

Article

# A Novel Hybrid Extreme Learning Machine Approach Improved by K Nearest Neighbor Method and Fireworks Algorithm for Flood Forecasting in Medium and Small Watershed of Loess Region

Juanhui Ren <sup>1</sup>, Bo Ren <sup>2,3</sup>, Qiuwen Zhang <sup>2,\*</sup> and Xiuqing Zheng <sup>1,\*</sup>

<sup>1</sup> College of Water Resources Science and Engineering, Taiyuan University of Technology, Taiyuan 030024, China

<sup>2</sup> College of Hydropower and information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

<sup>3</sup> Hydrology Bureau of Shanxi Province, Taiyuan 030001, China

\* Correspondence: qwzhang@hust.edu.cn (Q.Z.); zhengxiuqing@tyut.edu.cn (X.Z.); Tel.: +86-186-3616-1777 (Q.Z.); +86-0351-6010102 (X.Z.)

Received: 6 August 2019; Accepted: 2 September 2019; Published: 5 September 2019



**Abstract:** Sudden floods in the medium and small watershed by a sudden rainstorm and locally heavy rainfall often lead to flash floods. Therefore, it is of practical and theoretical significance to explore appropriate flood forecasting model for medium and small watersheds for flood control and disaster reduction in the loess region under the condition of underlying surface changes. This paper took the Gedong basin in the loess region of western Shanxi as the research area, analyzing the underlying surface and floods characteristics. The underlying surface change was divided into three periods (HSP<sub>1</sub>, HSP<sub>2</sub>, HSP<sub>3</sub>), and the floods were divided into three grades (great, moderate, small). The paper applied K Nearest Neighbor method and Fireworks Algorithm to improve the Extreme Learning Machine model (KNN-FWA-ELM) and proposed KNN-FWA-ELM hybrid flood forecasting model, which was further applied to flood forecasting of different underlying surface conditions and flood grades. Results demonstrated that KNN-FWA-ELM model had better simulation performance and higher simulation accuracy than the ELM model for flood forecasting, and the qualified rate was 17.39% higher than the ELM model. KNN-FWA-ELM model was superior to the ELM model in three periods and the simulation performance of three flood grades, and the simulation performance of KNN-FWA-ELM model was better in HSP<sub>1</sub> stage floods and great floods.

**Keywords:** loess region; medium and small watershed; flood forecasting; k nearest neighbor method; fireworks algorithm; extreme learning machine; hybrid approach

## 1. Introduction

In recent years, with the improvement of hydrological data acquisition technology and the development of artificial intelligence computing, data-driven flood forecasting models have gained increasing attention [1–8]. The development of intelligent computing has gone through three important stages. The first stage witnessed the emergence of mathematical statistics information technology based on Bayes theory. “Black box” models represented by artificial neural network emerged in the second stage, and, in the third stage, machine learning methods were discovered. With the improvement of computer technology, machine learning has gained increasing application fields, including the field of hydrology and water resources [9–15]. In the aspect of flood forecasting, machine learning can provide methods for constructing a data-driven high-speed flood forecasting model that can reflect the nonlinear

mechanism of flood process in watersheds. By now, the typical machine learning methods mainly include support vector machine (SVM), BP neural network, artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), deep learning, and extreme learning machine (ELM) [16]. Hadi et al. [17] made a comparison of the applicability between three data-driven models-ANN, ANFIS, and SVM to flood discharge simulation in three different watersheds. The results show that, in all the watersheds, both ANN model and ANFIS model simulate floods more accurately than SVM does; ANN model and ANFIS model have similar simulation process, but the former performs better; among all the models, ANN is the best in peak discharge simulation; the simulation accuracy of the models are reduced by the factors, such as small area, steep slope, flood fluctuation, and so on. He et al. [18] compared river discharge forecasting results in the typical semiarid hilly region between three data-driven models-ANN, ANFIS, and SVM, and the results indicate that SVM model boasts better forecasting performance than ANN and ANFIS model. Taormina et al. [19] combined ELM with Binary-coded Discrete Fully Informed Particle Swarm Optimization (BFIPS) to select an optimal input data set for rainfall-runoff simulation. Lima et al. [9] evaluated the forecasting accuracy and calculation speed of the discharge forecasting of ELM and ANN trained by Gradient Descent Algorithm, and the results suggest that ELM and ANN have similar performance, and the former shows faster calculation speed than the latter except for large data sets with multiple predictor variables. Lima et al. [20] selected two small watersheds in Colombia as a study area, used online sequence extreme learning machine model to make daily runoff simulation and forecasting in the watersheds, and compared it with online sequence multiple linear regression model. The result is that online sequence extreme learning machine is easily superior in forecasting to online sequence multiple linear regression model. Wang et al. [21] combined BP neural network model with Xin'anjiang hydrological model, which was used for Anding River watershed flood forecasting. The results show that the improved flood forecasting model can effectively decrease the calibration time of BP neural network and improve flood forecasting accuracy. Kong et al. [22] proposed a method for flood forecasting by integrated extreme learning machine based on similarity matching. The method can adaptively select appropriate ELM construction integration for different test samples, thus improving flood forecasting accuracy. Liu [23] applied parallel extreme learning machine model to flood forecasting in Wei River and the Hanjiang River watersheds, rendering satisfactory forecasting results and high operational efficiency. Evidently, it is applicable to flood forecasting in the foregoing watersheds. Kan et al. [24] combined artificial neural network with K Nearest Neighbor (KNN) method, constructed a flood forecasting model based on coupled machine learning, and applied it to flood forecasting in Tunxi River watershed. The results show that the model has high forecasting accuracy and reliability. Extreme learning machine [25] is a single hidden layer feed-forward neural network model characterized by simple construction, fast training speed, strong generalization, stable learning performance, and capability of avoiding local minimum values to some extent. Therefore, the method is favored by scholars in related fields, and has been widely applied in evapotranspiration forecasting [26], wind power forecasting [27], medium and long-term rainfall forecasting [28], drought forecasting [29,30], and many other fields.

As the economy and society develop, human activities impose ever-increasing influence on the natural environment. The loess region, with inherently fragile natural environment vulnerable to the long-run inappropriate human activities, has to face serious water loss and soil erosion. Due to complex topographical and geomorphological conditions and frequent storm rainfall, the loess region has become a place frequently threatened by flash flood disaster. Therefore, the storm flood in the loess region is very typical and representative and thus has aroused extensive attention from scholars. Wang et al. [31] conducted a statistical analysis on the data of rainfall floods in the four typical small watersheds of the loess region—Shejiagou watershed, Zizhou Tuanshangou watershed, Chabagou watershed, and Liujiagou watershed—and drew the conclusion that the flood in the loess region features high, short-duration, rapidly rising, and falling peak. Zhang et al. [32] constructed lumped, semi-distributed, and fully-distributed rainfall-runoff models for Gushanchuan watershed of the loess region based on LCM (Liu Changming Model) applicable to the excess infiltration runoff generation of

the loess plateau and compared and analyzed the simulation accuracy and calculation efficiency of the models. The results show that the fully-distributed model has the highest simulation accuracy, the lumped model has the highest calculation efficiency, and the semi-distributed model also renders high calculation efficiency while ensuring high simulation accuracy. Li [33] analyzed the change trend of the storm floods in Chabagou watershed, made flood simulation in the watershed based on multi-water source Time Varying Gain Model, used the parameters before and after the change of the underlying surface to simulate and analyze the floods of different grades, and estimated the influence of the change of the underlying surface on the flood. The results indicate that as the underlying surface of the watershed changes, the peak discharge and the flood volume become attenuated, and the peak discharge is attenuated more than the flood volume. The change of the underlying surface has an obvious peak clipping effect on moderate and great floods and attenuation effect on small floods. Yang [34] made flood simulation in Beiyu River watershed based on HEC-HMS (Hydrologic Engineering Center-Hydrological Model System) model, and the results tell that vegetation restoration can effectively improve the runoff generation conditions of the watershed and reduce peak discharge, which has an important guiding significance for flood control and disaster mitigation in watersheds. Feng [35] selected 20 floods in Chabagou watershed, constructed multiple linear regression model, BP neural network model, and detrended cross-correlation analysis (DCCA) method, and constructed a combined BP neural network model based on Multiple Linear Regression-Detrended Cross Correlation Analysis (MLR-DCCA) to forecasting flood discharge. The results indicate that the combined model has better simulation effect and higher simulation accuracy than the single model.

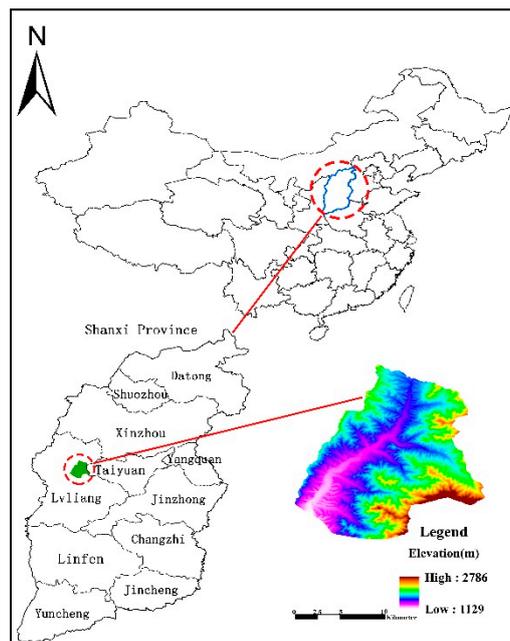
According to the current research status, scholars have achieved some research achievements in the flood forecasting model and flood simulation in medium and small watershed of the loess region, but there are still some issues that need to be clarified: (1) The research on flood forecasting in medium and small watershed in the loess region mostly concentrates on statistical methods and process-driven hydrological models, and most of the existing data-driven flood forecasting models do not consider the impact of underlying surface changes. The research results reflecting the impact of the underlying surface of medium and small watershed in the loess region on flooding and the flood forecasting are relatively scarce; (2) There are several articles in the literature on flood forecasting using a single machine learning model, but in the actual flood forecasting process, there is a lack of single machine learning methods. In the flood forecasting of a medium and small watershed in the loess region, the research on data-driven hybrid models obtained by combining different machine learning methods and achieving complementary advantages is less. Based on the extreme learning machine, this paper discussed its shortcomings in flood forecasting and used intelligent optimization algorithm and machine learning method to improve it. A hybrid flood forecasting model was proposed and applied to the flood forecasting of a medium and small watershed in the loess region. By comparing the simulated flood forecasting results of the hybrid and single model under different underlying surface conditions and flood characteristics, the applicability of the hybrid model in a medium and small watershed of the loess region was evaluated, and its advantages in flood forecasting were explored. The conclusions of the paper could provide significant guidance for the actual flood forecasting in a medium and small watershed in the loess region, which has important theoretical significance and practical value.

## 2. Study Area

### 2.1. Characteristics of Underlying Surface

#### 2.1.1. Geographical Position

Gedong basin is located in the northeast of Fangshan County of Lvliang City of Shanxi Province, with geographical coordinates of 111°03' to 111°35' east longitude and 37°38' to 38°11' north latitude, and area of about 724 km<sup>2</sup>, as shown in Figure 1. It is a typical medium and small watershed.



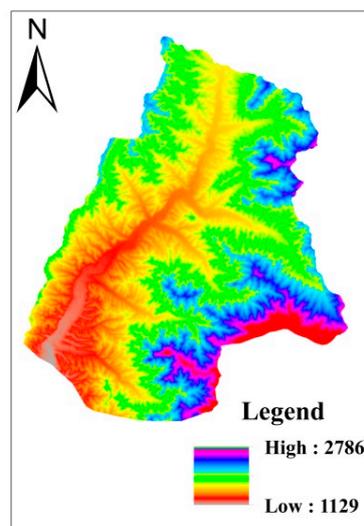
**Figure 1.** Location and area of Gedong basin.

### 2.1.2. Topographical and Geomorphological Conditions

Gedong basin is located in the loess plateau in the northwest of Shanxi Province, where the soil is loose, the storm rainfall and heavy rainfall mostly occur in summer and autumn, and the rainfall converging in a short time easily forms highly-erosive surface runoff and contributes to the thousands of gullies in the loess plateau. It features frequent rainfall storms, fast flow, and steep slopes.

According to the survey of the relevant geological disasters in the watershed, there are five geomorphological types in Gedong basin, namely denudated and eroded high and medium mountain, denudated and eroded high mountain, ridge loess hill, hilly loess hill, and sedimentary valley, of which the loess hill accounts for 40% of the total area of the watershed, so Gedong basin is a typical loess region.

The Digital elevation map, slope type, slope aspect, and slope gradient can well reflect the topographical features of the watershed, as shown in Figures 2–5, respectively.



**Figure 2.** Digital elevation map.

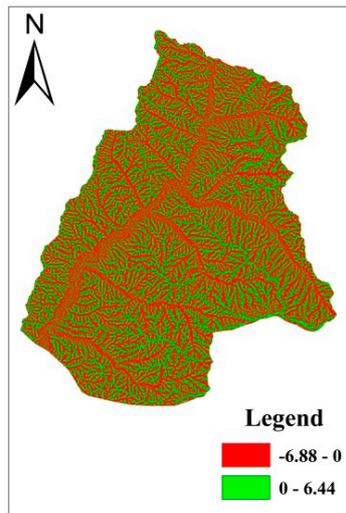


Figure 3. The slope type map.

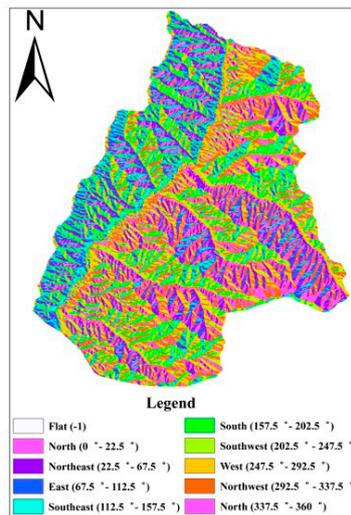


Figure 4. Slope aspect map.

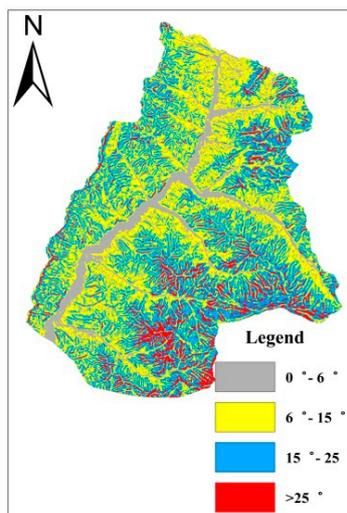


Figure 5. Slope gradient map.

From the digital elevation map, it is evident that the basin is high in the east and north and low in west and south generally. The slope type can be described and quantified by the curvature of slope surface: the curvature of liner slope and the convex slope is equal to or greater than 0, while that of concave slope and stepped slope is smaller than 0. According to slope type analysis, the slopes in the watershed are mainly concave and stepped. The basin is mainly represented by shady slope ( $0^{\circ}$ – $22.5^{\circ}$  and  $337.5^{\circ}$ – $360^{\circ}$ ,  $22.5^{\circ}$ – $67.5^{\circ}$ ,  $67.5^{\circ}$ – $112.5^{\circ}$ ,  $112.5^{\circ}$ – $157.5^{\circ}$ ) and sunny slope ( $157.5^{\circ}$ – $202.5^{\circ}$ ,  $201.5^{\circ}$ – $247.5^{\circ}$ ,  $247.5^{\circ}$ – $292.5^{\circ}$ ,  $292.5^{\circ}$ – $337.5^{\circ}$ ). In the loess region, compared with the shady slope, the sunny slope has poor water conditions, which are not suitable for vegetation growth, which makes the erosion severe and breaks the terrain. Therefore, the slope of the sunny slope is steeper than the shady slope. From the slope aspect analysis, the sunny slope accounts for about 60%, and the shady slope accounts for about 40%. It can be seen that the Gedong basin is dominated by a sunny slope, the terrain is relatively fragmented, and the slope is steep. From the slope gradient analysis, it is known that the part with slope gradient of  $0$ – $6^{\circ}$  accounts for 15% of the total area of the basin, that of  $25$ – $55^{\circ}$  accounts for 5%, and that of  $6$ – $15^{\circ}$  and  $15$ – $25^{\circ}$  accounts for 45% and 35%, respectively, which suggests the basin is dominated by the part with a slope gradient of  $6$ – $15^{\circ}$  and  $15$ – $25^{\circ}$ . In general, the basin terrain is relatively fragmented, and the slope is steep.

### 2.1.3. Soil and Land Use

The soil type in Gedong basin is simple, mainly composed of clay loam and sandy loam. The remote sensing images were used to invert the land use of the basin above Gedong Hydrological Station, respectively, in 1987, 1990, 1996, 2003, 2007, and 2012, mainly including forestland, grassland, cultivated land, construction land, and lands for other purposes (Figure 6). The area and percent of land-use types are shown in Table 1.

It can be seen in Table 1 that the land use of the basin is mainly used as forestland and then grassland and cultivated land. From 1987 to 2012, the area of forestland and grassland showed an increasing trend, while that of cultivated land constantly decreased. Compared with 1987, the percent of forestland and grassland increased by 17.73% and 6.78%, respectively, and that of cultivated land decreased by 23.64% in 2012, which indicated that the land use in the basin changed significantly and the policy of returning cultivated land to forestland and grassland made remarkable progress.

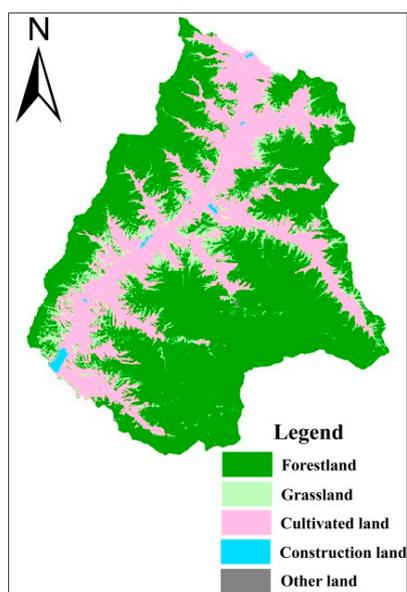


Figure 6. Land-use type map.

**Table 1.** Different land-use types and their percentage of the total area.

Land Use Type	1987		1990		1996	
	Area (km <sup>2</sup> )	Percentage (%)	Area (km <sup>2</sup> )	Percentage (%)	Area (km <sup>2</sup> )	Percentage (%)
Forestland	313.56	43.31	333.33	46.04	334.13	46.15
Grassland	124.96	17.26	130.46	18.02	142.34	19.66
Cultivated land	255.43	35.28	242.32	33.48	220.60	30.47
Construction land	3.48	0.48	5.14	0.71	5.57	0.77
Other land	26.57	3.67	12.74	1.75	21.36	2.95

Land Use type	2003		2007		2012	
	Area (km <sup>2</sup> )	Percentage (%)	Area (km <sup>2</sup> )	Percentage (%)	Area (km <sup>2</sup> )	Percentage (%)
Forestland	393.28	54.32	410.80	56.74	441.93	61.04
Grassland	158.41	21.88	159.42	22.02	174.05	24.04
Cultivated land	149.43	20.64	128.22	17.71	84.27	11.64
Construction land	13.18	1.82	18.17	2.51	23.53	3.25
Other land	9.70	1.34	7.39	1.02	0.22	0.03

The analysis of the underlying surface characteristics of the basin, relevant data, and field survey showed that the building of artificial terraces was started from the early 1960s and probably completed in 1979. Artificial terraces are strip-shaped platform-like or sectioned wave layer-like cultivated land built along the contour line, an effective way to water loss and soil erosion control in sloping cultivated land, playing a significant role in water storage, soil conversation, and yield increase. The land on the hillside is mostly built into stair-like sections (Figure 7), which is favorable to increase the infiltration of surface runoff and reduce soil erosion on slopes, thereby improving site conditions and increasing agricultural production gain. The terraced area is 76.75 km<sup>2</sup>, accounting for 10.6% of the total area of the basin.

In the basin, a small area of afforestation appeared from 1980 to 2000, after which the policy of returning cultivated land to forestland and grassland was implemented, and the large-scale afforestation has been initiated. Ren et al. [36] analyzed the flood evolution characteristics in different periods of Gedong basin and found that the characteristics of the underlying surface changed little from the 1980s to the 1990s. Therefore, according to the implementation of soil and water conservation measures in the basin, land use, and related literature, the Hydrologic Surface Period is divided into 1964–1979, 1980–2000, and 2001–2016, which are represented by HSP<sub>1</sub>, HSP<sub>2</sub>, and HSP<sub>3</sub>, respectively.

**Figure 7.** Terraces in Gedong basin.

## 2.2. Meteorological and Hydrological Characteristics

### 2.2.1. Meteorological Characteristics

The basin is characterized with warm temperate continental arid and semi-arid monsoon climate with obvious seasonal changes where the spring is low-temperature, dry, and windy, the summer is short, warm, hot, and much rainy, the autumn is cool with a pleasant climate, and the winter is long, dry, and cold. Its average annual temperature is 8.7 °C, coldest in January with average monthly

temperature of  $-9.5\text{ }^{\circ}\text{C}$  and extreme low temperature of  $-28.6\text{ }^{\circ}\text{C}$  (January 19, 1998), and hottest in July with average monthly temperature of  $21.3\text{ }^{\circ}\text{C}$  and extreme high temperature of  $35.6\text{ }^{\circ}\text{C}$  (June 22, 2005). Its average frost period is 160 days, and maximum frozen soil is 90 cm to 140 cm deep.

### 2.2.2. Hydrological Characteristics

In the Gedong basin, the average annual rainfall is 363.8 mm, of which July and August account for more than 80%. According to the historical statistical data of Gedong Hydrological Station, the average annual runoff is 109 million  $\text{m}^3$ . The hydrological data of the basin from 1964 to 2016 was sourced from Lvliang Branch of the Hydrology and Water Resources Survey Bureau of Shanxi Province. There are six rainfall gauging stations (i.e., Kai-Fu station, Ma Fang station, Ji Cui station, Wen Jiazhuang station, Shang Yangwan station, and Guo Jiazhuang station) and one runoff gauging station (Ge Dong station) distributed in the basin, as shown in Figure 8.



Figure 8. Stations in Gedong basin.

### 2.2.3. Flood Frequency Distribution Characteristics

Flood recurrence interval  $N$  (years) is the reciprocal of flood frequency, that is,  $N = 1/P$ , and flood frequency refers to the cumulative frequency of peak discharge. Flood recurrence interval or flood frequency can scientifically reflect the probability of flood occurrence and eliminate an influence factor-area of the watershed, which can hereby be used as a unified index for flood grading in watersheds. According to the Standard for Flood Control (GB50201-2014) enacted by the Ministry of Water Resources and the Ministry of Housing and Urban-Rural Development of the People’s Republic of China, a flood can be classified into five grades by flood recurrence interval as an index. The classification results are shown in Table 2.

Table 2. Classification of flood grade.

Flood Grade	Flood Recurrence Interval (Year)	Flood Frequency (%)	Flood Type
1	<5	>20	small
2	5–10	10–20	moderate
3	10–50	2–10	great
4	50–100	1–2	extraordinary
5	>100	<1	abnormal

Based on the studies by Ren et al. [36], the observed data was supplemented and improved. According to the above-mentioned national standard for flood control and the latest standard of flood frequency classification based on hydrological analysis of the Hydrology and Water Resources Survey Bureau of Shanxi Province, the frequency distribution curve of 46 floods in Gedong basin from 1964 to 2016 was revised. The results are shown in Figure 9.

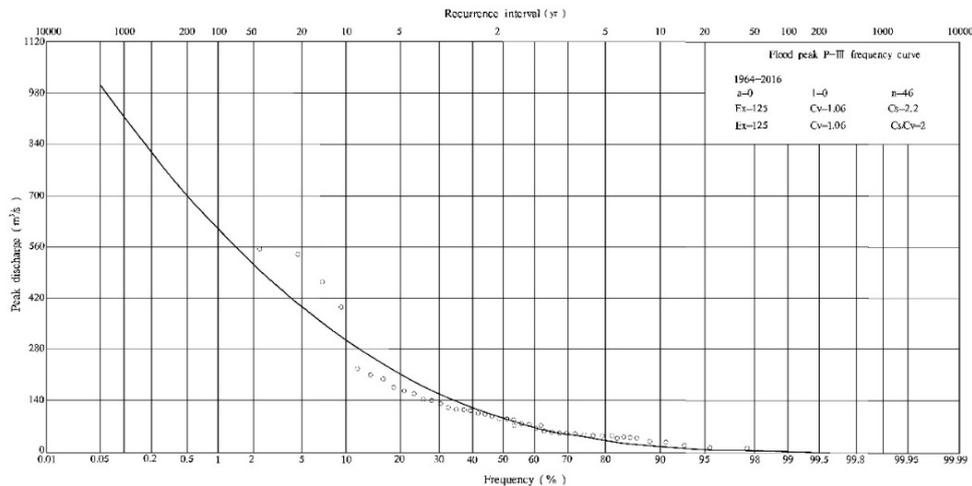


Figure 9. Frequency distribution curve.

It is known from the figure that there were 0 flood in the recurrence interval of more than 100 years, 0 flood in the occurrence interval of 50 to 100 years, four floods in the occurrence interval of 10 to 50 years, four floods in the occurrence interval of 5 to 10 years, and 38 floods in the occurrence interval of less than 5 years. According to the historical floods of Gedong basin and based on Table 2 and Figure 9, the 46 floods were classified into small, moderate, and great floods, of which small floods had peak discharge of less than 180 m<sup>3</sup>/s, moderate floods 180–350 m<sup>3</sup>/s, and great floods more than 350 m<sup>3</sup>/s.

### 3. Materials and Methods

#### 3.1. Principle of ELM

ELM, proposed by Professor Huang from the Nanyang Technological University of Singapore in 2004 [25], is a single hidden layer feed-forward neural network model. Compared with traditional neural network models, ELM is a nonlinear neural network model capable of randomly obtaining the input weight and offset and then calculating the output weight, which means that the output weight changes as the input weight and offset change [25,37–39]. This is the reason that ELM is faster in learning and better in generalization than any traditional neural network. The structure of the ELM is shown in Figure 10.

ELM establishment assumes there are  $N$  arbitrary samples  $(X_i, t_i)$ , of which  $X_i = [x_{i1}, x_{i2}, \dots, x_{in}] \in R^n$ ,  $t_i = [t_{i1}, t_{i2}, \dots, t_{in}]^T \in R^n$ , then an ELM model containing  $M$  hidden layer nodes can be expressed as:

$$\sum_{i=1}^M \beta_i g(\omega_i \cdot X_i + \varphi_i) = O_j \quad j = 1, 2, \dots, N \tag{1}$$

where  $\beta_j$  is the output weight;  $g(x)$  is the activation function of ELM;  $\omega_i = [\omega_{i1}, \omega_{i2}, \dots, \omega_{in}]^T$  is the input weight;  $\omega_i \cdot X_i$  is the inner product of  $\omega_i$  and  $X_i$ ;  $\varphi_i$  is the offset of the  $i$ -th hidden layer unit;  $O_j$  is the output value of ELM model.

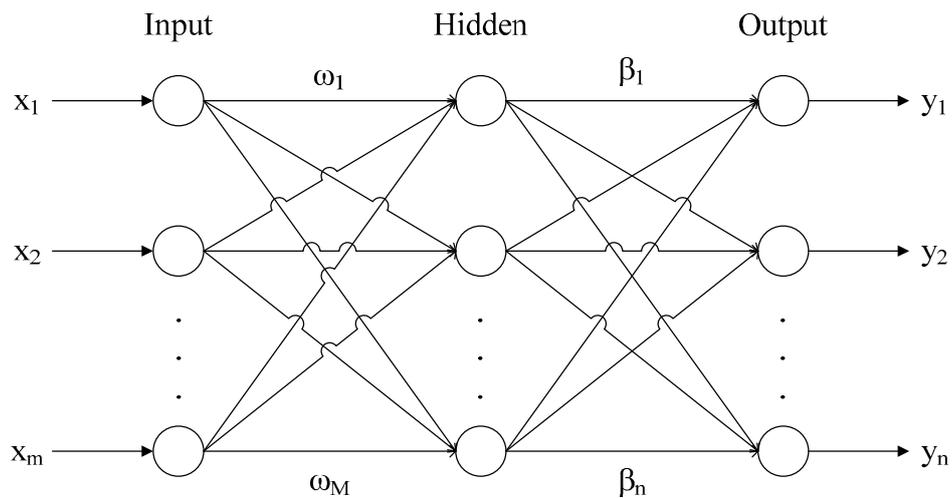


Figure 10. Structure of ELM.

The ultimate goal of the ELM model is to minimize the error between the output value  $O_j$  and the measured value  $t_j$ , which is expressed by the formula below:

$$\sum_{j=1}^N \|O_j - t_j\| = 0 \tag{2}$$

That is, find  $\beta_i$ ,  $\omega_i$ , and  $\varphi_i$  to make the following equation true:

$$\sum_{i=1}^M \beta_i g(\omega_i \cdot X_i + \varphi_i) = t_j \quad j = 1, 2, \dots, N \tag{3}$$

The above formula can be similar to a single hidden layer neural network, expressed as the matrix below:

$$H\beta = T \tag{4}$$

where  $H$  is the output of the single hidden layer node;  $\beta$  is the output weight;  $T$  is the ideal output value.

The output matrix  $H$  of the single hidden layer node is defined as:

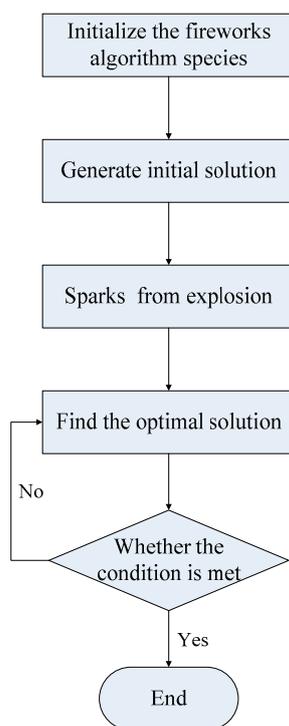
$$H = \begin{bmatrix} g(\omega_1 x_1 + \varphi_1) & \cdots & g(\omega_M x_1 + \varphi_M) \\ \vdots & \ddots & \vdots \\ g(\omega_1 x_N + \varphi_1) & \cdots & g(\omega_M x_N + \varphi_M) \end{bmatrix} \tag{5}$$

It can be seen that the input weight  $\omega_i$  and offset  $\varphi_i$  of the ELM model are generated randomly. Once the two values are determined, the output matrix  $H$  of the hidden layer can be determined, and ELM model can be converted into a linear system like  $H\beta = T$  to be solved, that is, the output weight  $\beta = H^{-1}T$ .

### 3.2. Principle of Fireworks Algorithm

Fireworks Algorithm (FWA), proposed by Tan and Zhu in 2010 and inspired by the natural phenomenon that fireworks explosion produces sparks [40], is a new type of swarm intelligence optimization algorithm with advantages in simulation accuracy and convergence performance. Similar to other optimization algorithms, it aims to find a solution to the optimization problems with constraints, but it provides a new way to find the solution, that is searching the solution according to the random explosion process of fireworks in a local space [41]. FWA mainly consists of explosion operator,

mutation operator, mapping rule, and selection strategy. Its basic principle is: firstly, assume that the number of fireworks at the beginning is  $N$ , evaluate the merit of each of the  $N$  fireworks, and determine the number of sparks generated by the explosion of each firework and the explosion magnitude of each firework. Secondly, introduce a mutation operator to generate variant sparks to diversify the algorithm. Thirdly, select a certain number of fireworks (including explosion sparks and Gaussian sparks) as the next generation of fireworks to pass down the excellent information. Finally, select the location of the sparks after the explosion of the fireworks. The flow of the fireworks algorithm is shown in Figure 11.



**Figure 11.** Chart of fireworks algorithm.

### 3.3. Principle of $K$ Nearest Neighbor Method

$K$  nearest neighbor method is a nonparametric statistical method mainly used for sample classification or regression [42]. Its basic principle is assuming that most samples in a feature space fall into a category, then  $k$  values of most similar samples in the feature space also belong to the category. The specific algorithm has two steps: (1) for a specific sample set, make measurement by a certain distance, and determine the value of  $k$ ; (2) in the determined KNN method, make classification according to the category which most samples belong to. The flow of the  $k$  nearest neighbor method is shown in Figure 12.

### 3.4. The Construction of the KNN-FWA-ELM Model

The application of the ELM model to flood forecasting in medium and small watershed of the loess region has the following shortcomings: (1) the ELM model randomly generates input weight and offset, and then calculates output weight, but the randomly generated input weight and offset are not necessarily optimal, which may result in a large error in model simulation; (2) the ELM model forecasts the discharge at the corresponding late time based on the rainfall at each rainfall gauging station at the current time and the rainfall at each rainfall gauging station and the observed discharge at a certain early time, which means the ELM is incapable of continuous flood forecasting during the absence of the observed data.

To make up for the shortcomings of the ELM model in the flood forecasting of a medium and small watershed in the loess region and further improve the forecasting ability of ELM model, this paper

improved the ELM model and proposed and constructed a data-driven hybrid model. In the hybrid model, the fireworks algorithm was used to optimize the input weight and offset, and the k nearest neighbor method was applied to predict the error of the discharge at the runoff gauging station and to achieve continuous flood forecasting when the observed data was missing and to correct the simulation results when the observed data was complete, improving the simulation and prediction accuracy.

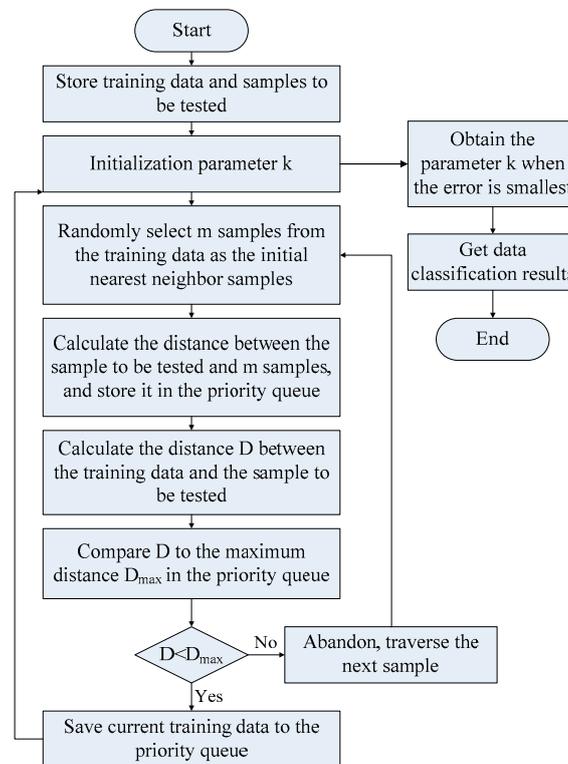


Figure 12. Flow chart of k nearest neighbor.

### 3.4.1. Model Parameters Setting

Based on the observed data of 46 floods in the Gedong basin from 1964 to 2016, the KNN-FWA-ELM and ELM models were employed to simulate the floods of 14, 18, and 14 in HSP<sub>1</sub>, HSP<sub>2</sub>, and HSP<sub>3</sub>, respectively. To obtain the desired simulation results, 10, 13, and 10 floods were randomly selected as the training set, and the remaining 4, 5, and 4 floods were used as the testing set in three periods. According to the observed floods data, the time interval was set to 5 min. The number of iterations was set to 1000 by trials and tests.

In KNN-FWA-ELM model, the activation function needed to be selected. The activation function was based on the actual flood process in the basin and was determined empirically. As the Sigmoid function (Equation (6)) can handle both large and small signals in the network, the Sigmoid function was selected as the activation function of KNN-FWA-ELM model.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

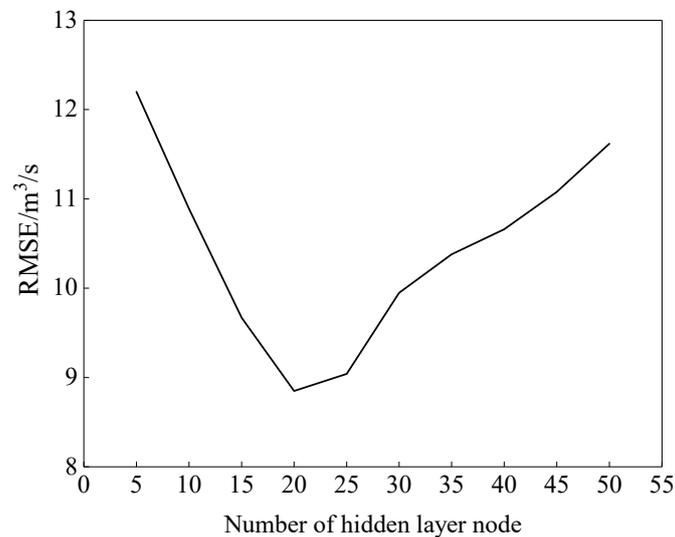
The determination of the number of hidden layer nodes directly affects the mapping ability of the network. If the number of nodes is too small, the information obtained by the network from the sample is insufficient, making it difficult to reflect the sample law of the training set. Conversely, if the number of nodes is too large, the irregular content of the training set may be learned, resulting in over-fitness. It can be known from the literature [43] that the number of hidden layer nodes can be determined by Equation (7).

$$L = \text{round}[\sqrt{k + m} + \text{rand}(1 - 10)] \quad (7)$$

where  $L$  is the number of hidden layer nodes;  $k$  is the number of model input factors;  $m$  is the number of model output factors. Specifically, the number of hidden layer nodes is determined by the following steps:

- (1) Set fewer hidden layer nodes;
- (2) Train and test the sample set;
- (3) Gradually increase the number of hidden layer nodes and use the same sample set for training and testing;
- (4) Compare the training and testing results of different hidden layer nodes, and obtain the number of hidden layer nodes when the error is the smallest.

With the above procedure, the number of the KNN-FWA-ELM model was obtained according to the Root Mean Squared Error (RMSE) curve. The root mean squared error curve is shown in Figure 13.



**Figure 13.** Relationship between RMSE (Root Mean Squared Error) and several hidden layer nodes.

Since the KNN-FWA-ELM model randomly generated input weight and offset, the result of each run changed after determining the number of hidden layer node. To obtain a desired and stable output, multiple runs were required to average. After many trials and tests, it was determined that KNN-FWA-ELM model needed to be run 30 times.

#### 3.4.2. Input and Output of the Model

In flood forecasting, the rainfall and the characteristics of the underlying surface are key influencing factors for discharge simulation and forecasting at the outlet of the basin, of which the rainfall is the main driving force for runoff generation in the basin, mainly including rainfall at each rainfall gauging station at time  $t$  and that at a certain early time; the characteristics of the underlying surface of the watershed is another key factor influencing the runoff generation because the data-driven model inherently cannot take into account the characteristics of the underlying surface of the basin; early-stage discharge, which is the discharge at the outlet at a certain early time, is employed to represent the conditions of the underlying surface of the basin. In short, the foregoing influence factors were used as the input data of the model and the discharge at the outlet at a certain late time corresponding to time  $t$  as the output data of the model to drive the model for flood simulation and forecasting.

The period by which the rainfall and discharge are shifted to a certain early time as the input data of the model and the period by which the discharge is shifted to a certain late time as the output

data of the model ( $\Delta T$ ) have a significant influence on model stability and forecasting accuracy. A too short period will render limited information acquired by the model and low simulation accuracy of the model; while too long period will lead to redundant information and the model over-fitting. In view of this, the partial mutual information (PMI) method [44–46] is applied, and the input factors for flood forecasting are screened to determine  $\Delta T$ . The specific screening process is as follows:

- (1) Based on the analysis of the characteristics of the floods in HSP<sub>1</sub>, HSP<sub>2</sub>, and HSP<sub>3</sub>, obtain the average lag time of the floods in three periods, which is 1.07 h, 1.60 h, and 2.37 h, respectively, and ensure the maximum value of the  $\Delta T$  is slightly larger than the average lag time of the floods in three periods, respectively.
- (2) Use the rainfall at each rainfall station at time  $t$ , at a certain early time, and the discharge at the outlet at time  $t$  as the input variables of the model, and ensure the maximum value of the period by which the rainfall shifted to a certain early time is the maximum value determined in the previous step.
- (3) Set the foregoing input variables as the input data and the discharge at the outlet at a certain late time corresponding to time  $t$  as the output data, make the iterative calculation by PMI method, and obtain the screened input variables, of which the difference between the corresponding time and time  $t$  is  $\Delta T$ .

In the paper,  $\Delta T$  was 0.92 h (55 min), 1.50 h (90 min), and 2.25 h (135 min) in three periods, respectively.

#### 3.4.3. Data Normalization

Thirteen input factors and 1 output factor of the model were obtained according to the procedures described in Section 3.4.2. Since different input and output factors have different dimensions and units, which affects the simulation and results analysis, it is necessary to standardize the data to eliminate the adverse influence. The data normalization adopted in the paper was min-max standardization by which each original data was linearly converted and the value of it was mapped to [0,1]. The conversion formula is as follows:

$$\hat{X} = \frac{X - \min}{\max - \min} \quad (8)$$

where  $X$  is the original data;  $\hat{X}$  is the normalized data; min is the minimum value in the original sample data; max is the maximum value in the original sample data.

#### 3.4.4. Model Construction

Based on the rainfall and discharge data, which are processed for normalization, the ELM model was applied to forecast the discharge at the outlet of the basin. The model is constructed as follows:

$$Q_s(t + \Delta t) = M_{ELM} \begin{bmatrix} P(t), P(t-1), \dots, P(t-\Delta t) \\ Q(t-1), Q(t-2), \dots, Q(t-t) \end{bmatrix} \quad (9)$$

where  $Q_s(t + \Delta t)$  is the simulated discharge at time  $t + \Delta t$ ;  $P$  is the rainfall at time  $t$  and at a certain early time;  $Q$  is the discharge corresponding to the rainfall at the early time.

KNN method is used for discharge error forecasting at the outlet of the basin. Firstly, select the error value of the forecasted discharge as a feature vector, and based on the feature vector at a certain time, select  $k$  samples most similar to the feature vector from the historical sample set, where the similar samples are determined by the minimum Euclidean distance between the feature vectors. Then, make inverse distance weighting (IDW) of the discharge errors of the selected similar samples and obtain the discharge error forecasting value at the outlet of the basin. The modeling method is as follows:

$$E_s(t + \Delta t) = M_{KNN} \begin{bmatrix} Q_s(t + \Delta t) \\ P(t), P(t-1), \dots, P(t - \Delta t) \\ Q(t-1), Q(t-2), \dots, Q(t - \Delta t) \end{bmatrix} \quad (10)$$

where  $E_s(t + \Delta t)$  is the simulated discharge error at time  $t + \Delta t$ .

$$Q(t + \Delta t) = Q_s(t + \Delta t) + E_s(t + \Delta t) \quad (11)$$

where  $Q(t + t)$  is the simulated discharge at time  $t + \Delta t$  of KNN-FWA-ELM model.

In the flood forecasting of KNN-FWA-ELM model, FWA is applied to select the optimal input weight and offset for the model after multiple iterations according to explosion process of fireworks, form a new matrix  $H$  by the optimal weight and offset, and finally calculate the output weight matrix, so as to optimize the parameters and improve the forecasting accuracy and generalization performance of the model. In the initial calculation of KNN, the observed discharge data is used as the early-stage discharge of model input; as the calculation continues, the simulated discharge can be used as the early-stage discharge of model input in the subsequent calculation. By performing this loop simulation, the information of the early-stage discharge can be effectively used, and the simulated value of early-stage discharge value, instead of the observed value, can serve as the input data of the model, which is favorable to the continuous flood forecasting. The specific modeling process is shown in Figure 14.

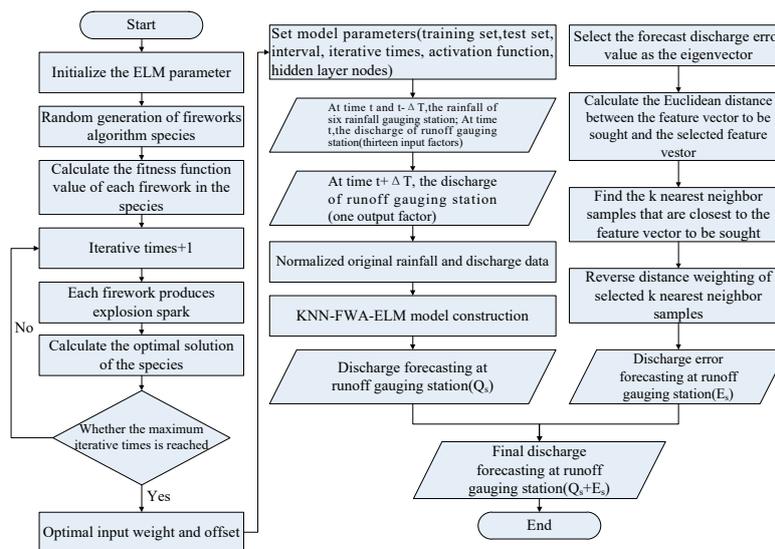


Figure 14. Establishment of the KNN-FWA-ELM model.

### 3.4.5. Evaluation Indexes for Forecasting Performance

There are many performance indices for evaluating the model’s performance [47]. Since each statistical index has its disadvantages, various indices should be used for evaluating the model performance effectively [48,49]. To evaluate the performance of the hybrid KNN-FWA-ELM model, the peak flood error ( $\Delta Q$ ), the peak current difference ( $\Delta h$ ), the coefficient of determination ( $R^2$ ), Nash–Sutcliffe efficiency coefficient ( $NS$ ), root mean squared error ( $RMSE$ ), mean squared relative error ( $MSRE$ ), and mean absolute relative error ( $MARE$ ) were used as the evaluation indicators in this study, which have been widely and commonly used for evaluating the performance of flood forecasting and hydrological simulation. The formulas are defined as follows (Equations (12)–(18)), respectively.

$$\Delta Q = \frac{Q_s - Q_o}{Q_o} \times 100\% \quad (12)$$

where  $\Delta Q$  is the peak flood error (%);  $Q_s$  is the simulated discharge ( $m^3/s$ );  $Q_o$  is the observed discharge ( $m^3/s$ ).

$$\Delta h = h_{speak} - h_{opeak} \tag{13}$$

where  $\Delta h$  is the peak current difference (h);  $h_{speak}$  is the simulated peak current time (h);  $h_{opeak}$  is the observed peak current time (h).

$$R^2 = \frac{(\sum_{t=1}^T (Q_s^t - \overline{Q_s})(Q_o^t - \overline{Q_o}))^2}{\sum_{t=1}^T (Q_s^t - \overline{Q_s})^2 \sum_{t=1}^T (Q_o^t - \overline{Q_o})^2} \tag{14}$$

$$NS = 1 - \frac{\sum_{t=1}^T (Q_s^t - Q_o^t)^2}{\sum_{t=1}^T (Q_o^t - \overline{Q_o})^2} \tag{15}$$

where  $Q_s^t$  is the simulated discharge at time t ( $m^3/s$ );  $Q_o^t$  is the observed discharge at time t ( $m^3/s$ );  $\overline{Q_s}$  is the average of simulated discharge ( $m^3/s$ );  $\overline{Q_o}$  is the average of observed discharge ( $m^3/s$ ).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Q_{s,i} - Q_{o,i})^2}{n}} \tag{16}$$

$$MSRE = \frac{1}{n} \sum_{i=1}^n \frac{(Q_{s,i} - Q_{o,i})^2}{Q_{o,i}^2} \tag{17}$$

$$MARE = \frac{1}{n} \sum_{i=1}^n \frac{|Q_{s,i} - Q_{o,i}|}{Q_{o,i}} \tag{18}$$

where  $Q_{s,i}$  is the *i*th simulated discharge ( $m^3/s$ );  $Q_{o,i}$  is the *i*th observed discharge ( $m^3/s$ ).

### 4. Results and Discussion

#### 4.1. The Flood Forecasting of ELM Model

A total of 46 floods in 1964–2016 were selected for simulation, and the simulation results by the ELM model in HSP<sub>1</sub>, HSP<sub>2</sub>, and HSP<sub>3</sub> were obtained, respectively. The simulation results are summarized in Table 3.

**Table 3.** Evaluation of flood simulation results by ELM (extreme learning machine) model.

Periods	Data Set	Flood Events	$\Delta Q/\%$	$\Delta h/h$	NS	$R^2$	RMSE/ $m^3/s$	MSRE	MARE	Qualified or Not
HSP <sub>1</sub>	Training	19640716	-19.21	-0.33	0.83	0.91	2.86	0.05	0.17	Qualified
	Training	19670822	-34.43	-0.42	0.60	0.63	85.43	0.48	0.65	Not qualified
	Training	19680727	-7.15	-0.17	0.78	0.78	5.19	0.67	0.56	Qualified
	Training	19690728	-17.24	-0.25	0.79	0.79	4.39	0.04	0.16	Qualified
	Training	19700809	-19.18	-0.08	0.86	0.89	36.59	3.15	0.92	Qualified
	Training	19710815	-10.27	-0.50	0.81	0.82	7.30	1.16	0.57	Qualified
	Training	19720719	-18.69	0	0.82	0.83	15.19	6.09	1.04	Qualified
	Training	19730716	-32.93	0.25	0.64	0.75	11.92	4.64	1.92	Not qualified
	Training	19760728	-19.55	-1.50	0.72	0.83	9.08	1.01	0.46	Qualified
	Training	19780717	-29.89	0	0.69	0.85	19.51	17.63	2.66	Not qualified
	Testing	19650707	-5.93	-0.17	0.70	0.94	3.67	2.19	1.09	Qualified
	Testing	19660816	-10.89	-0.75	0.83	0.88	19.88	2.02	0.86	Qualified
	Testing	19770705	-7.58	-0.25	0.81	0.86	10.26	0.37	0.53	Qualified
	Testing	19790723	-16.01	-0.67	0.83	0.85	4.37	0.30	0.39	Qualified

Table 3. Cont.

Periods	Data Set	Flood Events	$\Delta Q/\%$	$\Delta h/h$	NS	$R^2$	RMSE/m <sup>3</sup> /s	MSRE	MARE	Qualified or Not
HSP <sub>2</sub>	Training	19800820	-34.75	-0.17	0.66	0.66	0.62	0.15	0.29	Not qualified
	Training	19840701	-23.62	0.42	0.67	0.69	5.69	0.70	0.61	Not qualified
	Training	19850805	-14.02	-0.42	0.76	0.79	10.40	36.65	1.43	Qualified
	Training	19860729	-3.33	-0.50	0.68	0.80	22.72	6.55	4.27	Not qualified
	Training	19880723	-17.22	0	0.66	0.74	53.29	1.48	0.89	Not qualified
	Training	19910915	-15.80	-0.50	0.80	0.84	15.62	131.22	4.93	Qualified
	Training	19920802	-51.04	0.25	0.65	0.70	3.82	2.92	1.49	Not qualified
	Training	19940715	-12.87	-0.33	0.82	0.84	2.54	0.26	0.42	Qualified
	Training	19950801	-10.67	-0.67	0.84	0.87	8.02	2.43	1.16	Qualified
	Training	19960809	-19.96	-0.42	0.65	0.67	11.11	2.13	0.74	Not qualified
	Training	19970729	-27.15	0.33	0.64	0.72	6.92	1.38	1.06	Not qualified
	Training	19990721	-12.76	-0.17	0.71	0.74	9.20	0.61	0.61	Qualified
	Training	20000704	-28.59	-0.25	0.76	0.82	12.05	7.41	2.39	Not qualified
	Testing	19810620	-17.70	-1.75	0.72	0.74	6.60	0.28	0.44	Qualified
	Testing	19820815	-16.70	-0.42	0.70	0.70	7.05	1.89	0.69	Qualified
	Testing	19870630	-10.68	-0.25	0.72	0.81	7.92	8.11	2.14	Qualified
Testing	19890722	-8.84	-0.25	0.81	0.91	4.41	0.18	0.30	Qualified	
Testing	19980713	-23.91	0	0.69	0.71	20.36	0.81	0.69	Not qualified	
HSP <sub>3</sub>	Training	20010826	-14.62	0.17	0.32	0.36	6.02	2.84	1.61	Not qualified
	Training	20030607	-18.50	-0.58	0.79	0.81	3.81	1.68	0.67	Qualified
	Training	20040812	-18.34	-0.25	0.72	0.87	3.60	1.79	0.71	Qualified
	Training	20050812	-14.09	-0.25	0.80	0.80	6.00	2.28	1.07	Qualified
	Training	20071006	-9.67	-0.50	0.79	0.86	4.74	1.14	0.32	Qualified
	Training	20090907	-14.99	-0.42	0.73	0.81	5.22	1.04	0.15	Qualified
	Training	20110814	-22.72	0	0.05	0.51	0.86	0.08	0.22	Not qualified
	Training	20130811	-11.30	0	0.74	0.85	6.57	1.23	0.66	Qualified
	Training	20140709	-17.32	2.42	0.36	0.71	2.66	0.13	0.34	Not qualified
	Training	20160815	-11.87	-0.07	0.78	0.79	9.23	1.21	0.69	Qualified
	Testing	20020627	-24.65	-0.17	0.87	0.90	100.38	0.49	0.52	Not qualified
	Testing	20060814	-6.28	-0.25	0.79	0.84	4.25	2.23	1.11	Qualified
	Testing	20120731	-6.75	-0.75	0.72	0.75	2.60	2.08	0.18	Qualified
Testing	20150802	-9.13	0	0.72	0.73	11.80	1.18	0.72	Qualified	

$\Delta Q$  is the peak flood error;  $\Delta h$  is the peak current difference; NS is the Nash–Sutcliffe efficiency coefficient;  $R^2$  is the coefficient of determination; RMSE is the root mean squared error; MSRE is the mean squared relative error; MARE is the mean absolute relative error.

The Table 3 shows that, among the 46 floods simulated by ELM model, the absolute of  $\Delta Q$  of 35 floods was within 20%, rendering the overall qualified rate as 76.09%, with the variation range of [3.33, 19.55]; the absolute of  $\Delta h$  of 46 floods was within 3h, rendering the overall qualified rate as 100%, with the variation range of [0, 2.42]; the NS of 32 floods was above 0.7, rendering the overall qualified rate as 69.57%, with the maximum value of 0.87; the  $R^2$  of 40 floods was above 0.7, rendering the overall qualified rate as 86.96%, with the maximum value of 0.94. The variation range of RMSE of all the floods was [0.62, 100.38]; the variation range of MSRE of all the floods was [0.04, 131.22]; the variation range of MARE of all the floods was [0.15, 4.93]. When  $\Delta Q$ ,  $\Delta h$ , NS,  $R^2$ , RMSE, MSRE, and MARE were all put into consideration, a total of 30 simulated floods were acceptable, rendering the overall qualified rate as 65.22%.

In HSP<sub>1</sub>, HSP<sub>2</sub>, and HSP<sub>3</sub>, the simulation qualified rate of the floods was 78.57%, 50%, and 71.43%, respectively, and the average  $\Delta Q$  of the acceptable simulated floods was -13.79%, -13.34%, and -12.08%, respectively, indicating that  $\Delta Q$  of the simulated floods gradually decreased from 1964 to 2016. The average  $\Delta h$  of the acceptable simulated floods was -0.42 h, -0.53 h, and -0.32 h, respectively, which meant that  $\Delta h$  of the simulated floods in three periods was all negative, and the absolute of  $\Delta h$  of the simulated floods in HSP<sub>3</sub> was the smallest. The average NS of the acceptable simulated floods was 0.80, 0.76, and 0.76, respectively, suggesting that the simulation in HSP<sub>1</sub> was the most reliable. The average  $R^2$  of the acceptable simulated floods was 0.85, 0.80, and 0.81, respectively, showing that the linear correlation between the simulated values and the observed values in HSP<sub>2</sub> was slightly lower than that in the other two periods. The average RMSE of the acceptable simulated

floods was 10.80, 7.97, and 5.79, respectively, showing that the deviation between simulated values and the observed values gradually decreased and the simulation got better and better in three periods from 1964 to 2016. The average MSRE of the acceptable simulated floods was 1.55, 20.18, and 1.59. The average MARE of the acceptable simulated floods was 0.61, 1.35, and 0.63. According to the Standard for Hydrological Information and Hydrological Forecasting (GB/T 22482-2008), this simulation reached class C, indicating that the ELM model was applicable to flood forecasting in medium and small watersheds of the loess region. Overall, the simulation performance of the ELM model on floods in three periods was best in HSP<sub>1</sub>, followed by HSP<sub>3</sub> and HSP<sub>2</sub>.

4.2. The Flood Forecasting of KNN-FWA-ELM Model

A total of 46 floods in 1964–2016 were selected for simulation, and the simulation results in HSP<sub>1</sub>, HSP<sub>2</sub>, and HSP<sub>3</sub>, respectively, were obtained by KNN-FWA-ELM model. The simulation results are shown in Table 4.

Table 4. Evaluation of flood simulation results by KNN-FWA-ELM model.

Periods	Data Set	Flood Events	$\Delta Q/\%$	$\Delta h/h$	NS	R <sup>2</sup>	RMSE/m <sup>3</sup> /s	MSRE	MARE	Qualified or Not
HSP <sub>1</sub>	Training	19640716	-13.67	-0.33	0.91	0.91	2.04	0.02	0.09	Qualified
	Training	19660816	-2.27	-0.42	0.86	0.90	18.06	1.43	0.61	Qualified
	Training	19670822	-8.58	-0.50	0.89	0.95	43.18	0.31	0.49	Qualified
	Training	19680727	-6.25	0	0.84	0.84	4.35	0.40	0.39	Qualified
	Training	19690728	-13.33	-0.42	0.84	0.85	3.79	0.04	0.12	Qualified
	Training	19710815	-8.26	0	0.85	0.85	6.54	0.68	0.30	Qualified
	Training	19720719	-15.87	0	0.86	0.86	13.59	4.36	0.73	Qualified
	Training	19730716	-20.55	0.17	0.69	0.81	10.95	4.59	1.90	Not qualified
	Training	19760728	-14.96	-1.58	0.85	0.89	6.60	0.81	0.35	Qualified
	Training	19790723	-11.15	-0.58	0.85	0.88	4.05	0.25	0.28	Qualified
	Testing	19650707	-2.70	-0.17	0.87	0.94	2.40	0.74	0.74	Qualified
	Testing	19700809	-2.09	-0.08	0.88	0.91	32.72	2.01	0.75	Qualified
	Testing	19770705	-3.42	0	0.86	0.89	8.65	0.31	0.40	Qualified
Testing	19780717	-8.24	-0.50	0.82	0.93	14.88	8.78	1.35	Qualified	
HSP <sub>2</sub>	Training	19800820	-9.43	-0.17	0.70	0.71	0.58	0.13	0.26	Qualified
	Training	19820815	-9.07	-0.50	0.74	0.75	6.54	1.88	0.65	Qualified
	Training	19840701	-21.65	0.33	0.69	0.70	5.55	0.65	0.60	Not qualified
	Training	19850805	-11.33	-0.33	0.83	0.84	8.83	33.29	1.33	Qualified
	Training	19860729	-2.91	-0.33	0.79	0.90	18.20	5.70	1.05	Qualified
	Training	19880723	-1.03	-0.08	0.84	0.87	36.44	0.68	0.49	Qualified
	Training	19910915	-13.32	-0.25	0.85	0.90	13.23	118.65	4.41	Qualified
	Training	19920802	-34.51	0	0.67	0.79	3.76	1.48	0.99	Not qualified
	Training	19940715	-3.90	-0.25	0.87	0.87	2.22	0.24	0.36	Qualified
	Training	19950801	-4.13	-0.33	0.87	0.88	7.08	2.00	1.06	Qualified
	Training	19960809	-9.40	-0.33	0.70	0.70	10.36	1.85	0.68	Qualified
	Training	19970729	-27.05	0.08	0.68	0.80	6.50	1.36	1.03	Not qualified
	Training	19990721	-1.95	0	0.76	0.79	8.32	0.58	0.59	Qualified
	Testing	19810620	-9.36	-1.33	0.81	0.82	5.36	0.24	0.39	Qualified
	Testing	19870630	-1.01	0	0.82	0.85	6.30	8.03	2.12	Qualified
Testing	19890722	-1.50	-0.33	0.88	0.91	3.52	0.09	0.23	Qualified	
Testing	19980713	-6.13	-0.08	0.81	0.83	15.93	0.69	0.59	Qualified	
Testing	20000704	-21.09	-0.17	0.83	0.85	10.18	7.38	2.35	Not qualified	
HSP <sub>3</sub>	Training	20020627	-20.07	-0.08	0.71	0.75	73.07	0.47	0.51	Not qualified
	Training	20030607	-4.21	-0.08	0.85	0.85	3.38	1.52	0.51	Qualified
	Training	20040812	-8.15	0	0.86	0.88	2.57	1.67	0.68	Qualified
	Training	20050812	-3.88	-0.17	0.84	0.87	5.27	1.98	0.90	Qualified
	Training	20060814	-2.33	-0.25	0.85	0.85	3.69	1.95	0.91	Qualified
	Training	20071006	-3.75	-0.25	0.86	0.88	4.04	1.13	0.26	Qualified
	Training	20090907	-5.64	-0.25	0.78	0.85	4.72	0.96	0.13	Qualified
	Training	20110814	-21.11	-0.08	0.25	0.68	0.76	0.07	0.17	Not qualified
	Training	20140709	-12.44	4	0.45	0.73	2.46	0.11	0.31	Not qualified
	Training	20150802	-7.69	0.08	0.80	0.81	9.97	0.96	0.68	Qualified
	Testing	20010826	-4.99	0.08	0.73	0.82	3.88	2.57	1.30	Qualified
	Testing	20120731	-4.06	-0.67	0.85	0.86	1.91	2.06	0.18	Qualified
	Testing	20130811	-5.86	0	0.82	0.91	5.49	1.18	0.50	Qualified
Testing	20160815	-5.34	-0.17	0.86	0.86	7.47	1.15	0.53	Qualified	

The Table 4 shows that, among the 46 floods simulated by KNN-FWA-ELM model, the absolute of  $\Delta Q$  of 39 floods was within 20%, with the overall qualified rate 84.78% and the variation range of [1.01, 15.87]; the absolute of  $\Delta h$  of 45 floods was within 3 h, with the overall qualified rate 97.83% and the variation range of [0, 1.58]; the  $NS$  of 40 floods was above 0.7, rendering the overall qualified rate as 86.96%, with the maximum value of 0.91; the  $R^2$  of 44 floods was above 0.7, rendering the overall qualified rate as 95.65%, with the maximum value of 0.95. The variation range of  $RMSE$  of all the floods was [0.58, 43.18]; the variation range of  $MSRE$  of all the floods was [0.02, 118.65]; the variation range of  $MARE$  of all the floods was [0.09, 4.41]. When  $\Delta Q$ ,  $\Delta h$ ,  $NS$ ,  $R^2$ ,  $RMSE$ ,  $MSRE$ , and  $MARE$  were all put into consideration, a total of 38 floods were acceptable, rendering the overall qualified rate as 82.61%.

In HSP<sub>1</sub>, HSP<sub>2</sub>, and HSP<sub>3</sub>, the simulation qualified rate of the floods was 92.86%, 77.78%, and 78.57%, respectively, and the average  $\Delta Q$  of the acceptable simulated floods was  $-8.52\%$ ,  $-6.03\%$ , and  $-5.08\%$ , respectively, indicating that  $\Delta Q$  of the simulated floods gradually decreased from 1964 to 2016. The average  $\Delta h$  of the acceptable simulated floods was  $-0.35$  h,  $-0.31$  h, and  $-0.15$  h, respectively, indicating that the average  $\Delta h$  of the simulated floods in three periods was all negative, and the absolute of  $\Delta h$  of the simulated floods was the smallest in HSP<sub>3</sub>. The average  $NS$  of the acceptable simulated floods was 0.86, 0.81, and 0.83, respectively, suggesting that the simulation was the most reliable in HSP<sub>1</sub>. The average  $R^2$  of the acceptable simulated floods was 0.89, 0.83, and 0.86, respectively, showing that the linear correlation between the simulated values and the observed values in HSP<sub>1</sub> was higher than that in the other two periods. The average  $RMSE$  of the acceptable simulated floods was 12.37, 10.21, and 4.76, respectively, showing that the deviation between the simulated values and the observed values gradually decreased and the simulation got better and better in three periods from 1964 to 2016. The average  $MSRE$  of the acceptable simulated floods was 1.55, 12.43, and 1.56. The average  $MARE$  of the acceptable simulated floods was 0.51, 1.01, and 0.60. According to the Standard for Hydrological Information and Hydrological Forecasting (GB/T 22482–2008), this simulation reached class B, indicating that the KNN-FWA-ELM model was applicable to flood forecasting in a medium and small watershed of the loess region. In summary, the simulation performance of the KNN-FWA-ELM model on floods in three periods was best in HSP<sub>1</sub>, followed by HSP<sub>3</sub> and HSP<sub>2</sub>.

#### 4.3. Comparison and Analysis of Simulation Results Between KNN-FWA-ELM Model and ELM Model

##### 4.3.1. Comparison and Analysis of Simulation Results of All Floods

Based on the simulation qualified rate,  $\Delta Q$ ,  $\Delta h$ ,  $NS$ ,  $R^2$ ,  $RMSE$ ,  $MSRE$ , and  $MARE$ , a comprehensive comparison of the simulation results of 46 floods in 1964–2016 between KNN-FWA-ELM and ELM model was carried out, and the results are summarized in Table 5.

**Table 5.** Comparison of all flood simulation results between KNN-FWA-ELM and ELM model.

Evaluation Index	ELM Model	KNN-FWA-ELM Model
	1964–2016	1964–2016
Qualified rate/%	65.22	82.61
$\Delta Q/\%$	$-13.08$	$-6.61$
$\Delta h/h$	$-0.42$	$-0.28$
$NS$	0.77	0.83
$R^2$	0.82	0.86
$RMSE/m^3/s$	8.28	9.37
$MSRE$	7.15	5.56
$MARE$	0.84	0.72

Table 5 shows that the flood simulation qualified rate of each of the two models was higher than 60%, suggesting the simulation results were all acceptable, and the simulation qualified rate of KNN-FWA-ELM model was 17.39% higher than that of the ELM model, and the peak flood simulation value of two models were generally negative. The absolute of  $\Delta Q$  of KNN-FWA-ELM model was

6.47% less than that of the ELM model. The peak current time of the two models was early, and the absolute of  $\Delta h$  of KNN-FWA-ELM model was 0.14 h shorter than that of the ELM model. According to the values of  $NS$ , the simulation reliability of KNN-FWA-ELM model was 7.79% higher than that of the ELM model. The  $R^2$  of each model was good, but that of KNN-FWA-ELM model was 4.88% higher than the ELM model. According to the values of  $RMSE$ , the deviation between the simulated values and the observed value of the ELM model was smaller than that of KNN-FWA-ELM model. The  $MSRE$  and  $MARE$  of KNN-FWA-ELM model was 22.24% and 14.29% lower than that of the ELM model, respectively. Tables 3 and 4 show that the ELM model had a lower great flood simulation qualified rate (19660816, 19670822, 19700809, and 19880723 were the great floods, the qualified rate of the ELM model and KNN-FWA-ELM model for great floods simulation was 50% and 100%, respectively). During the great flood process, the overall fluctuation was large, and the flood peak discharge was high. There would be a large deviation between the simulation and the observed value at a certain time, and the  $RMSE$  is sensitive to the extraordinarily large or small deviation in the simulation process, resulting in the large value of  $RMSE$ . Therefore, the values of  $RMSE$  of the ELM model were smaller than that of KNN-FWA-ELM model. In general, KNN-FWA-ELM model had better simulation results and higher simulation accuracy than the ELM model.

A comparison of the simulated floods with the observed ones between KNN-FWA-ELM and ELM model was carried out to demonstrate the differences between the two models, and the results are illustrated in Figure 15.

Figure 15 illustrates that the simulated floods of both KNN-FWA-ELM model and the ELM model showed fluctuations from the observed floods, and KNN-FWA-ELM model fitted better than that of the ELM model. Such fluctuation is related to the random parameter generation of the data-driven model, and the change of underlying surface conditions have a different level of effects on the flooding process.

It can be seen that in the ELM model, input weight and offset randomly determined might not be optimal; while with regard to KNN-FWA-ELM model, FWA was set to find the optimal values of input weight and offset and KNN method was employed to make the full use of early discharge data, which could correct the simulation results if the observed discharge data are completed and give continuous forecasting if the observed discharge data are absent, thus reducing the instability and improving the simulation accuracy. According to the comparison between the KNN-FWA-ELM model and the ELM model (as shown in Table 5) and the aforementioned analysis, it is evident that the KNN-FWA-ELM model performed better in simulation than the ELM model.

#### 4.3.2. Comparison of Simulation Results of Floods in Different Periods

In this section, a comparison of the simulation results in  $HSP_1$ ,  $HSP_2$ , and  $HSP_3$  was conducted. The average evaluation values of the two models in three periods are shown in Table 6.

**Table 6.** Comparison of all flood simulation results in different periods.

Evaluation Index	ELM Model			KNN-FWA-ELM Model		
	HSP <sub>1</sub>	HSP <sub>2</sub>	HSP <sub>3</sub>	HSP <sub>1</sub>	HSP <sub>2</sub>	HSP <sub>3</sub>
Qualified rate/%	78.57	50.00	71.43	92.86	77.78	78.57
$\Delta Q$ /%	-13.79	-13.34	-12.08	-8.52	-6.03	-5.08
$\Delta h$ /h	-0.42	-0.53	-0.32	-0.35	-0.31	-0.15
$NS$	0.80	0.76	0.76	0.86	0.81	0.83
$R^2$	0.85	0.80	0.81	0.89	0.83	0.86
$RMSE/m^3/s$	10.80	7.97	5.79	12.37	10.21	4.76
$MSRE$	1.55	20.18	1.59	1.55	12.43	1.56
$MARE$	0.61	1.35	0.63	0.51	1.01	0.60

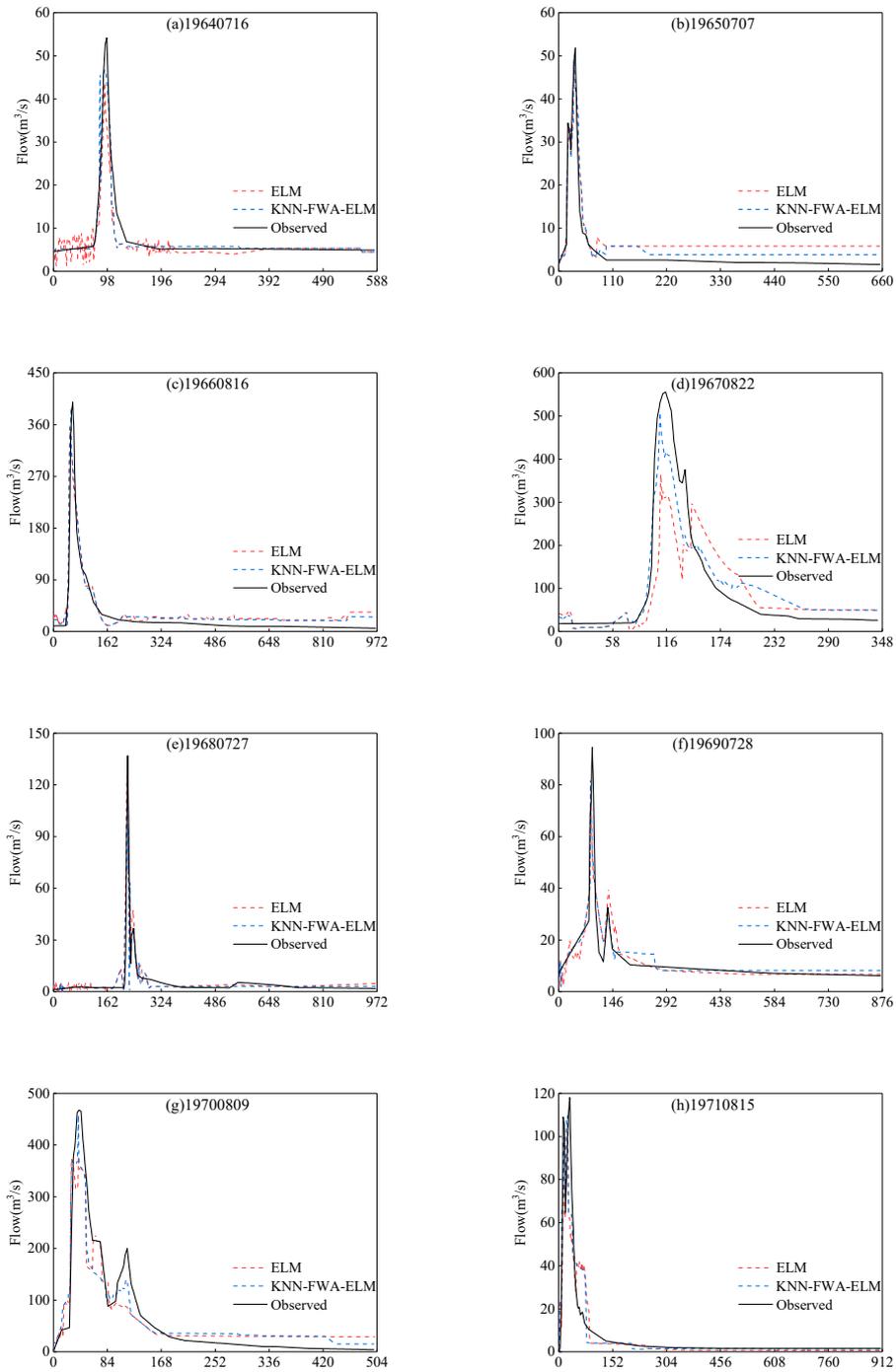


Figure 15. Cont.

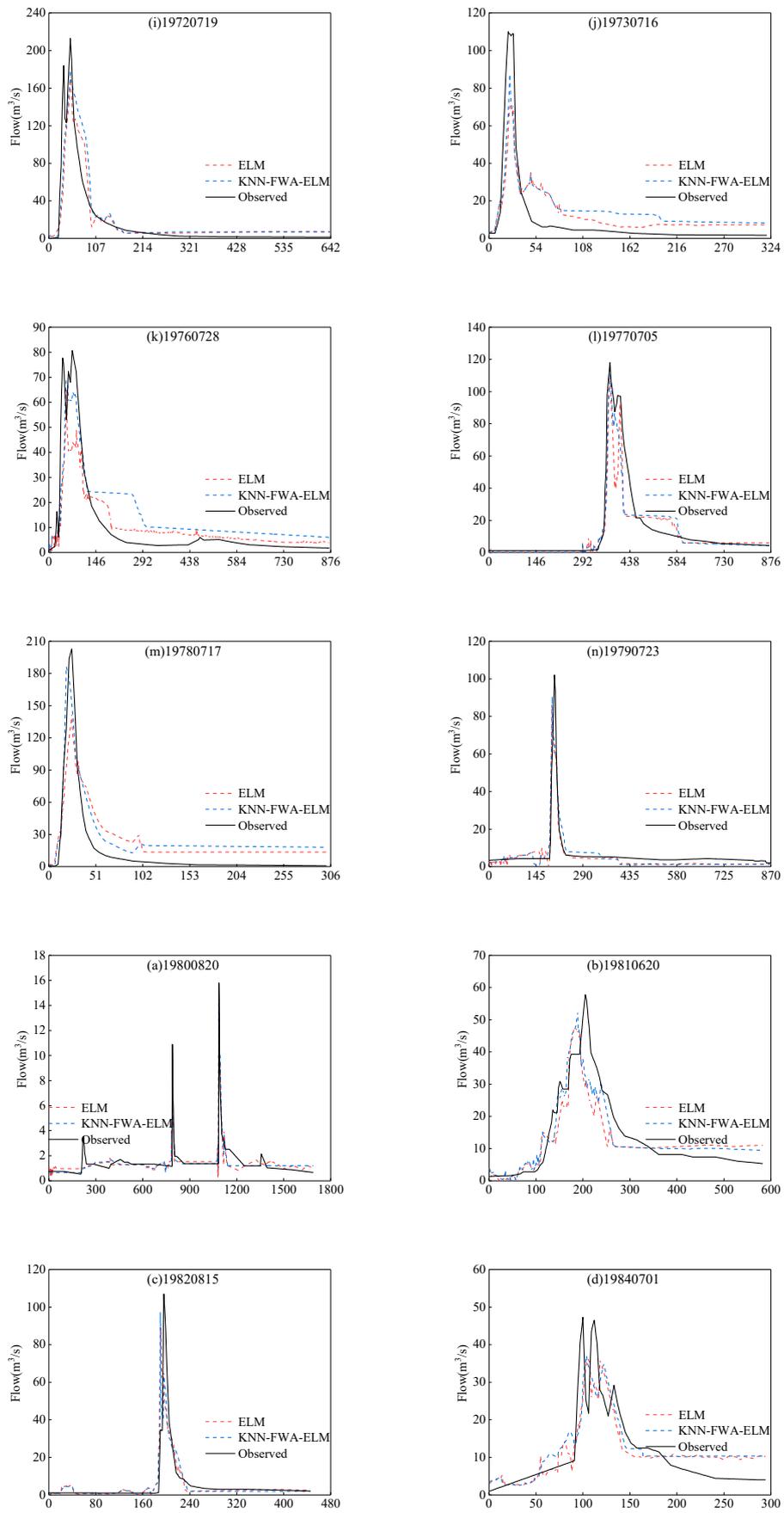


Figure 15. Cont.

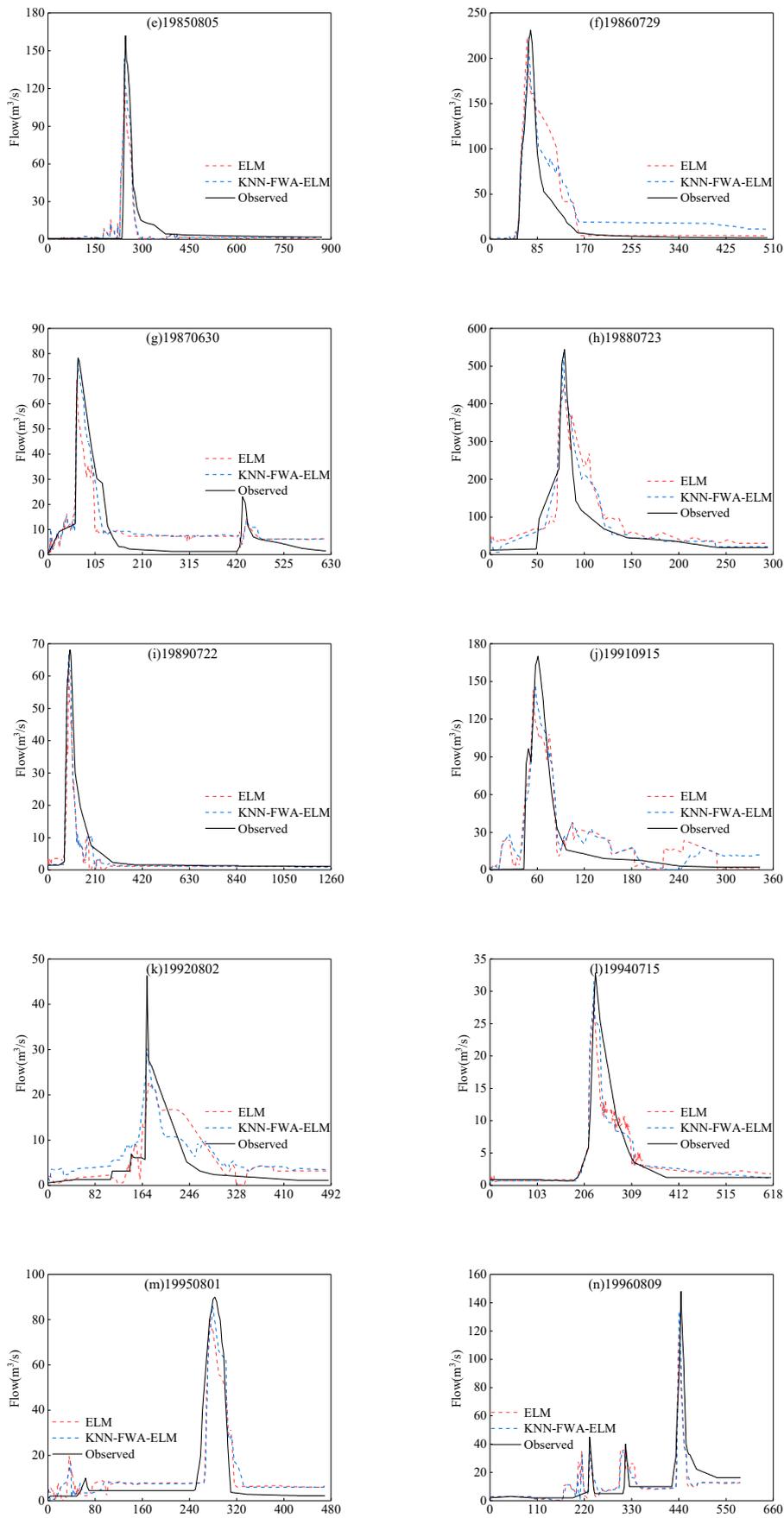


Figure 15. Cont.

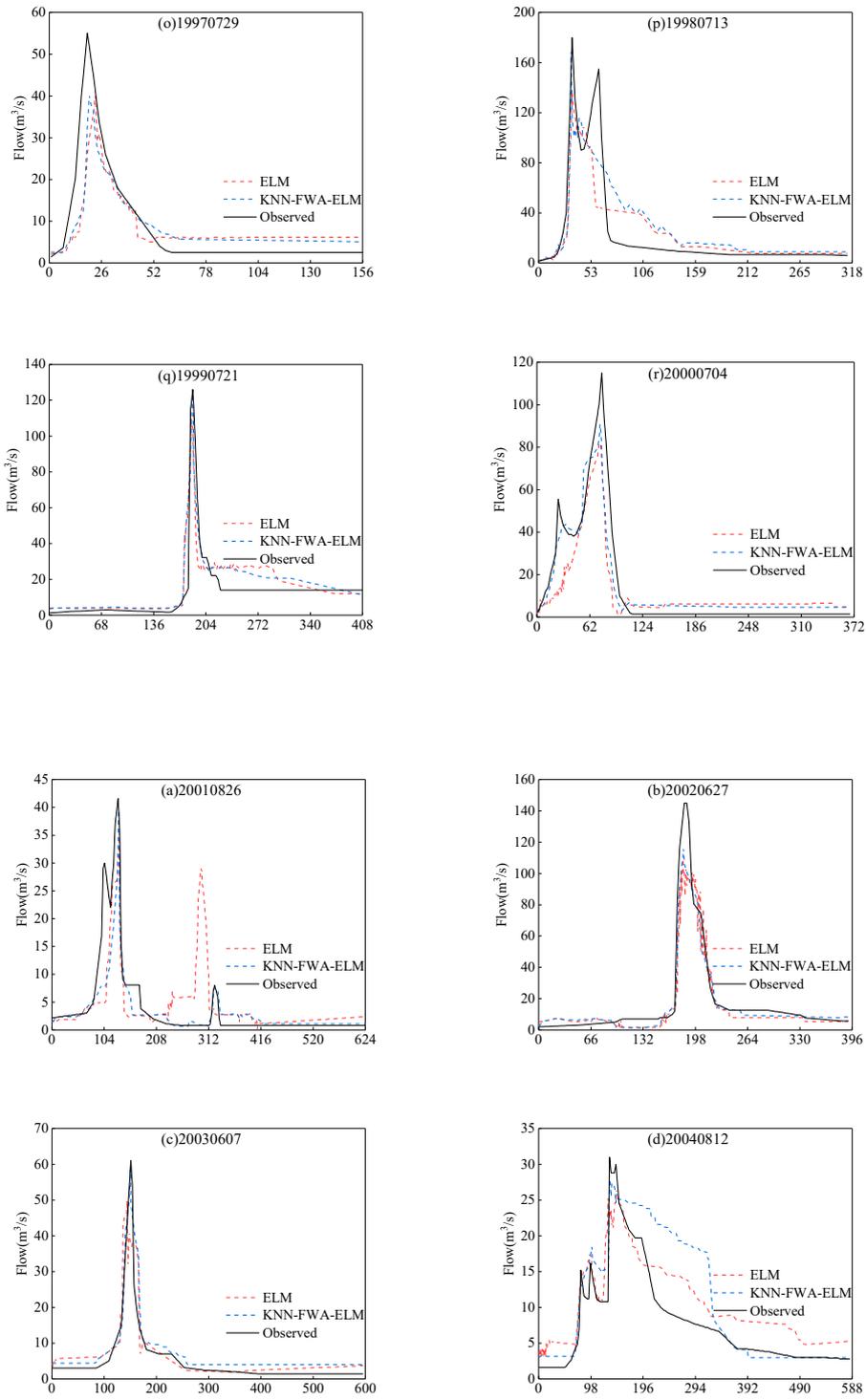


Figure 15. Cont.

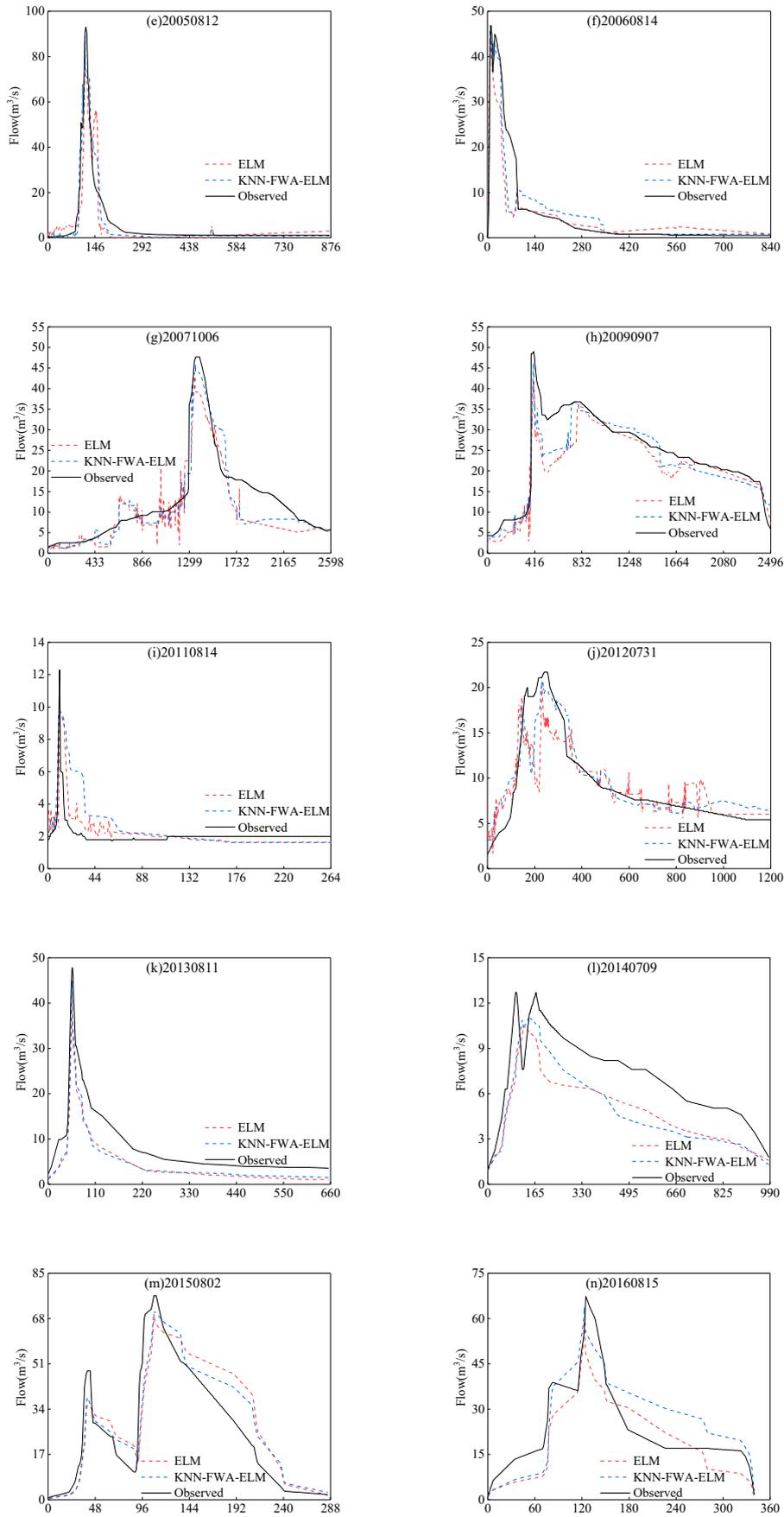


Figure 15. Comparison of KNN-FWA-ELM and ELM simulation processes.

It can be seen from Table 6 that the simulation qualification rate of KNN-FWA-ELM model was higher than that of the ELM model, which increased by 18.19%, 55.56%, and 9.99%, respectively, in HSP<sub>1</sub>, HSP<sub>2</sub>, and HSP<sub>3</sub>. The absolute of  $\Delta Q$  of KNN-FWA-ELM model was smaller than that of the ELM model, which decreased by 38.22%, 54.80%, and 57.95%, respectively. The absolute of  $\Delta h$  of KNN-FWA-ELM model was smaller than that of the ELM model, shortened by 0.07 h, 0.22 h, and 0.17 h, respectively. The reliability of KNN-FWA-ELM model was higher than that of the ELM model, with an increase of 7.50%, 6.58%, and 9.21%, respectively. The  $R^2$  of KNN-FWA-ELM model was higher than that of the ELM model, with an increase of 4.71%, 3.75%, and 6.17%, respectively. In the two periods of HSP<sub>1</sub> and HSP<sub>2</sub>, the *RMSE* of KNN-FWA-ELM model was greater than that of the ELM model due to the distribution of great floods and the low qualified rate of ELM model for the simulation of great floods. In HSP<sub>3</sub>, the *RMSE* of KNN-FWA-ELM model was 17.79% lower than that of the ELM model. In the two periods of HSP<sub>2</sub> and HSP<sub>3</sub>, the *MSRE* of KNN-FWA-ELM model was 38.40% and 1.89% lower than that of the ELM model, respectively. The *MSRE* is a good efficiency index for moderate data values. In HSP<sub>2</sub>, the *MSRE* of some flood events was relatively large; this is because, during the moderate-flow process of these events, there is a big difference between the simulated value and the observed value. The *MARE* of KNN-FWA-ELM model was smaller than that of the ELM model, which decreased by 16.39%, 25.19%, and 4.76%, respectively. Compared with HSP<sub>2</sub> and HSP<sub>3</sub>, the great floods are predominantly distributed in HSP<sub>1</sub>, according to the relevant literature reports [50]. The flow velocity of the great floods is stable, the nonlinear characteristics are weak, and the rest two periods are given priority to with moderate and small floods. The flood magnitude is relatively small, and the nonlinear characteristics are strong, leading to slightly poor simulation results.

To measure the simulation performance of KNN-FWA-ELM model, eight evaluation indexes were adopted. They were qualified rate,  $\Delta Q$ ,  $\Delta h$ , *NS*,  $R^2$ , *RMSE*, *MSRE*, and *MARE*. The comparison between KNN-FWA-ELM model and the ELM model are given in Table 6. Depending on the aforementioned analysis, it is concluded that the hybrid KNN-FWA-ELM model could achieve better-simulated results with significant improvement based on eight statistical evaluation indexes for flood forecasting in HSP<sub>1</sub>, HSP<sub>2</sub>, and HSP<sub>3</sub>, with HSP<sub>1</sub> as the top performance.

#### 4.3.3. Comparison of Simulation Results of Floods Under Different Grades

Based on the division of flood grades in Section 2.2.3, KNN-FWA-ELM model and the ELM model were compared and analyzed for simulation results of great, moderate, and small floods in 46 floods from 1964 to 2016.

The simulation performance of the two models on great floods is shown in Table 7.

**Table 7.** Comparison of great flood simulation results.

Evaluation Index		$\Delta Q/\%$	$\Delta h/h$	<i>NS</i>	$R^2$	<i>RMSE/m<sup>3</sup>/s</i>	<i>MSRE</i>	<i>MARE</i>	Qualified or Not
ELM model	19660816	-10.89	-0.75	0.83	0.88	19.88	2.02	0.86	Qualified
	19670822	-34.43	-0.42	0.60	0.63	85.43	0.48	0.65	Not qualified
	19700809	-19.18	-0.08	0.86	0.89	36.59	3.15	0.92	Qualified
	19880723	-17.22	0	0.66	0.74	53.29	1.48	0.89	Not qualified
	Average <sup>1</sup>	-15.04	-0.42	0.85	0.89	28.24	2.59	0.89	-
KNN-FWA-ELM model	19660816	-2.27	-0.42	0.86	0.90	18.06	1.43	0.61	Qualified
	19670822	-8.58	-0.50	0.89	0.95	43.18	0.31	0.49	Qualified
	19700809	-2.09	-0.08	0.88	0.91	32.72	2.01	0.75	Qualified
	19880723	-1.03	-0.08	0.84	0.87	36.44	0.68	0.49	Qualified
	Average <sup>1</sup>	-3.49	-0.27	0.87	0.91	32.60	1.11	0.59	-

<sup>1</sup> The average value in Table 7 is obtained from the qualified flood events.

It can be seen in Table 7 that the qualified rate of the ELM model and KNN-FWA-ELM model for great floods simulation was 50% and 100%, respectively. Compared with the ELM model, the average  $\Delta Q$  of KNN-FWA-ELM model was reduced by 76.80%. The average  $\Delta h$  of KNN-FWA-ELM model was shortened by 0.15 h. The average *NS* and  $R^2$  were improved by 2.35% and 2.25%, respectively.

The average *RMSE* of KNN-FWA-ELM model was greater than that of the ELM model. The average *MSRE* and *MARE* of KNN-FWA-ELM model were smaller than that of the ELM model, decreasing by 57.14% and 33.71%, respectively. In summary, the KNN-FWA-ELM model had a better simulation performance on the great floods than the ELM model.

The simulation performance of two models on moderate floods is shown in Table 8.

**Table 8.** Comparison of moderate flood simulation results.

Evaluation Index	$\Delta Q/\%$	$\Delta h/h$	<i>NS</i>	$R^2$	<i>RMSE/m<sup>3</sup>/s</i>	<i>MSRE</i>	<i>MARE</i>	Qualified or Not	
ELM model	19720719	-18.69	0	0.82	0.83	15.19	6.09	1.04	Qualified
	19780717	-29.89	0	0.69	0.85	19.51	17.63	2.66	Not qualified
	19860729	-3.33	-0.50	0.68	0.80	22.72	6.55	4.27	Not qualified
	19980713	-23.91	0	0.69	0.71	20.36	0.81	0.69	Not qualified
	Average <sup>1</sup>	-18.69	0	0.82	0.83	15.19	6.09	1.04	-
KNN-FWA-ELM model	19720719	-15.87	0	0.86	0.86	13.59	4.36	0.73	Qualified
	19780717	-8.24	-0.50	0.82	0.93	14.88	8.78	1.35	Qualified
	19860729	-2.91	-0.33	0.79	0.90	18.20	5.70	1.05	Qualified
	19980713	-6.13	-0.08	0.81	0.83	15.93	0.69	0.59	Qualified
	Average <sup>1</sup>	-8.29	-0.23	0.82	0.88	15.65	4.88	0.93	-

<sup>1</sup> The average value in Table 8 is obtained from the qualified flood events.

It can be seen in Table 8 that the qualified rate of the ELM model and KNN-FWA-ELM model for moderate floods simulation was 25% and 100%, respectively. The KNN-FWA-ELM model reduced the average  $\Delta Q$  by 55.64% compared with the ELM model. The absolute of average  $\Delta h$  of KNN-FWA-ELM model was greater than that of the ELM model. The average reliability of the two models was the same. The average  $R^2$  of KNN-FWA-ELM model was 6.02% higher than that of the ELM model. The average *RMSE* of KNN-FWA-ELM model was 3.03% greater than that of the ELM model. The average *MSRE* and *MARE* of KNN-FWA-ELM model were smaller than that of the ELM model, which were decreased by 19.87% and 10.58%, respectively. Overall, it could be concluded that KNN-FWA-ELM model was better than the ELM model in simulating moderate floods.

There were 38 small floods in the basin from 1964 to 2016. The simulation performance of the two models on small floods is shown in Table 9.

**Table 9.** Comparison of small flood simulation results.

Evaluation Index	ELM Model	KNN-FWA-ELM Model
Qualified rate/%	71.05	78.95
$\Delta Q/\%$	-12.74	-6.80
$\Delta h/h$	-0.43	-0.29
<i>NS</i>	0.77	0.83
$R^2$	0.82	0.85
<i>RMSE/m<sup>3</sup>/s</i>	6.54	5.44
<i>MSRE</i>	7.53	6.25
<i>MARE</i>	0.83	0.71

It can be seen in Table 9 that the qualified rate of KNN-FWA-ELM model for small floods simulation increased by 11.12% compared with the ELM model. The average  $\Delta Q$  of KNN-FWA-ELM model decreased by 46.62%. The average  $\Delta h$  of KNN-FWA-ELM model was shortened by 0.14 h. The average *NS* and  $R^2$  of KNN-FWA-ELM model increased by 7.79% and 3.66%, respectively. The average *RMSE*, *MSRE*, and *MARE* of KNN-FWA-ELM model were lower than that of the ELM model, which were decreased by 16.82%, 17%, and 14.46%, respectively. The above results demonstrated that KNN-FWA-ELM model performed better than the ELM model in simulating small floods.

Based on the eight statistical evaluation indexes and the aforementioned analysis, it is evident that KNN-FWA-ELM model had better simulation performance for great, moderate, and small floods than the ELM model, and the KNN-FWA-ELM model for great floods simulation performance was

better due to the steady flow velocity of the great floods and the weak nonlinear characteristics of the flood process. While the magnitude of the moderate and small floods was relatively small, and the nonlinear characteristics were strong. Therefore, KNN-FWA-ELM model got preferable simulation performance on great floods. This is consistent with previous research of machine learning flood forecasting model [24].

## 5. Conclusions

In this study, a novel hybrid KNN-FWA-ELM approach, in which k nearest neighbor method, fireworks algorithm, and extreme learning machine are integrated, was proposed for flood forecasting in a medium and small watershed of loess region. This integrated approach could increase the robustness of flood forecasting results. We finally have drawn the following conclusions.

For the special topography, micro-geomorphology, vegetation, and other underlying surface conditions of the loess region, the advantages and disadvantages of the ELM model in the flood forecasting of typical medium and small watershed in the loess region were analyzed. To realize model parameter optimization and continuous flood forecasting, MATLABR2016b was used as the platform to self-program, and the ELM model was improved by using fireworks algorithm and k nearest neighbor method. FWA was adopted to find the optimal values of input weight and offset, and KNN method was employed to make the full use of early discharge data. A hybrid data-driven model (KNN-FWA-ELM) for flood forecasting in a medium and small watershed in the loess region was proposed and established, and the changes of the underlying surface characteristics of the basin in different periods were considered in the input and output of the model.

The KNN-FWA-ELM and ELM models were applied to the Gedong Basin, and the floods in different periods and different grades were simulated. The qualified rate,  $\Delta Q$ ,  $\Delta h$ ,  $NS$ ,  $R^2$ ,  $RMSE$ ,  $MSRE$ , and  $MARE$  were employed to measure the flood simulation results of the two models. The performance comparison of their simulated results in the present study demonstrated that the qualified rate of KNN-FWA-ELM model was 82.61%, and the ELM model was 65.22%; the absolute of  $\Delta Q$  of KNN-FWA-ELM model was 6.47% less than that of the ELM model; the absolute of  $\Delta h$  of KNN-FWA-ELM model was 0.14 h shorter than that of the ELM model; the higher values of the  $NS$  and  $R^2$  indicated that KNN-FWA-ELM model gave a better performance than the ELM model; the smaller values of  $MSRE$  and  $MARE$  indicated the high simulation precision of KNN-FWA-ELM model. In short, the developed hybrid KNN-FWA-ELM model could provide a significant improvement in flood forecasting.

Depending on the performance comparison of the two models in HSP<sub>1</sub>, HSP<sub>2</sub>, and HSP<sub>3</sub>, it is concluded that KNN-FWA-ELM model could achieve better-simulated results with significant improvement based on eight statistical evaluation indexes for flood forecasting and demonstrate the best performance on HSP<sub>1</sub>.

Based on the eight statistical evaluation indexes and the performance comparison of the two models for great, moderate, and small floods, it is evident that KNN-FWA-ELM model had better simulation performance and gave the best results for great floods.

The proposed hybrid KNN-FWA-ELM model obtained a good flood forecasting result in a medium and small watershed of loess region in our study, but there are drawbacks, such as the number of the hidden layer nodes were determined by a human. In our future studies, we will determine the number of optimal hidden layer nodes through the ELM itself by changing the nodes dynamically, and we will compare the simulated flood forecasting results of the hybrid KNN-FWA-ELM model and the hydrological model under different underlying surface conditions and different flood grades. We will also apply this hybrid model to other medium and small watershed of loess region in China. In short, the developed hybrid model could provide a significant improvement to flood forecasting.

**Author Contributions:** Q.Z. and X.Z. conceived and designed the research theme. B.R. collected the data and designed methods. J.R. analyzed the data and interpreted the results. J.R. and B.R. wrote and edited the paper.

**Funding:** This research was funded by the Government Financial Grants Project (ZNGZ2015-036) and the APC was funded by the Government Financial Grants Project (ZNGZ2015-036).

**Acknowledgments:** This research was supported by the Government Financial Grants Project (ZNGZ2015-036). We thank the editors of the journal and the reviewers for their useful comments and suggestions to improve the paper quality greatly.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Sun, Y.X.; Tang, D.S.; Sun, Y.F.; Cui, Q.F. Comparison of a fuzzy control and the data-driven model for flood forecasting. *Nat. Hazards* **2016**, *82*, 827–844. [[CrossRef](#)]
2. Zhou, Y.L.; Guo, S.L.; Chang, F.J. Explore an evolutionary recurrent ANFIS for modeling multi-step-ahead flood forecasts. *J. Hydrol.* **2019**, *570*, 343–355. [[CrossRef](#)]
3. Nguyen, P.K.T.; Chua, L.H.C.; Son, L.H. Flood forecasting in large rivers with data-driven models. *Nat. Hazards* **2014**, *71*, 767–784. [[CrossRef](#)]
4. Badrzadeh, H.; Sarukkalige, R.; Jayawardena, A.W. Hourly runoff forecasting for flood risk management: Application of various computational intelligence models. *J. Hydrol.* **2015**, *529*, 1633–1643. [[CrossRef](#)]
5. Seo, Y.M.; Kim, S.W.; Singh, V.J. Multistep-ahead flood forecasting using wavelet and data-driven methods. *KSCE J. Civ. Eng.* **2015**, *19*, 401–417. [[CrossRef](#)]
6. Amirian, E.; Deiam, M.; Chen, Z.X. Performance forecasting for polymer flooding in heavy oil reservoirs. *Fuel* **2018**, *216*, 83–100. [[CrossRef](#)]
7. Nguyen, P.K.T.; Chua, L.H.C. The data-driven approach as an operational real-time flood forecasting model. *Hydrol. Process.* **2012**, *26*, 2878–2893. [[CrossRef](#)]
8. Matos, J.P.; Portela, M.M.; Schleiss, A.J. Towards safer data-driven forecasting of extreme streamflows. *Water Resour. Manag.* **2018**, *32*, 701–720. [[CrossRef](#)]
9. Lima, A.R.; Cannon, A.J.; Hsieh, W.W. Nonlinear regression in environmental sciences using extreme learning machines: A comparative evaluation. *Environ. Model. Softw.* **2015**, *73*, 175–188. [[CrossRef](#)]
10. Hsieh, W.W. *Machine Learning Methods in the Environmental Sciences*; Cambridge University Press: Cambridge, England, 2009.
11. Cherkassky, V.; Krasnopolsky, V.; Solomatine, D.P. Computational intelligence in earth sciences and environmental applications: Issues and challenges. *Neural Netw.* **2006**, *19*, 113–121. [[CrossRef](#)]
12. Shamseldin, A.Y. Artificial neural network model for river flow forecasting in a developing country. *J. Hydroinform.* **2010**, *12*, 22–35. [[CrossRef](#)]
13. Maier, H.R.; Jain, A.; Dandy, G.C. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environ. Model. Softw.* **2010**, *25*, 891–909. [[CrossRef](#)]
14. Chang, F.J.; Chen, P.A.; Lu, Y.R. Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control. *J. Hydrol.* **2014**, *517*, 836–846. [[CrossRef](#)]
15. Wu, W.; Dandy, G.C.; Maier, H.R. Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modeling. *Environ. Model. Softw.* **2014**, *54*, 108–127. [[CrossRef](#)]
16. Siqueira, H.; Boccato, L.; Attux, R. Echo state networks and extreme learning machines: A comparative study on seasonal streamflow series prediction. In *Neural Information Processing, Proceedings of the International Conference on Neural Information Processing, Doha, Qatar, 12–15 November 2012*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 491–500.
17. Hadi, S.J.; Tombul, M. Forecasting daily streamflow for basins with different physical characteristics through data-driven methods. *Water Resour. Manag.* **2018**, *32*, 3405–3422. [[CrossRef](#)]
18. He, Z.; Wen, X.; Liu, H. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *J. Hydrol.* **2014**, *509*, 379–386. [[CrossRef](#)]
19. Taormina, R.; Chau, K.W. Data-driven input variable selection for rainfall–runoff modeling using binary-coded particle swarm optimization and Extreme Learning Machines. *J. Hydrol.* **2015**, *529*, 1617–1632. [[CrossRef](#)]

20. Lima, A.R.; Cannon, A.J.; Hsieh, W.W. Forecasting daily streamflow using online sequential extreme learning machines. *J. Hydrol.* **2016**, *537*, 431–443. [[CrossRef](#)]
21. Wang, J.J.; Shi, P.; Jiang, P. Application of BP Neural Network Algorithm in Traditional Hydrological Model for Flood Forecasting. *Water* **2017**, *9*, 48. [[CrossRef](#)]
22. Kong, J.; Li, S.J.; Zhu, Y.L. Flood forecasting for small and medium-sized rivers by ensemble extreme learning machine. *J. Hydrol.* **2018**, *38*, 67–72. (In Chinese)
23. Liu, J. The application of parallel extreme learning machine in flood forecasting. *J. Northwest. Univ. (Nat. Sci. Ed.)* **2015**, *45*, 545–550. (In Chinese)
24. Kan, G.Y.; Hong, Y.; Liang, K. Research on the flood forecasting based on coupled machine learning model. *China Rural Water Hydropower* **2018**, *10*, 165–169. (In Chinese)
25. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [[CrossRef](#)]
26. Abdullah, S.S.; Malek, M.A.; Abdullah, N.S.; Kisi, O.; Yap, K.S. Extreme learning machines: A new approach for prediction of reference evapotranspiration. *J. Hydrol.* **2015**, *527*, 184–195. [[CrossRef](#)]
27. Wan, C.; Song, Y.H.; Xu, Z.; Yang, G.Y.; Nielsen, A.H. Probabilistic wind power forecasting with hybrid artificial neural networks. *Electr. Power Compon. Syst.* **2016**, *44*, 1656–1668. [[CrossRef](#)]
28. Bai, Y.; Chen, Z.Q.; Xie, J.J. Daily reservoir inflow forecasting using multiscale deep feature learning with hybrid models. *J. Hydrol.* **2016**, *532*, 193–206. [[CrossRef](#)]
29. Deo, R.C.; Tiwari, M.K.; Adamowski, J.F.; Quilty, J.M. Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. *Stoch. Environ. Res. Risk Assess.* **2017**, *31*, 1211–1240. [[CrossRef](#)]
30. Ali, M.; Deo, R.C.; Downs, N.J.; Maraseni, T. Multi-stage committee based extreme learning machine model incorporating the influence of climate parameters and seasonality on drought forecasting. *Comput. Electron. Agric.* **2018**, *152*, 149–165. [[CrossRef](#)]
31. Wang, W.Z.; Jiao, J.Y. Statistic analysis on process of gully runoff and sediment yield under different rain pattern in Loess Plateau Region. *Bull. Soil Water Conserv.* **1996**, *16*, 12–18. (In Chinese)
32. Zhang, Y.C.; Liu, C.M.; Yang, S.T.; Liu, X.Y.; Cai, M.Y.; Dong, G.T.; Luo, Y. Comparison of LCM hydrological models with lumped, semi-distributed and distributed building structures in typical watershed of Yellow River Basin. *Acta Geogr. Sin.* **2014**, *69*, 90–99. (In Chinese)
33. Li, J. Study on Flood Runoff Variation Characteristics in Dalin River Basin. Master's Thesis, Xi'an University of Technology, Xi'an, China, June 2017. (In Chinese).
34. Yang, H. Analysis on the Causes and Characteristics of River Runoff Change in Beiyuhe Basin. Master's Thesis, Lanzhou University, Lanzhou, China, May 2018. (In Chinese).
35. Feng, X.W. Prediction of Secondary Flood in Semi-Arid Area Based on Multiple Combination Models. Master's Thesis, Xi'an University of Technology, Xi'an, China, June 2018. (In Chinese).
36. Ren, J.H.; Zheng, X.Q.; Zhao, X.H.; Chen, J.F.; Li, A.M.; Chen, Y.P. Characteristics of flood evolution and their contribution factors in Gedong basin of loess hilly region in Western Shanxi Province. *Water Res. Power.* **2017**, *35*, 47–50. (In Chinese)
37. Huang, G.B.; Zhou, H.; Ding, X. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2012**, *42*, 513–529. [[CrossRef](#)]
38. Zhu, Q.Y.; Qin, A.K.; Suganthan, P.N. Evolutionary extreme learning machine. *Pattern Recognit.* **2005**, *38*, 1759–1763. [[CrossRef](#)]
39. Huang, G.B.; Ding, X.; Zhou, H. Optimization method based extreme learning machine for classification. *Neurocomputing* **2010**, *74*, 155–163. [[CrossRef](#)]
40. Tan, Y.; Zhu, Y. Fireworks Algorithm for Optimization. In Proceedings of the International Conference on Advances in Swarm Intelligence, Berlin, Germany, 12–15 June 2010; pp. 355–364.
41. Cheng, S.; Qin, Q.; Chen, J. Analytics on Fireworks Algorithm Solving Problems with Shifts in the Decision Space and Objective Space. *Int. J. Swarm Intell. Res.* **2015**, *6*, 52–86. [[CrossRef](#)]
42. Altman, N.S. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *Am. Stat.* **1992**, *46*, 175–185.
43. Huang, Y.W.; Lai, D.H. Hidden Node Optimization for Extreme Learning Machine. *AASRI Procedia* **2012**, *3*, 375–380. [[CrossRef](#)]
44. Sharma, A. Seasonal to interannual rainfall probabilistic forecastings for improved water supply management: Part 1—A strategy for system predictor identification. *J. Hydrol.* **2000**, *239*, 232–239. [[CrossRef](#)]

45. Sharma, A.; Luk, K.C.; Cordery, I. Seasonal to interannual rainfall probabilistic forecastings for improved water supply management: Part 2—Predictor identification of quarterly rainfall using ocean-atmosphere information. *J. Hydrol.* **2000**, *239*, 240–248. [[CrossRef](#)]
46. Sharma, A. Seasonal to interannual rainfall probabilistic forecastings for improved water supply management: Part 3—A nonparametric probabilistic forecasting model. *J. Hydrol.* **2000**, *239*, 249–258. [[CrossRef](#)]
47. Dawson, C.W.; Abrahart, R.J.; See, L.M. HydroTest: A web-based toolbox of evaluation metrics for the standardized assessment of hydrological forecasts. *Environ. Model. Softw.* **2007**, *22*, 1034–1052. [[CrossRef](#)]
48. Jain, A.; Srinivasulu, S. Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms and artificial neural network techniques. *Water Resour. Res.* **2004**. [[CrossRef](#)]
49. Dawson, C.W.; Wilby, R.L. Hydrological modelling using artificial neural networks. *Prog. Phys. Geogr.* **2001**, *25*, 80–108. [[CrossRef](#)]
50. Zhang, D.H.; Zhang, J.C.; Liu, F.G. Some comments on nonlinear effect in catchment hydrology. *Adv. Water Sci.* **2007**, *5*, 776–784. (In Chinese)



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).