeman AJ, et al. Characterising performance of environmental models. Environ Model Softw 2013;40:1–20.
[55] Ding Y, Zhu Y, Feng J, Zhang P, Cheng Z. Interpretable spatio-temporal attention LSTM model for flood forecasting. Neurocomputing 2020;403:348–59.
[56] China Meteorological Administration (CMA). Blue paper of climate change 2019 in China. Beijing: China Meteorological Administration (CMA); 2019.
[57] Han S, Xu D, Wang S. Decreasing potential evaporation trends in China from 1956 to 2005: accelerated in regions with significant agricultural influence? Agric Meteorol 2012;154–155:44–56.