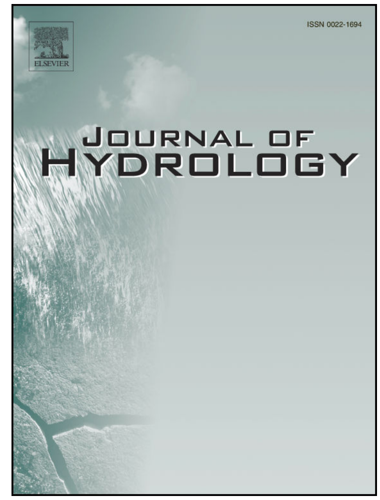
Journal Pre-proofs



Research papers

Predicting flood susceptibility using long short-term memory (LSTM) neural network model

Zhice Fang, Yi Wang, Ling Peng, Haoyuan Hong

PII: S0022-1694(20)31195-1
DOI: https://doi.org/10.1016/j.jhydrol.2020.125734
Reference: HYDROL 125734

To appear in: *Journal of Hydrology*

Received Date: 22 April 2020
Revised Date: 26 October 2020
Accepted Date: 2 November 2020

Please cite this article as: Fang, Z., Wang, Y., Peng, L., Hong, H., Predicting flood susceptibility using long short-term memory (LSTM) neural network model, *Journal of Hydrology* (2020), doi: https://doi.org/10.1016/j.jhydrol.2020.125734

Predicting flood susceptibility using long short-term memory (LSTM) neural network model

Zhice Fang [1], Yi Wang [1,*], Ling Peng [2], Haoyuan Hong [3,4,5,6,*]

[1] Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan 430074, China

[2] China Institute of Geo-Environment Monitoring, Beijing 100081, China

[3] Department of Geography and Regional Research, University of Vienna, Universitätsstraße 7, 1010 Vienna, Austria

[4] Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Ministry of Education, Nanjing, 210023, China

[5] State Key Laboratory Cultivation Base of Geographical Environment Evolution (Jiangsu Province), Nanjing, 210023, China

[6] Jiangsu Center for Collaborative Innovation in Geographic Information Resource Development and Application, Nanjing, Jiangsu 210023, China

[*]Correspondence Author: Yi Wang (cug.yi.wang@gmail.com); Haoyuan Hong (171301013@stu.njnu.edu.cn)

Abstract

Identifying floods and producing flood susceptibility maps are crucial steps for decision-makers to prevent and manage disasters. Plenty of studies have used machine learning models to produce reliable susceptibility maps. Nevertheless, most studies ignore the importance of developing appropriate feature engineering methods. In this study, we propose a local spatial sequential long short-term memory neural network (LSS-LSTM) for flood susceptibility prediction in Shangyou County, China. Three main contributions of this study are summarized as follows. First, it is a new perspective that the deep learning technique of LSTM is used for flood susceptibility prediction. Second, we integrate an appropriate feature engineering method with LSTM to predict flood susceptibility. Third, we implement two optimization techniques of data augmentation and batch normalization to further improve the performance of the proposed method. The LSS-LSTM method can not only capture both attribution information of flood conditioning factors and local spatial information of flood data, but also retain the powerful sequential modelling capability to deal with flood spatial relationship. Experimental results demonstrate that the LSS-LSTM method achieves satisfying prediction performance (93.75% and 0.965) in terms of accuracy and area under the ROC curve.

Key words: Flood susceptibility prediction; long short-term memory neural network; deep learning; feature engineering.

1. Introduction

Floods are one of the most common and disastrous natural hazards in the world (Giovannettone et al., 2018). As reported by the United Nations Office for Disaster Risk Reduction (UNDRR), 150,016 floods occurred between 1996 and 2015, severely affecting natural systems and human activities (Hong et al., 2018a). Flood susceptibility is the possibility of flooding in an area based on a range of geo-environmental conditions (Ahmadlou et al., 2018; Bui et al., 2019b). Flood susceptibility prediction (FSP) can provide helpful guidance for decision-makers to effectively manage and prevent flood hazards. Therefore, producing reliable and accurate susceptibility maps is important for flood-prone areas.

In recent years, machine learning techniques for FSP have exhibited powerful capability and achieved successful results, including decision tree (Bui et al., 2019c; Choubin et al., 2019; Khosravi et al., 2018), support vector machine (Choubin et al., 2019; Tehrany et al., 2014), random forest (Chapi et al., 2017; Zhao et al., 2019), and artificial neural network (Campolo et al., 2003; Gebrehiwot et al., 2019). Recently, some studies attempt to use hybrid strategy to obtain more powerful models for FSP. For example, subsampling and bootstrapping algorithm are combined with machine learning models to predict flood susceptibility (Dodangeh et al., 2020). Researchers integrate frequency ratio and logistic regression model for FSP (Costache et al., 2020b). Reduced-error pruning tree models are integrated with bagging and random subspace ensemble strategies (Chen et al., 2019). Moreover, some researchers use meta-heuristic optimization algorithms to find the optimal parameters of intelligence models (Ahmadlou et al., 2018; Bui et al., 2019a; Bui et al., 2019b; Wang et al., 2019c). These methods use different strategies to capture flood occurrence characteristics from existed background information and then predict unknown flood locations.

Feature engineering is an essential step in machine learning, which use domain knowledge of the data to create features that make models work better (Turner et al., 1999). In FSP, feature engineering can convert raw flood data into specific data representations that better portray the susceptibility prediction task to the predictive models. This operation determines the processing perspective of flood susceptibility models when facing flood data. Therefore, it is very important to develop an appropriate feature engineering method for machine learning models to better understand and learn the information of flood occurrence. In general, when using machine learning methods for FSP, the one-dimensional vector-based feature engineering method is widely used because of its convenience in operation (Chapi et al., 2017; Khosravi et al., 2019; Wang et al., 2019b). This method converts raw flood data into a set of one-dimensional feature vectors. Specifically, the entire study area is first converted to a raster form with a specified spatial resolution. Then, flood susceptibility prediction can be regarded as a binary classification process to distinguish whether a grid cell (pixel) in the study area will have flood disasters. Each grid cell is composed of a set of feature values (flood conditioning factors). Therefore, machine learning methods can predict flood susceptibility by learning these feature vectors. However, there are some drawbacks when using the machine learning methods mentioned previously for FSP. First, standard guidelines of feature engineering for flood data remain controversial. Second, various machine learning models have their own feature learning characteristics, and the appropriate feature engineering methods can maximize the classification ability of these models (Zheng and Casari, 2018). But there are few studies aimed at explore the feature engineering method for specific models for FSP. Therefore, it is essential to develop an appropriate feature engineering method for a specific prediction model to achieve reliable flood susceptibility maps.

Over the past few years, deep learning techniques have achieved inspiring results in many fields, such as pattern recognition (Hu et al., 2015), scene annotation (Zhou et al., 2014) and natural language processing (Collobert and Weston, 2008). Recently, several deep learning techniques have been successfully used for disaster susceptibility prediction, such as convolutional neural network (Fang et al., 2020a; Fang et al., 2020b; Sameen et al., 2019; Wang et al., 2019a; Zhang et al., 2019), recurrent neural network (RNN) (Wang et al., 2020b), fully connected sparse autoencoder neural network (Huang et al., 2019) and deep neural network (Bui et al., 2020; Bui et al., 2019d). Among these incredible techniques, RNN is of great interest because it can periodically capture sequential data by using a special recurrent hidden unit (LeCun et al., 2015). However, the conventional RNN has gradient vanishing and exploding problems, and is difficult to tackle long-term sequential input (Bengio et al., 1994). To tackle the above problems, a modified RNN of long short-term memory neural network (LSTM) is proposed and achieves better performance in solving sequence tasks than conventional RNNs (Hochreiter and Schmidhuber, 1997; Ma et al., 2015). In particular, LSTM has been used for flood forecasting and achieved impressive results (Le et al., 2019; Liu et al., 2018). However, the application of LSTM in regional flood susceptibility analysis is still rare. In addition, in our previous study (Wang et al., 2020a), we find that involving spatial information to flood susceptibility model can improve the prediction accuracy, but there still exists redundant spatial information in local space. The special forget mechanism of LSTM structure can remember key information and discard useless information, which can solve the above problem to a certain extent (Sak et al., 2014).

In this study, we propose a local spatial sequential long short-term memory neural network (LSS-LSTM) for FSP in Shangyou County, China. The three main contributions of this study are outlined below. First, it is a new perspective that the deep learning technique of LSTM is used as a classifier for FSP. Second, we combine an appropriate feature engineering method with LSTM to transform raw flood data into spatial sequences. Third, we implement two powerful optimization techniques of data augmentation and batch normalization to further improve the performance of the LSS-LSTM method. Based on these contributions, the LSS-LSTM method can not only capture both attribution information of flood conditioning factors and local spatial information of flood data, but also retain powerful sequential modelling ability to deal with flood spatial relationship.

2. Study area and available data

2.1. Study area

Shangyou County is located in the southern Jiangxi Province, with an area of about 1543 km^2 between the coordinates of 25°42′N to 26°01′N and 114°00′E to 114°40′E. The altitude of the study area is between 110 m and 1901 m above sea level (**Fig. 1**). Shangyou County is located in the hilly mountains in the middle of Luoxiao Mountains. The northeast, northwest, and southwest of the county are mountains, and the southeast is hills and valley basins. The altitude of the low hills is less than 200 meters above sea level and the relative altitude is less than 50 meters. The terrain of Shangyou County slopes from northwest to southeast, and the hills and valleys are mainly distributed in the southeast of the region. In Shangyou County, approximately 90% of natural land is covered by vegetation (including grass and forest). The agricultural land accounts for only 6%, and the rest areas are other types of land use.

The geological structure of Shangyou County is diverse and complex. The area is located in the uplift zone of the southern section of the Huaxia Plate, spanning the Luoxiao-Zhuguang uplift and the Yushan uplift. The magmatic activity in the study area is frequent, and the structural deformation is strong. The faults are most developed in the northwest and northeast directions. The neotectonic movement is not obvious, mainly intermittent uplift. The exposed strata are mainly the Sinian, Cambrian, Ordovician, Devonian, Carboniferous, Cretaceous and Quaternary strata. Magmatic rocks are dominated by Caledonian and Yanshanian granites. The main types of rock are metamorphic rocks and magmatic rocks. The rock mass is highly weathered and has fissures. From a climate perspective, this county belongs to the humid monsoon climate zone of the subtropical hilly region. During the period from 1959 to 2014, the annual average temperature and sunshine hours were 18.6 ℃ and 1708.3 h, respectively, and the annual average precipitation was between 933.7 and 2147.6 mm. In general, Shangyou County has abundant precipitation and extreme climatic conditions, and flood disasters often occur after heavy rainfall.

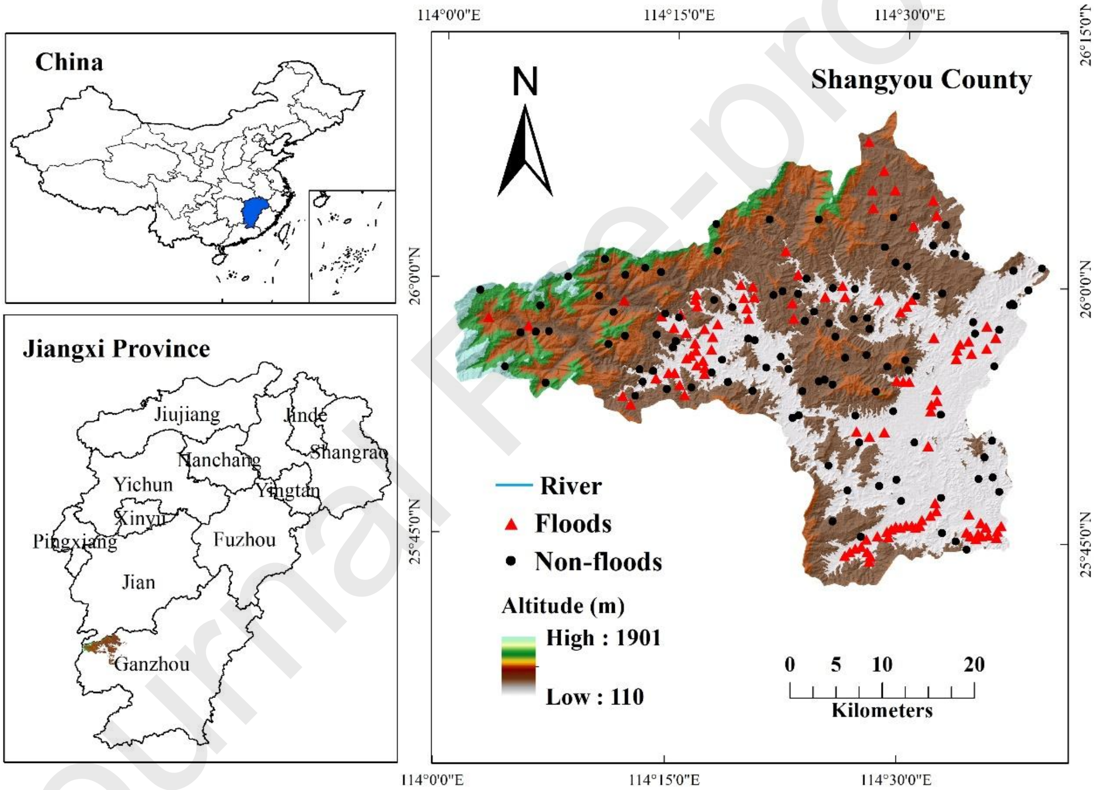


Fig. 1. Location of the study area.

2.2. Flood inventory map

Producing a reliable flood inventory map is a crucial step in flood susceptibility assessment (Gebrehiwot et al., 2019; Termeh et al., 2018). This map provides the detailed location information of inundated area. In fact, flood events are always polygons in the study area. To show the spatial distribution of floods, the practice that floods are represented as points has been used in previous publications (Diakakis et al., 2012; Salvati et al., 2010). When modelling flood susceptibility in a specific area, there is also no need to use entire polygons. In this study, the flood areas are identified by collecting previous records, extensive field surveys, and unmanned aerial vehicle data. The flood areas contain four flood events that occurred from July 26 to 29, 2006. Then, we collect 108 historical flood locations from the flood polygons to construct the flood inventory map. All the available data are obtained from Jiangxi Meteorological Bureau[1] and the Department of Civil Affairs of Jiangxi province[2]. Non-flood points are not directly available in this study area, and there is no standard guide to select accurate non-flood data. Therefore, we randomly sampled the same number of non-flood points (108) from areas without floods. This is a simple and universal sampling process widely used in previous studies (Bui et al., 2019a; Chen et al., 2019; Costache et al., 2020a). The distribution of flood and non-flood points is presented in **Fig. 1**.

2.3. Flood conditioning factors

Since flooding is triggered by a variety of environmental factors, it can ensure the reliability and accuracy of FSP results by choosing appropriate conditioning factors (Bui et al., 2020; Chapi et al., 2017). In this study, we selected flood conditioning factors based primarily on previous studies and expert knowledge. For example, flat

[1] http://jx.cma.gov.cn
[2] http://www.jxmzw.gov.cn

areas have a high potential for flooding as water flows down from higher terrain (Li et al., 2012). As for Shangyou County, river flood disasters are more likely to occur in areas of lower height and slope. Curvature indicates the degree of deformation of the slope surface. Hudson and Kesel (2000) concluded that the areas with curvature values between 1 and 2 are prone to flooding. Aspect is another key factor, as the windward slope is prone to precipitation. Aspect is related to the intensity of solar radiation, which affects the surface vegetation and soil moisture. Soil types reflect water permeability and storage capacity and directly affect drainage processes (Chapi et al., 2017; Choubin et al., 2019). Heitmuller et al. (2015) concluded that lithology determines the shape of channel and affects the development of floodplains. In addition, lithology affects the formation of soil characteristics to some extent (Tehrany et al., 2019; Zazo et al., 2018). The distance of river factor was chosen because the river network is the main way for flood discharging and expanding (Shafizadeh-Moghadam et al., 2018). Different types of land use directly or indirectly affect water infiltration and evapotranspiration (Bui et al., 2019b; Giovannettone et al., 2018; Tiwari et al., 2016). Normalized difference vegetation index (NDVI) displays the density of surface vegetation coverage, and Huang et al. (2012) studied the relationship between NDVI and flooding. Furthermore, areas with sparse vegetation cover have a high potential for flooding since its poor water storage capacity (Caprario and Finotti, 2019). NDVI is defined as follows:

$$NDVI = \frac{R_{NIR} - R_R}{R_{NIR} + R_R} \tag{1}$$

where R_{NIR} and R_R are the spectral reflectance of the near-infrared band and the red band in the electromagnetic spectrum, respectively.

The reason for choosing the annual average precipitation factor is that the floods in Shangyou County mostly occurred during or after heavy rainfall. The stream power index (SPI) reflects the erosive force of the current, which affects the stability of the terrain. Fuller (2008) studied the relationship between geomorphic conditions and floods and claimed that high stream power could lead to catastrophic channel variation. The sediment transport index (STI) factor has been widely used in flood susceptibility analysis (Chapi et al., 2017; Chen et al., 2019; Tehrany et al., 2019). STI represents the influence of terrain on erosion and reflects the intensity of sediment movement due to water movement (Werner et al., 2005). In addition, Billi (2011) concluded that the active capacity of sediment transportation can increase the frequency of floods. The factors of SPI and STI are calculated as follows (Moore et al., 1993; Moore and Wilson, 1992):

$$SPI = A_s \tan\beta \tag{2}$$

$$STI = \left(\frac{A_s}{22.13}\right)^{0.6}\left(\frac{\sin\beta}{0.0896}\right)^{1.3} \tag{3}$$

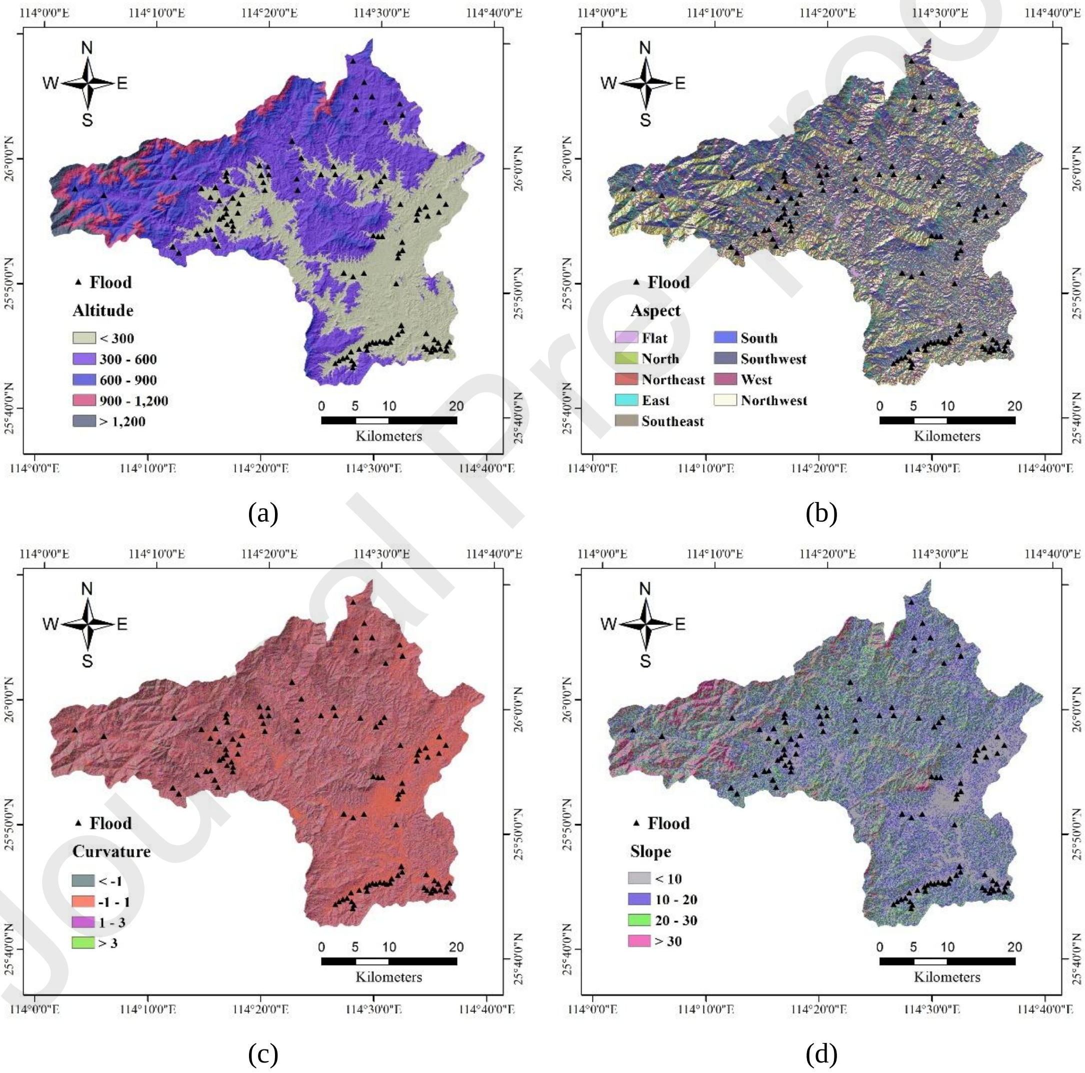
where A_s and β represent the area of the basin and the slope gradient, respectively.

Topographic wetness index (TWI) portrays soil saturated situation associated with water accumulation in the basin (Mahmoud and Gan, 2018; Tehrany et al., 2015). The TWI factor is calculated as follows (BEVEN and Kirkby, 1979):

$$TWI = \ln\left(\frac{\alpha}{\tan\beta}\right) \tag{4}$$

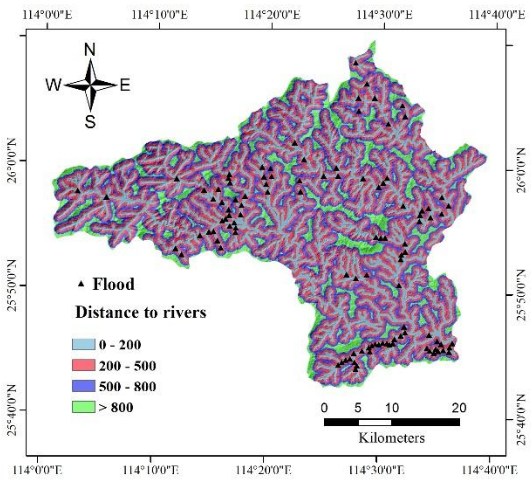
where α is the upslope area per unit contour length and β represents the slope angle.

As mentioned previously, we considered 13 flood conditioning factors for FSP in this study based on theoretical analysis and literature review (**Fig. 2**). Related information of all factors is listed in **Table 1**. It should be noted that the DEM data was acquired from ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) GDEM Version 2[3] with a spatial resolution of 30 × 30 m.

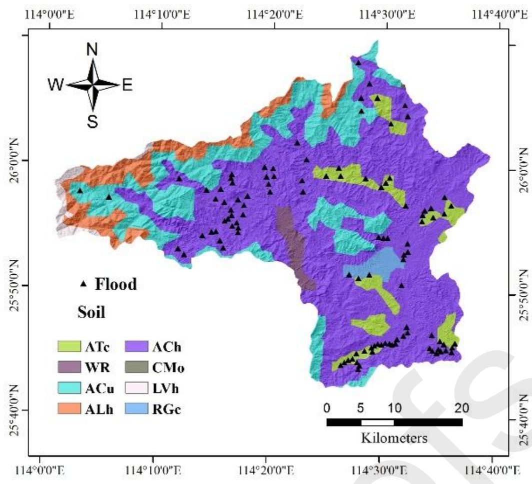


---

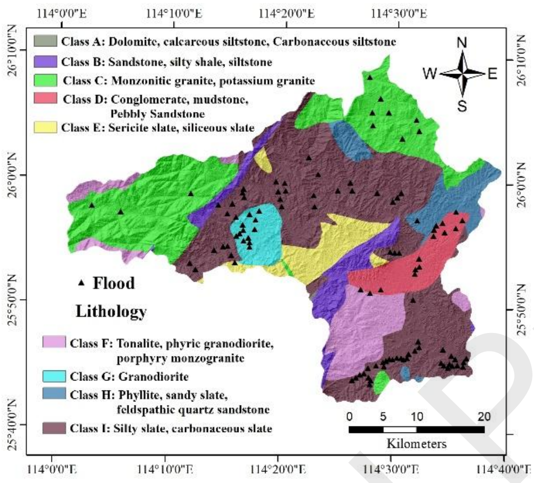
3 http://gdem.ersdac.jspacesystems.or.jp


Flood
Distance to rivers
0 - 200
200 - 500
500 - 800
> 800
0 5 10 20
Kilometers

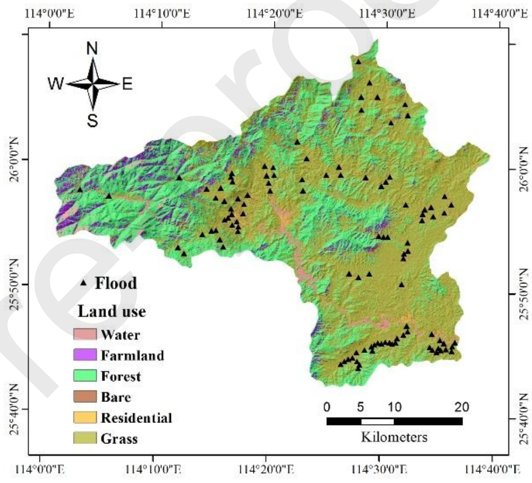
(e)


Flood
Soil
ATc
ACh
WR
CMo
ACu
LVh
ALh
RGc
0 5 10 20
Kilometers

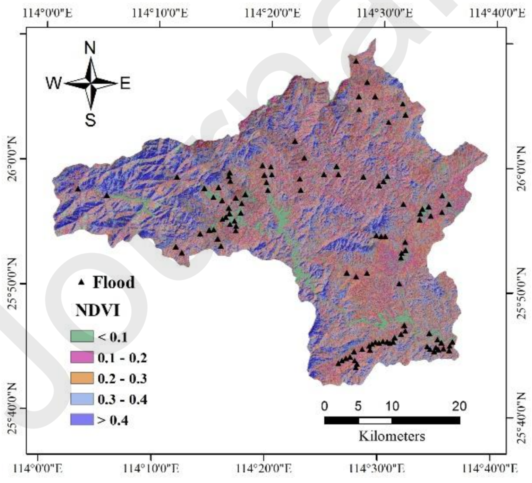
(f)


Class A: Dolomite, calcarcous siltstone, Carbonaceous siltstone
Class B: Sandstone, silty shale, siltstone
Class C: Monzonitic granite, potassium granite
Class D: Conglomerate, mudstone, Pebbly Sandstone
Class E: Sericite slate, siliceous slate
Flood
Lithology
Class F: Tonalite, phyric granodiorite, porphyry monzogranite
Class G: Granodiorite
Class H: Phyllite, sandy slate, feldspathic quartz sandstone
Class I: Silty slate, carbonaceous slate
0 5 10 20
Kilometers

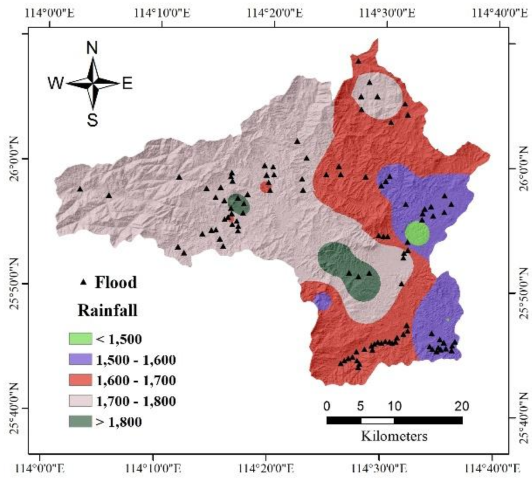
(g)


Flood
Land use
Water
Farmland
Forest
Bare
Residential
Grass
0 5 10 20
Kilometers

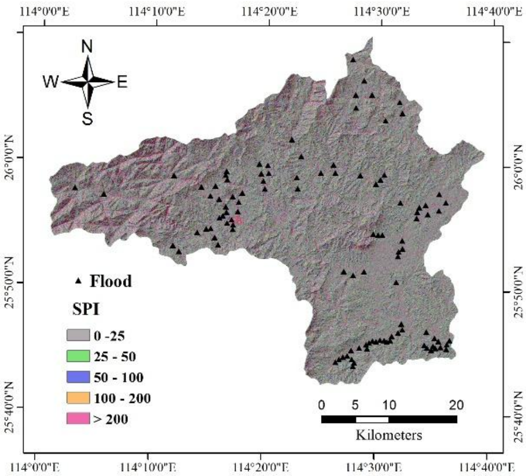
(h)


Flood
NDVI
< 0.1
0.1 - 0.2
0.2 - 0.3
0.3 - 0.4
> 0.4
0 5 10 20
Kilometers

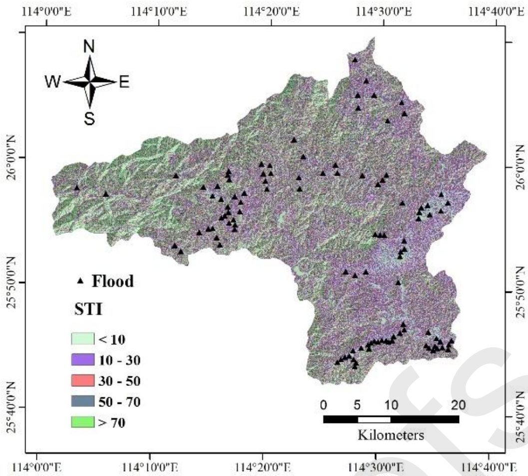
(i)


Flood
Rainfall
< 1,500
1,500 - 1,600
1,600 - 1,700
1,700 - 1,800
> 1,800
0 5 10 20
Kilometers

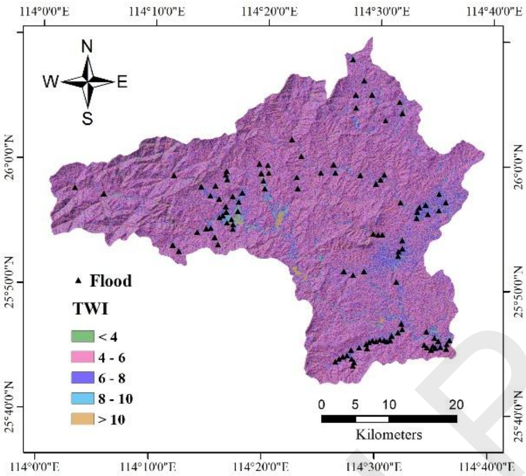
(j)



(k)



(l)



(m)

Fig. 2. **Thematic maps of flood conditioning factors. (a) Altitude, (b) aspect, (c) curvature, (d) slope, (e) distance to rivers, (f) soil, (g) lithology, (h) land use, (i) NDVI, (j) rainfall, (k) SPT, (l) STI and (m) TWI.**

Table 1 Information of the landslide conditioning factors.

Flood conditioning factors	Source	Scare/Resolution
Altitude Aspect Curvature Slope SPI STI TWI	DEM-derived	30 m
Lithology	China Geology Survey[4]	1:2,000,000
Land use NDVI	Landsat 7 ETM + images (Scene ID: LE71220422001324SGS00)	30 m
Soil	Institute of Soil Science, Chinese Academy of Sciences[5]	1:1,000,000
Distance to rivers	DEM-derived	30 m
Rainfall	Jiangxi Meteorological Bureau[6]	1:50,000

3. Methodology

3.1. Data preparation

Data preparation is an essential step before flood susceptibility modelling. The factors of altitude, aspect, curvature, and slope were calculated from the DEM data using ArcGIS software. The river networks were extracted from the topographic map and the distance to rivers factor was calculated by using Euclidean tool. The land use factor was derived from a Landsat 7 ETM+ satellite image with a classification accuracy of 85% by using the conventional maximum likelihood algorithm. This classification algorithm has excellent performance in land use classification task

[4] http://www.cgs.gov.cn
[5] http://www.issas.ac.cn
[6] http://www.weather.org.cn

(Paola and Schowengerdt, 1995). The NDVI factor was calculated from this satellite image using the ENVI software. The factors of SPI, STI, and TWI were calculated from the DEM data by using the SAGA software. Different scale data was first vectorized on the ArcGIS platform. Then, specific flood conditioning factors was extracted from these vector data and converted into a raster form. All the factors were converted to a raster format of 30 m spatial resolution, which is consistent with the DEM data. These factors were also reclassified into different categories based on previous studies, expert knowledge, and characteristics of flood spatial distributions (Costache et al., 2020a; Khosravi et al., 2019; Sameen et al., 2019). In this study, 70% flood and non-flood locations (76 and 76) were randomly selected for training models, whereas the remaining 30% flood and non-flood locations (32 and 32) were used to construct the test set.

3.2. Information gain ratio

To analyze the relationship between flood conditioning factors and flood occurrence, information gain ratio (IGR) method was used to evaluate the importance of flood conditioning factors. The IGR method is a commonly used feature selection method that has been widely used in flood susceptibility analysis (Bui et al., 2020; Chapi et al., 2017; Khosravi et al., 2019). Assuming that the training set *S* contains *n* classes, the expected information is calculated as follows:

$$\mathrm{H}(\mathrm{S}) = -\sum_{i=1}^{n} p_i \log_2 \left(p_i \right) \tag{5}$$

where p_i is the probability that a sample belongs to class C_i. The factor *A* has *m* values and its average entropy is calculated as follows:

$$\mathrm{E}(\mathrm{A}) = -\sum_{i=1}^{m} p_i H(S) \tag{6}$$

The split information value denotes the potential information obtained by splitting *S* into *m* parts corresponding to *m* outcomes on attribute *A* and can be calculated as follows:

$$SplitInfo_A(\mathrm{S}) = -\sum_{i=1}^{m} X_i (S_i/S) \log_2 (S_i/S) \tag{7}$$

Finally, the variable importance value (VI) is defined as follows:

$$VI(A) = \frac{H(S) - E(A)}{SplitInfo_A(S)} \tag{8}$$

Factors with higher VI values are more important for prediction models. If the values are equal to 0, it can be considered that the corresponding factors have no contribution to flood occurrence and should be removed from flood susceptibility modelling.

3.3. RNN and LSTM neural network

As a class of artificial neural network, RNN has received great success in the fields involving sequential data analysis (Choi et al., 2017; Ma et al., 2015; Mou et al., 2017). From the graphic structure of the regular RNN shown in **Fig. 3**, we can see that RNN can store information of the previous hidden state and apply it to the output along with the current input. In this manner, RNN is able to capture dynamic representations from sequential data by using a specific recurrent hidden state (LeCun et al., 2015).

LSTM, as a special type of RNN, is presented to capture long-term dependence of sequential data (Hochreiter and Schmidhuber, 1997). A LSTM network consists of an input layer, a hidden layer and an output layer, and its structure is similar to that of RNN. But the difference between RNN and LSTM is that the latter replaces the basic unit of the regular RNN with a memory block (Graves et al., 2013). As shown in **Fig. 4**, the memory block contains three gate functions which play different roles in information flow process.

Let $x = \{x_1, x_2, ..., x_N\}$ be a sequential input and $y = \{y_1, y_2 ..., y_N\}$ denotes the output sequence, the forget gate is a key state that determine whether the current information should be forgotten or remembered. For a certain time step *t*, it can be calculated as follows:

$$f_t = \sigma\left(W_{fx} x_t + W_{fh} h_{t-1} + b_f\right) \tag{9}$$

where W_{fx} and W_{fh} are the forget weight matrix and the forget-hidden weight matrix, respectively, b_f is the bias of the forget gate, and σ is the sigmoid function. The input gate i_t determines the information updating and $\tilde{c}_t$ memorizes the new information, which are defined as follows:

$$i_t = \sigma\left(W_{ix} x_t + W_{ih} h_{t-1} + b_i\right) \tag{10}$$

$$\tilde{c}_t = \tanh\left(W_{cx} x_t + W_{ch} h_{t-1} + b_c\right) \tag{11}$$

where W_{ix} and W_{ch} denotes the weight matrix, b_i and b_c are the bias vectors of input gate and updating cell state, respectively.

Then, the new memory cell state c_t is updated as follows:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{12}$$

where c_{t-1} is the previous memory cell state and $\odot$ represents the element-wise product.

Finally, the output gate controls the output activations. The hidden layer sent to next time step is defined as follows:

$$h_t = o_t \odot \tanh(c_t) \tag{13}$$

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \tag{14}$$

where W_{ox} is the output weight matrix, W_{oh} denotes the output-hidden weight matrix and b_o is the bias of the output gate.

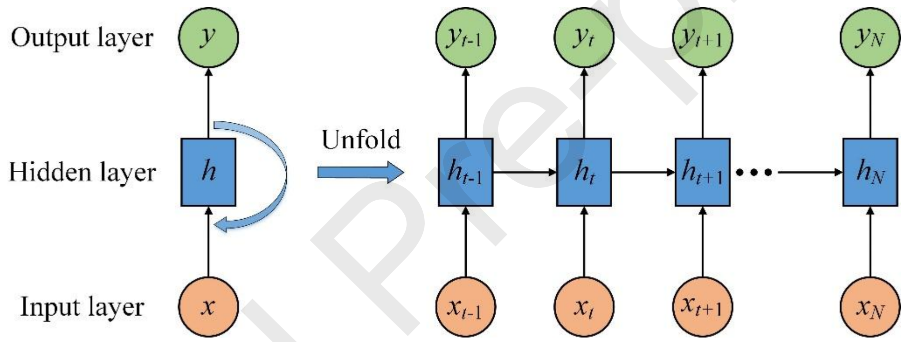


Fig. 3. Graphic structure of the regular RNN. x, h and y are the input layer, hidden layer and output layer, respectively. t is a certain time step.

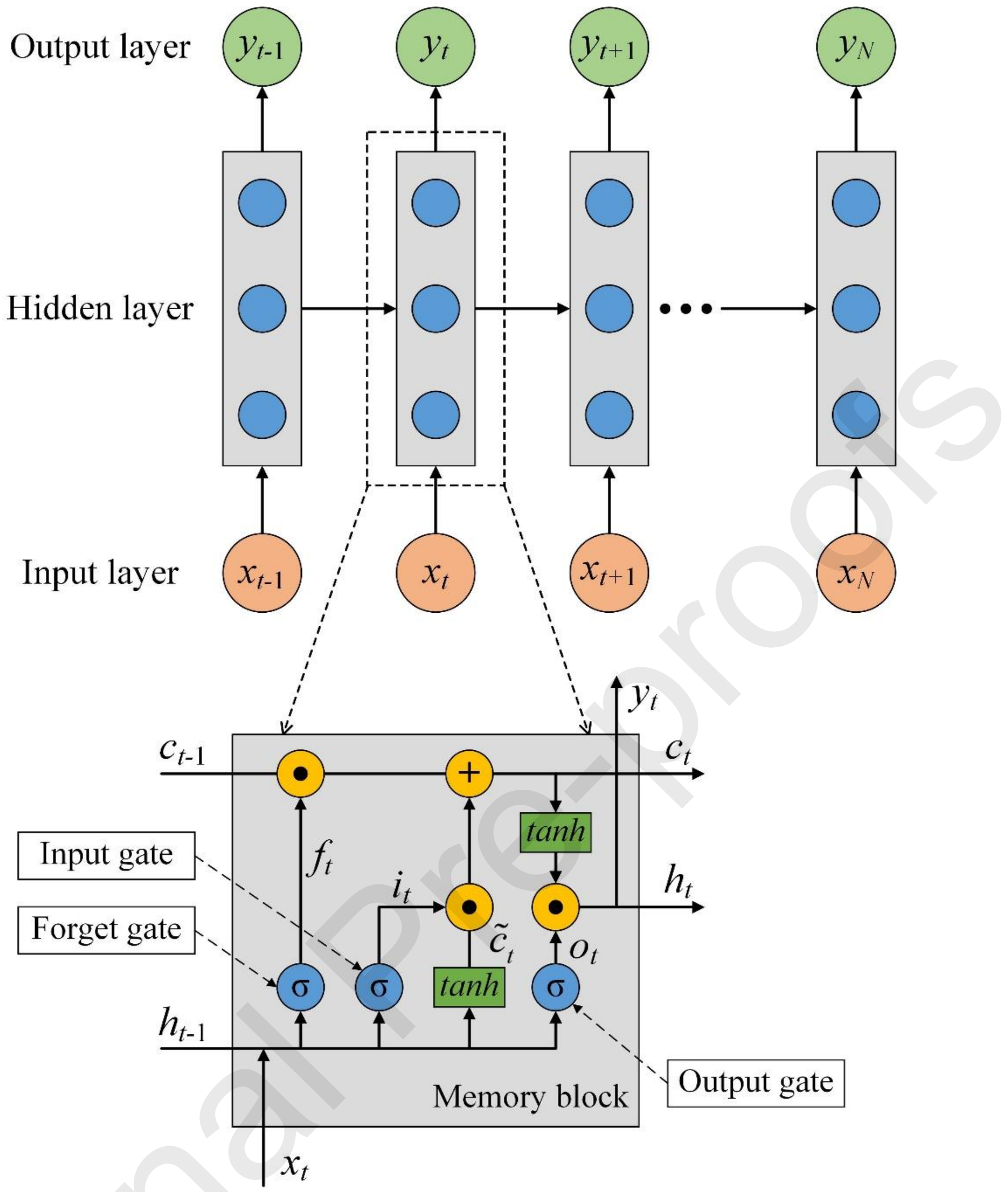


Fig. 4. Architecture of the LSTM network. x **and** y **are the input layer and output layer, respectively.** t **is a certain time step.** i_t , f_t **and** o_t **are the input gate, forget gate and output gate, respectively.** $\tilde{c}_t$ **memorizes the new information and** c_t **is new memory cell gate.**

3.4. Modelling process of LSS-LSTM

The proposed LSS-LSTM method mainly consists of three steps: layer stacking, feature engineering and LSTM construction, as illustrated in **Fig. 5**. In the layer stacking step, each flood conditioning factor can be viewed a single-band image with a size of 2156 × 1722, and all the conditioning factor layers are stacked together to form a multi-band image. In the feature engineering step, we produce each image patch pixel by pixel from the multi-band image. As shown in the **Fig. 5** (b), each central pixel and its neighboring pixels in a 3 × 3 window are first extracted, and the resultant image patch have a size of 3 × 3 × 13 that is composed of 9 vectors. Image patches contain the characteristics of the factors and the spatial information. Then, these vectors are sorted to construct sequential data with a size of 9 × 13 according to spatial continuity. In the LSTM construction step, a LSTM structure is constructed to possess these extracted sequential data. The sorted vectors are progressively sent to the LSTM architecture and the results are output only at the final time step. To the best of our knowledge, flood events are not only related to its morphological, geological and hydrological conditions, but also to the neighboring environment information (Giovannettone et al., 2018; Sampson et al., 2015). According to the inherent nature of LSTM mentioned previously, useful information for previous vectors that contribute to flood prediction can be memorized and passed to subsequent hidden layer states. Irrelevant and redundant information will be discarded by using the forget gate of LSTM. In the final time step, all important information is aggregated and contributes to flood susceptibility analysis.

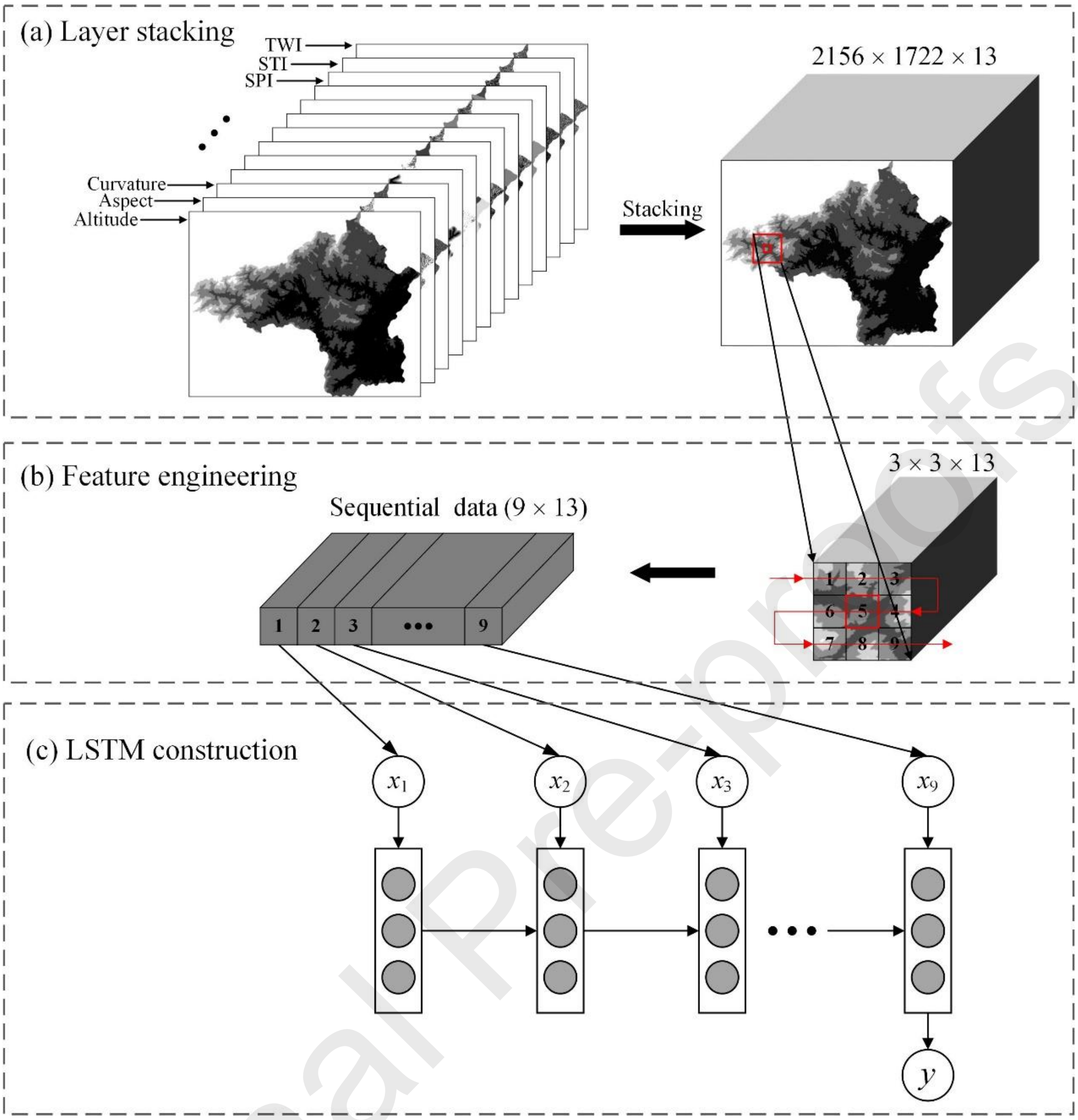


Fig. 5. Modelling process of LSS-LSTM. In panel (a), all the conditioning factors are stacked together to form a multi-band image. Then, in panel (b), each pixel and its neighboring pixels in a 3 × 3 window are extracted, and 9 pixel vectors are sorted into a sequential data based on spatial continuity. In panel (c), the sequential data is sent to LSTM network. *x* and *y* denote the input and output, respectively.

3.5. Model optimization

Over-fitting is a common problem in applying deep learning methods (Schmidhuber, 2015). Specifically, over-fitting occurs when learning models are so closely fitted to training data and causes a negative impact on predicting new data. In this study, two optimization techniques of data augmentation and batch normalization were used to solve the over-fitting problem during the modelling process.

3.5.1. Data augmentation

Since it is very difficult to obtain sufficient flood samples for model construction, the data augmentation technique is used to increase the number of training samples, which can improve the generalization ability of the prediction model. In general, the training set can be augmented by rotating and flipping each extracted image patch. **Fig. 6** shows seven types of transformation for an image patch with an uppercase letter F in the augmentation process. We can obtain 4 different samples by rotating the image patch with 90˚, 180˚ and 270˚, including the image patch itself, and the other four new samples are obtained by flipping the four image patches in the horizontal direction. Finally, the training set can be increased by 8 times. For example, the original training set contains 152 samples in this study. After the data augmentation procedure, the final training set for LSS-LSTM modelling includes 1216 samples.

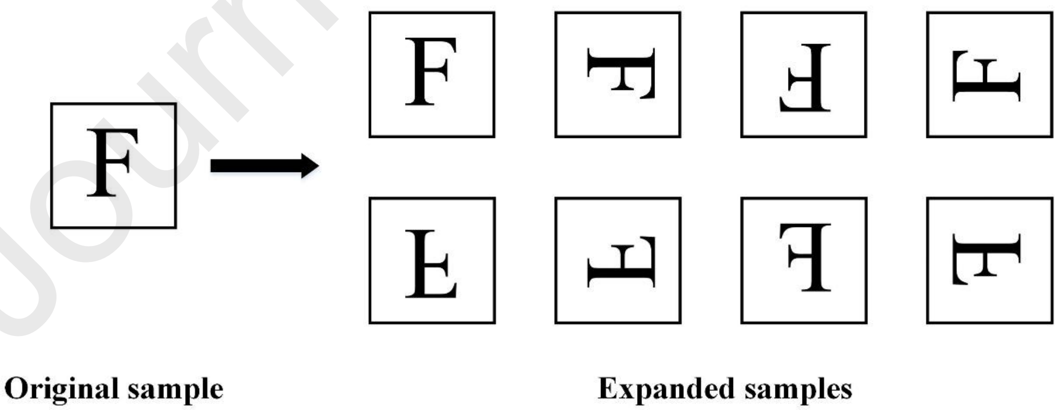


Fig. 6. Schematic diagram of data augmentation.

3.5.2. Batch normalization

Batch normalization can normalize the neural network layer by adjusting and scaling the activations, which can improve the generalization ability and convergence speed of the model (Ioffe and Szegedy, 2015). This technique is conducted to fix the mean and variance of layer input. During the training process of the LSS-LSTM method, the training data is sent to the neural network in batches. For one of the batches B that have m samples, the mean and variance are first calculated as follows:

$$\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{15}$$

$$\sigma^2 = \frac{1}{m}\sum_{i=1}^{m}\left(x_i - \mu_B\right)^2 \tag{16}$$

Then, each sample is normalized separately as follows:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}, \quad i = 1, 2, \dots, m \tag{17}$$

where ε is a very small constant for numerical stability. Next, the normalized values are scaled and shifted by a pair of parameters that are calculated as follows:

$$y_i = \gamma \hat{x}_i + \beta, \quad i = 1, 2, \dots, m \tag{18}$$

where γ and β are learnt in the subsequent optimization process. Finally, the output of the batch normalization transformation is passed to the neural network layer.

3.6. Model evaluation criteria

Model evaluation is a crucial step to assess the effectiveness of various FSP methods (Ahmadlou et al., 2018; Wang et al., 2019c). In this study, the receiver operating characteristic (ROC) curve that plots "sensitivity" on *y*-axis and "1-specificity" on *x*-axis and the area under ROC (AUC) is used for evaluation. The AUC value ranges from 0 to 1, and a higher value indicates a better model performance. Furthermore, several statistical criteria shown in **Table 2** were used to assess the performance of the FSP model as well.

Table 2 Statistical criteria for evaluation.

Evaluation criterion	Description & formula
True Positive (TP)	The number of flood samples that are correctly classified.
False Positive (FP)	The number of non-flood samples that are misclassified.
True Negative (TN)	The number of non-flood samples that are correctly classified.
False Negative (FN)	The number of non-flood samples that are misclassified.
Accuracy	$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$
Sensitivity	$Sensitivity = \frac{TP}{TP+FN}$
Specificity	$Specificity = \frac{TN}{TN+FP}$

4. Results and discussion

4.1. Relationship analysis between conditioning factors and flood occurrence

The VI values of different flood conditioning factors are shown in **Fig. 7**. All factors had a positive impact on the occurrence of floods. Specifically, the factors of slope and aspect achieved the highest and lowest VI values of 0.2929 and 0.0142, respectively. This is because river flood disasters often occur on flat terrain with low slopes. In addition, other studies confirmed the similar observations (Bui et al., 2020; Chen et al., 2019; Termeh et al., 2018).

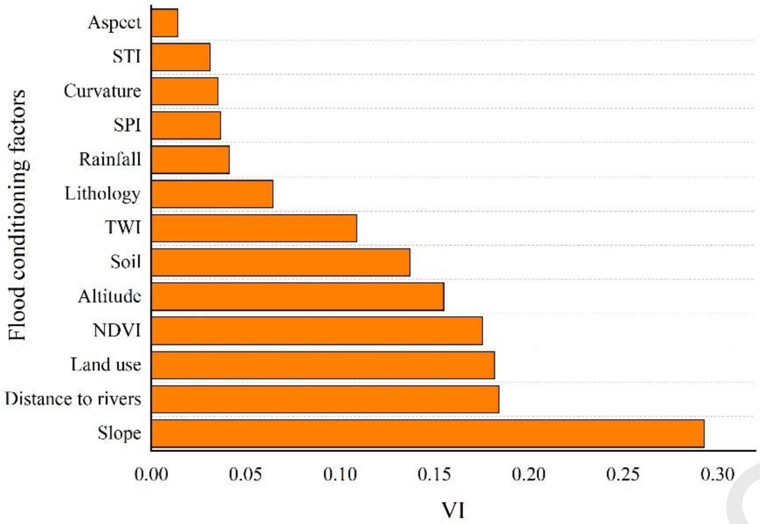
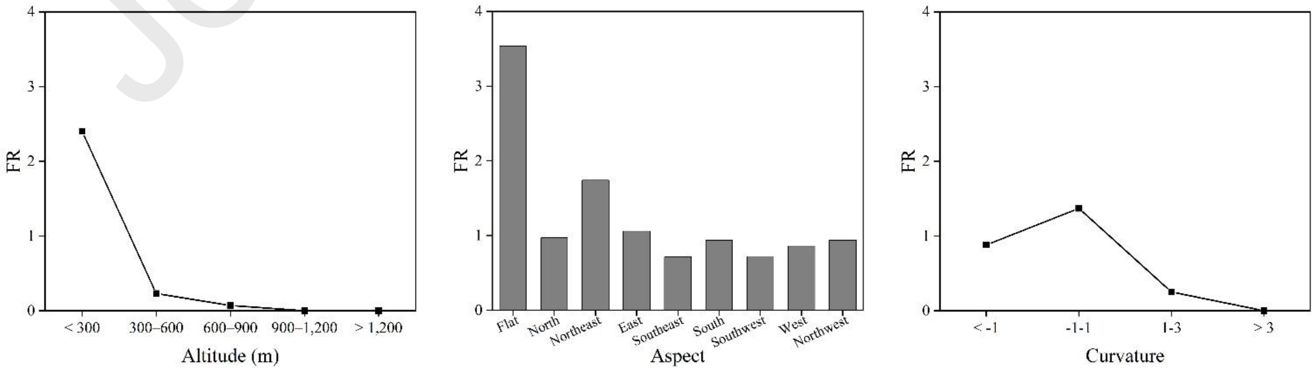


Fig. 7. Importance of flood conditioning factors using the IGR method. Factors with higher VI (variable importance) values are more important for prediction models, whereas factors with VI of zero indicate no contribution to floods.

To further explore the relationship between the conditioning factors and the occurrence of floods, we calculated the frequency ratio (FR) values of various factors. FR can measure the potential of flooding in the area corresponding to each category of a conditioning factor. The FR value is calculated by the ratio of percentage of floods to the percentage of domain in a specific class (Arabameri et al., 2019; Tehrany et al., 2015). The higher the FR value, the more prone to flooding in the corresponding areas (Termeh et al., 2018). **Fig. 8** shows the FR values of different conditioning factors (see detailed information in **Table A1** in *Appendix*). As for the altitude factor, areas with altitude lower than 300 m have higher FR values than other regions, and 89.81% of the historical floods occurred in the corresponding area. This observation is associated with our previous analysis that flat areas are more prone to floods. Furthermore, our results are relevant to other studies (Arabameri et al., 2019; Shafizadeh-Moghadam et al., 2018). In terms of the aspect factor, the flat class has the highest FR value of 3.54. However, since flat class contain very few grid cells and floods, this class is not instructive for floods. In addition, the FR values of aspect of other classes are closer, indicating the relationship between the aspect factor and flood occurrence is weak. This observation can also explain why the lowest VI value of aspect is achieved using the IGR method. Area with curvature value between -1 and 1 has the highest FR value. The corresponding area is determined to be a flat area, with 74.07% of the historical floods. The distance to river factor is key because rivers are the main channels for flood drainage and expansion (Bui et al., 2019d). From the **Fig. 8** we can find that the FR value decreases with increasing distance to rivers. In addition, FR values of the class with distance to river between 0 and 200 m are much higher than other classes, and the corresponding area accounts for 92.59% of the historical floods. This is why the distance to rivers factor is the second most important variable according to the results of the IGR method. Land use is another important factor for flood occurrence. Among these categories, the percentage of floods in grassland areas is the highest (83.33%), because grassland is usually located in flat areas. Moreover, no floods occur in forested areas because the forest area has strong water storage capacity and can mitigate flood disasters (Caprario and Finotti, 2019; Chapi et al., 2017). For lithology, Class I has the highest FR value of 1.61, while Class A has a FR value of 0 because the areas of Class A is mainly composed of dolomite, and its drainage density is lower, so the possibility of flooding is lower. For NDVI, the FR value decreases as the NDVI value increases. Higher NDVI values indicate better condition for vegetation growth . As a result, vegetated areas can store large amount of water and reduce the possibility of floods. Regarding the rainfall factor, all categories except the lowest rainfall class achieve high FR values. Other studied have also confirmed that heavy rainfall can increase the likelihood of floods (Giovannettone et al., 2018; Rahmati et al., 2016; Zhao et al., 2018). The areas with slope less than 10° has the highest FR value and other areas have near-zero FR values. It can be found that flat terrain is more prone to flooding. In the case of soil, the classes of ACh and ATc have higher FR values of 3.91 and 1.18 than other soil types. Several studies have confirmed that soils are related to the occurrence of floods because soil types directly determine soil permeability and structure (González-Arqueros et al., 2018; Tehrany et al., 2014; Tehrany et al., 2015). The class with SPI value larger than 200 has the highest FR value of 1.51, and the class with value between 25 and 50 has the lowest FR value of 0.32. For the STI factor, areas with SPI value less than 10 class has the highest FR value of 1.50, and the corresponding area accounts for 73.15% of the historical flood locations. This observation is consistent with previous studies that low STI a high flooding potential (Hong et al., 2018b; Tehrany et al., 2019). For the TWI factor, the FR value increases as the TWI value increases because higher TWI values indicate higher water accumulation levels, which is also consistent with other studies (Chapi et al., 2017; Shafizadeh-Moghadam et al., 2018; Tehrany et al., 2015).

In this study, we extracted flood conditioning factors from different sources and screened them based on the above relationship analysis. The quality of data sources is important for flood susceptibility modelling. The ASTER GDEM product is available for free to acquire, which have been widely used to extract flood conditioning factors in previous studies (Iosub et al., 2020; Khosravi et al., 2019; Tien Bui et al., 2019). Moreover, some advanced DEM products (such as shuttle radar topography mission (SRTM) DEM and multi-error-removed improved-terrain (MERIT) DEM) have higher accuracy than ASTER GDEM in hydrological analysis. Therefore, in future research, it is necessary to discuss the impact of different advanced DEM data for flood susceptibility modelling. In addition, we should note that there is a temporal mismatch between flood occurrence and its conditioning factors. Generally, lithology and soil can be regarded as constant factors and may not change over time. However, DEM, land use and NDVI will be affected by major changes seasonally within decades. Even within the same year, certain flood conditioning factors may change significantly. This is an important problem existed widely and is difficult to solve in FSP. Most researchers treated the flood conditioning factors as constant and ignore the temporal mismatch (Bui et al., 2019d; Chapi et al., 2017; Giovannettone et al., 2018). It should be noted that Roy *et al*. (2020) discussed the flood susceptibility results based on multi-temporal land use and rainfall factors. It will be necessary to dynamically analyze flood susceptibility with different temporal factors in the future.



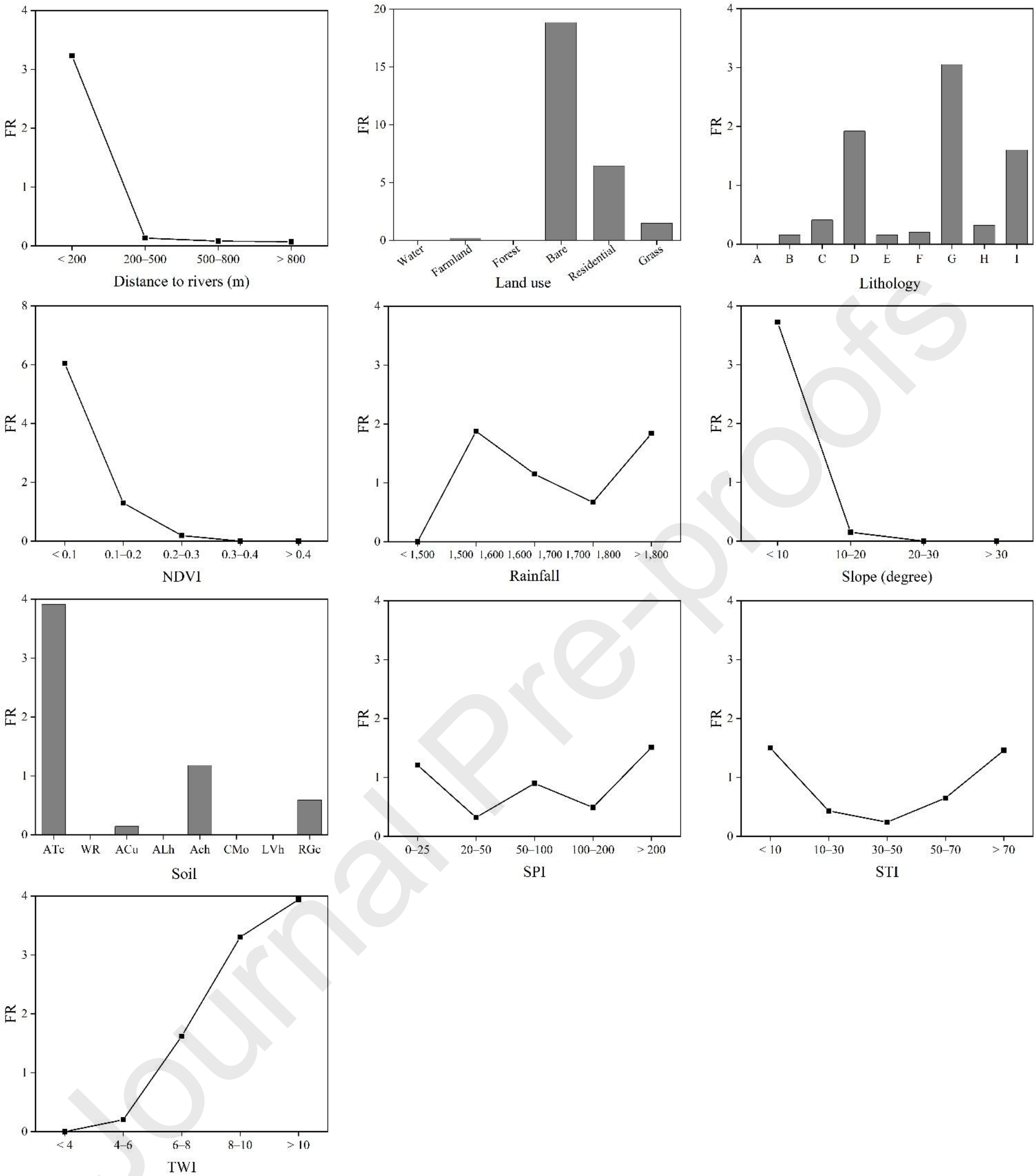


Fig. 8. Spatial relationship between conditioning factors and flood occurrence using frequency ratio (FR) model. A higher FR value indicates that the corresponding class is more prone to flooding.

4.2. Model performance

In the LSS-LSTM modelling process, the training strategy and hyperparameter setting have a significant impact on model performance. The purpose of the training process is to minimize the loss value and iteratively update the parameters using a specific optimization method. As shown in **Fig. 9**, the loss values of the training and validation data gradually decreased as the epochs increases, indicating a satisfactory training process. In this study, all the hyperparameters used in the LSS-LSTM method were optimized using the grid search method based on the five-fold cross-validation procedure, and **Table 3** lists the search space and the optimized results. The final network architecture of LSS-LSTM model contains one input layer, one hidden layer with 25 LSTM cells and one hidden layer. Moreover, the batch normalization layer was added before each activation function layer. All experiments were performed in Python under the framework of Keras[7] and Scikit-Learn[8].

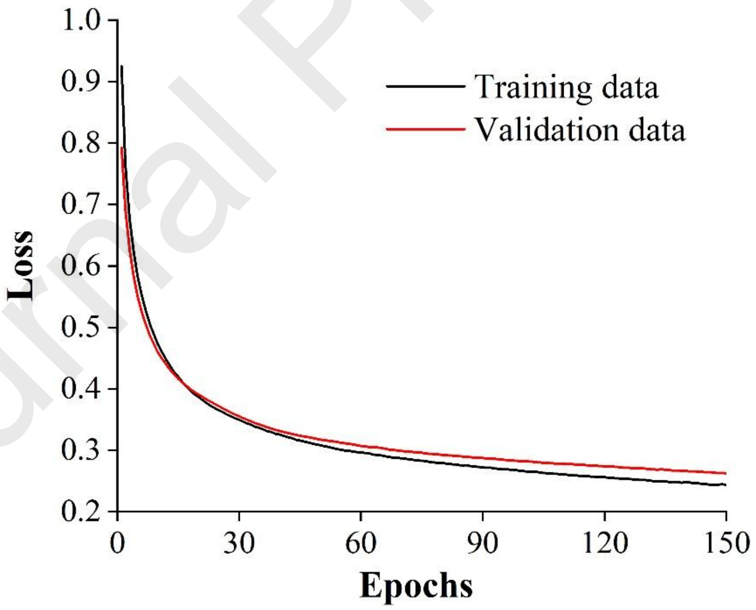


Fig. 9. The loss value variation during training process. The convergence of training and validation loss values to a lower level indicates a satisfactory training result.

[7] https://keras.io
[8] https://scikit-learn.org

Table 3 The search space and the optimized hyperparameters.

Hyperparameters	Search space	Search results	Description
Activation function	[ReLU, Tanh]	Tanh	Converting the linear relationships into nonlinear ones
Optimizer	[Adam, Adadelta, Adagrad, RMSprop]	Adagrad	Providing the direction to update the weights of the network
Learning rate	[0.001, 0.002, 0.005, 0.1, 0.2, 0.5, 1]	0.002	Controlling the learning speed of the model
Batch size	[100, 200, 300, 400, 500]	500	Number of samples processed by the neural network per iteration
Hidden size	[10, 15, 20, 25, 30]	25	Modulating the output size of hidden layers in the LSTM network

After the modelling construction, we used the LSS-LSTM method to predict flood susceptibility. **Table 4** presents the prediction performance. The LSS-LSTM method achieved an accuracy of 93.75%, which means that the method can effectively distinguish between flood samples and non-flood samples. In this experiment, flood susceptibility model is used to predict the probability of flood occurrence in a given area. Therefore, it is necessary to accurately predict the flood area and a higher accuracy value is crucial for our results. Moreover, another key point is that the flood model cannot miss any potential flood regions. This is because if we fail to find the area where the flood disaster may occur, prevention and management of this area may be ignored, which may cause devastating damage for society. Therefore, sensitivity is a remarkable index in the field of FSP. Results show that LSS-LSTM method achieved a very high sensitivity value of 96.67, indicating that the model has the ability to find almost all potential flood locations. **Fig. 10** shows the ROC curve using the test set. As claimed in previous studies (Arabameri et al., 2019; Kanani-Sadat et al., 2019), a AUC value larger than 0.9 is an excellent prediction result. Therefore, the proposed LSS-LSTM method achieved a relative good prediction performance with an AUC value of 0.965.

Table 4 Prediction performance of LSS-LSTM.

Method	Accuracy	Sensitivity	Specificity
LSS-LSTM	93.75%	96.67%	91.18%

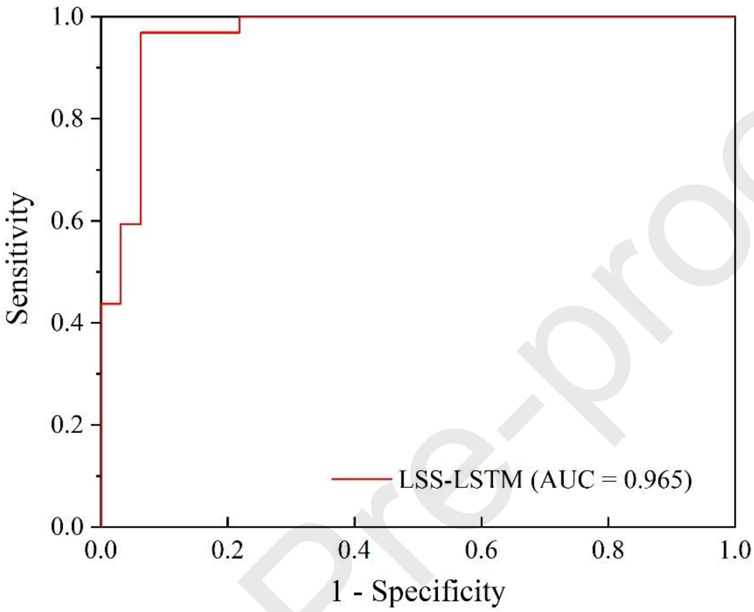


Fig. 10. ROC curve of the LSS-LSTM method.

In our experiments, we calculated the susceptibility indices of 1,714,419 grid cells in the study area to build the flood susceptibility map. All the susceptibility indices were sorted in ascending order and divided into five classes using the natural (Jenks) breaks method (Chapi et al., 2017; Shafizadeh-Moghadam et al., 2018; Tehrany et al., 2019). **Fig. 11** shows the final flood susceptibility map obtained by the proposed LSS-LSTM method. The majority of floods are located in the very high and high susceptible zones, and these very high susceptible areas are mainly located in the eastern and southern parts of the study area, which are low in altitude and slope and close to rivers. The sub-region (a) in the susceptibility map is the location of

Shangyoujiang Reservoir, and the sub-region (b) is close to the Shangyou River, which is the main tributary of the Gangjiang River system in the Yangtze River basin.

To quantitatively analyze the resultant susceptibility map, the flood density (FD) index, that is the percentage of flood pixels (PFP) divides by the percentage of susceptible class pixels (PSP), was used for evaluation. As shown in **Table 5**, the very high susceptibility class achieved the highest FD value of 13.39, followed by the classes of high (2.99), and moderate (0.29). Meanwhile, no flood occurred in the very low and low susceptible areas, indicating that the flood susceptibility map is reliable in the low susceptible area. It is also instructive for management, making it easier for people to focus on these areas with high susceptibility. In addition, the obtained trends of flood density distribution are completely consistent with several previous studies (Bui et al., 2020; Shafizadeh-Moghadam et al., 2018; Termeh et al., 2018).

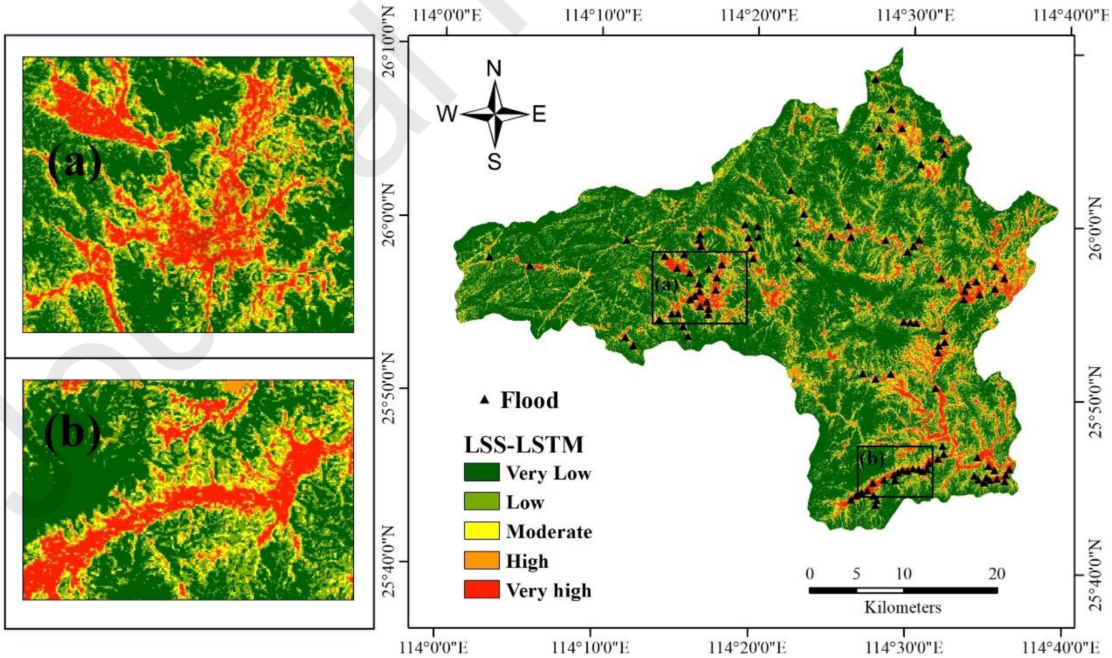


Fig. 11. Flood susceptibility map of the LSS-LSTM method. The continuous susceptibility values were reclassified into five susceptible classes using the natural (Jenks) breaks method.

Table 5 Flood density analysis of the flood susceptibility map by LSS-LSTM. The flood density (FD) is calculated by the ratio of the percentage of flood pixels (PFP) to the percentage of susceptible class pixels (PSP).

Susceptibility Class	No. of class pixels	PSP	No. of flood pixels	PFP	FD
Very low	965,654	59.70	0	0.00	0.00
Low	380,243	23.57	0	0.00	0.00
Moderate	161,723	6.37	3	2.78	0.29
High	106,032	3.97	20	18.52	2.99
Very high	100,767	6.39	85	78.70	13.39

4.3. Model sensitivity analysis

Generally, results of a robust flood susceptibility model should not change a lot if the input data changes with a reasonable range. To effectively demonstrate the prediction results of the LSS-LSTM method is universal rather than accidental, this study analysed two random manipulations occurred in LSS-LSTM modelling process to measure model sensitivity.

For the first random manipulation, we randomly selected the training and test sets for 10 times. Thus, a total of 10 LSS-LSTM models were constructed for evaluation. **Table 6** reports the results of several evaluation criteria for the proposed method performed 10 times. All the evaluation criteria demonstrate stable and reasonable fluctuations. For example, the mean and standard deviation (std) of AUC were 0.958 and 0.016, and those of ACC were 91.95% and 2.06%, respectively. In summary, the results indicate that the proposed method is not sensitive to the randomness of training/test splitting process and is robust to flood susceptibility analysis.

For the second random manipulation, we randomly changed the stacking order of conditioning factors for 10 times when all the conditioning factors are stacked together, as shown in **Fig. 5** (a). Results of LSS-LSTM performed 10 times is shown in **Table 7**. All the evaluation criteria have a reasonable fluctuations. For example, the mean and std of AUC were 0.958 and 0.011, and those of ACC were 92.35% and 1.30%, respectively. Compared to the results of LSS-LSTM with different training/test sets (**Table 6**), the results of LSS-LSTM with different factor stacking orders are less fluctuating. The phenomenon indicates that model performance is less sensitive to the variation of factors stacking order. This is because modelling samples contains different flood information, and each random splitting process may generate totally new training/test sets. However, when we change the stacking order of conditioning factors in LSS-LSTM modelling process, the total amount of information contained in the factors did not change. Furthermore, from the inherent structure of LSTM we can know that this network is not sensitive to factors order.

Table 6 Results of the LSS-LSTM method that carried out 10 times with different training/test sets.

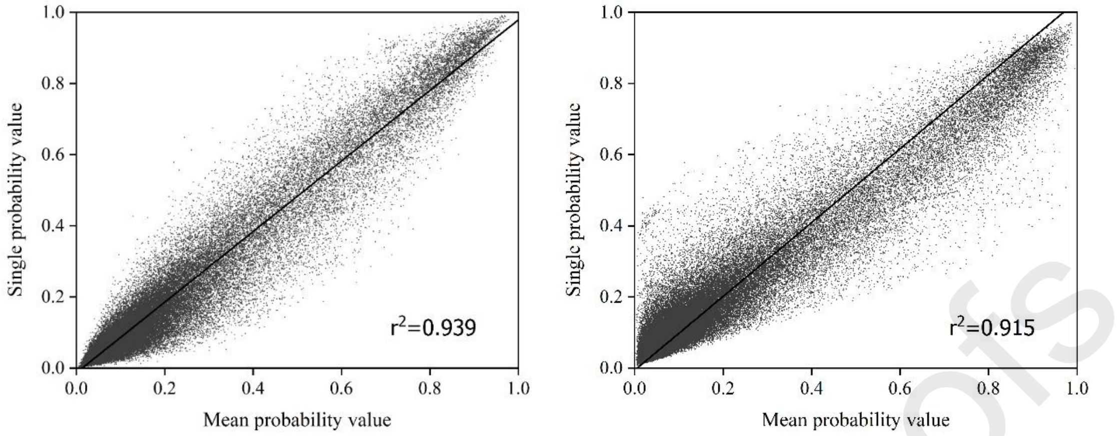
Evaluation criteria	Number of experiments										Statistics			
	1	2	3	4	5	6	7	8	9	10	Min	Max	Average	Std
AUC	0.965	0.955	0.946	0.985	0.946	0.976	0.959	0.976	0.938	0.938	0.938	0.985	0.958	0.016
Accuracy (%)	93.75	89.06	90.63	92.19	92.19	93.75	90.63	92.19	89.06	95.90	89.06	95.90	91.95	2.06
Sensitivity (%)	96.67	83.78	100	96.55	100	96.67	90.63	93.55	93.10	96.67	83.78	100	94.76	4.59
Specificity (%)	91.18	96.30	84.21	88.57	86.49	91.18	90.63	90.91	85.71	91.18	84.21	96.30	89.64	3.12

Table 7 Results of the LSS-LSTM method that carried out 10 times with different stacking orders of flood conditioning factors.

Evaluation criteria	Number of experiments										Statistics			
	1	2	3	4	5	6	7	8	9	10	Min	Max	Average	Std
AUC	0.965	0.941	0.955	0.974	0.963	0.966	0.947	0.965	0.940	0.962	0.940	0.974	0.958	0.011
Accuracy (%)	93.75	92.19	93.75	90.63	92.19	93.75	90.63	92.19	93.75	90.63	90.63	93.75	92.35	1.30
Sensitivity (%)	96.67	96.55	96.67	96.43	100	96.67	96.43	100	100	96.43	96.43	100	97.59	1.58
Specificity (%)	91.18	88.57	91.18	86.11	86.49	91.18	86.11	86.49	88.89	86.11	86.11	91.18	88.23	2.15

4.4. Model uncertainty analysis

To analyze the uncertainty of applying the LSS-LSTM method in FSP, we constructed two uncertainty scenarios: 10 susceptibility estimates obtained from LSS-LSTM method with different training/test sets and 10 susceptibility estimates obtained from the LSS-LSTM method with different factors stacking orders. For convenience, we named the above two scenarios as uncertainty scenario A and uncertainty scenario B, respectively. Due to the large amount of statistical calculation cost, for each uncertainty scenario, we first calculated the average of the 10 susceptibility estimates and arranged them in ascending order. Then, we selected 85720 grid cells from the average susceptibility estimates based on a systematic sampling with a periodic interval of 20, which can represent the susceptibility value distribution of the study area. A comparison between the single susceptibility estimate (the same as the **Fig. 11**) and the average susceptibility estimate is shown in **Fig. 12**. Guzzetti et al. (2006) compared the mean value of 50 susceptibility estimates and a single susceptibility estimate, which proved that the correlation between them is very high. Peng et al. (2014) performed a comparison between the average of 5 susceptibility estimates and a single susceptibility estimate, and obtained a high correlation ($r^2 = 0.909$) as well. As shown in **Fig. 12**, two uncertainty scenarios showed very high correlations ($r^2 = 0.939$ and $r^2 = 0.915$) between the single susceptibility estimate and the average susceptibility estimate, indicating the predicted susceptibility by the LSS-LSTM method is robust.



(a) (b)

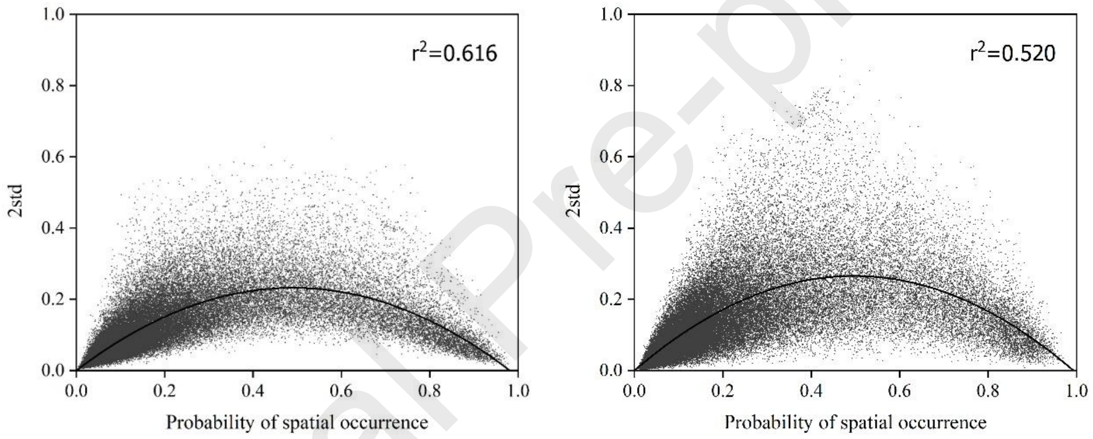
Fig. 12. Comparison between a single susceptibility estimate and the average susceptibility estimate. (a) Uncertainty scenario A. The average susceptibility estimate was calculated based on the 10 estimates derived from different training/test sets. (b) Uncertainty scenario B. The average susceptibility estimate was calculated based on 10 estimates derived from different factors stacking orders.

In addition, in order to quantify the uncertainty of flood prediction methods, we adopted a measure strategy proposed by Guzzetti et al. (2006). **Fig. 13** plots the mean susceptibility estimate on the *x*-axis against two standard deviations (2std) of the susceptibility estimate on the *y*-axis. For the two uncertainty scenarios, the 2std values increases from very low susceptibility to moderate and then decreases to very high susceptibility. Specifically, the 2std values are relatively low (< 0.35) for the low and high susceptibility zones, which indicates the LSS-LSTM method is capable of achieving stable predictions in these two susceptible zones. This conclusion is instructive and significant because it is necessary to accurately and stably predict flood locations. On the other hand, minimizing the problem of predicting potential flood areas as non-flooded areas is also important for further hazard management. Furthermore, the scatter distribution shown in **Fig. 13** is sparser for moderate susceptibility, indicating that the LSS-LSTM method is unable to stably predict whether a grid cell with moderate susceptibility is flood or non-flood. The variation in **Fig. 13** can be fitted by the following equations:

$$y = -0.964x^2 + 0.946x,\ 0 \le x \le 1,\ r^2 = 0.616 \tag{19}$$

$$y = -1.078x^2 + 1.07x,\ 0 \le x \le 1,\ r^2 = 0.520 \tag{20}$$

where x is the estimated susceptibility and y denotes the 2std value. In this way, Eq. (19) and Eq. (20) can be used to quantitatively assess the model uncertainty for each grid cell.



(a) (b)

Fig. 13. Mean susceptibility estimate (*x*-axis) against two standard deviations of the susceptibility estimate (*y*-axis). (a) Uncertainty scenario A. *x*-axis denotes the mean susceptibility estimate of 10 estimates obtained from different training/test sets. *y*-axis is the two standard deviations (2std) of the susceptibility estimate. (b) Uncertainty scenario B. *x*-axis denotes the mean susceptibility estimate of 10 estimates obtained from different factors stacking orders. *y*-axis is 2std of the susceptibility estimate.

4.5. Hyperparameters sensitivity analysis

It is crucial for flood susceptibility modelling to accurately set hyperparameters (Rijal et al., 2018; Santos et al., 2019), especially for constructing a deep learning neural network. In this subsection, we discussed the impact of three hyperparameters for flood susceptibility analysis.

In the first experiment, we analyzed the impact of batch normalization on flood susceptibility. The LSS-LSTM models optimized with and without batch normalization were compared. **Fig. 14** presents the training and validation accuracy variations of the two models during the training process. The LSS-LSTM method using batch normalization achieved higher training and validation accuracy values than those by the LSS-LSTM method without batch normalization, indicating that the prediction capability of LSS-LSTM can be effectively improved by using the batch normalization. This is because batch normalization can solve the problem of internal covariate shift existing in the training process (Ioffe and Szegedy, 2015). More specifically, the input distribution of each hidden layer in the LSS-LSTM method is transformed to a normal distribution, which can avoid the vanishing gradient problem and accelerate convergence. Hence, the purpose of improving accuracy and generalization can be achieved.

In the second experiment, we compared the LSS-LSTM models optimized with and without data augmentation. In **Fig. 15**, the LSS-LSTM method using data augmentation achieved higher training and validation accuracy values than those by the LSS-LSTM method without data augmentation. The data augmentation approach is a simple and convenient trick that artificially expands the size of the training set. Some publications have proven its effectiveness in improving generalization ability and application accuracy of deep learning methods in several fields (Han et al., 2018; Ma et al., 2019; Renda et al., 2019). In the field of flood susceptibility analysis, real flood samples are very limited and the collection of the flood samples is time-consuming, which may hinder the application of deep learning in this field. The results of applying the data augmentation technique in our study demonstrate its effectiveness in improving the prediction capability of LSS-LSTM, which can provide reference and help for other researchers in applying deep learning method for FSP.

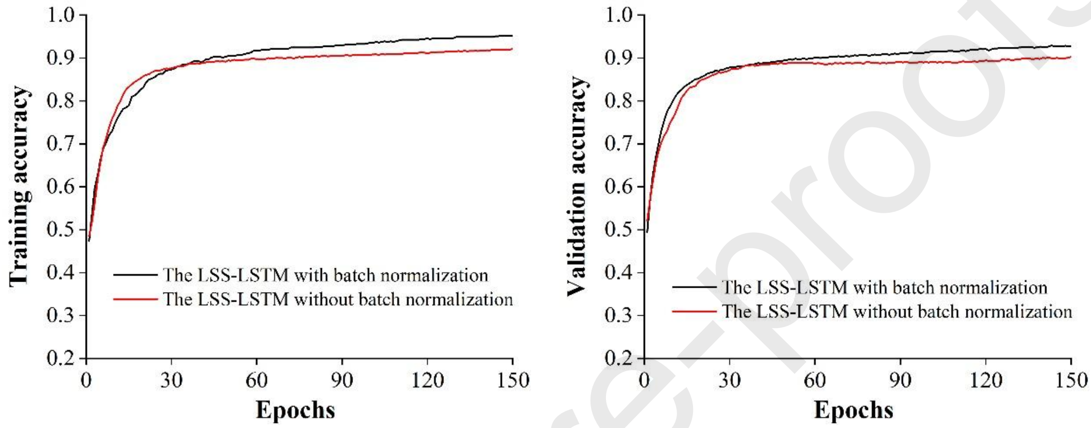


Fig. 14. Loss variation using the LSS-LSTM models with and without batch normalization.

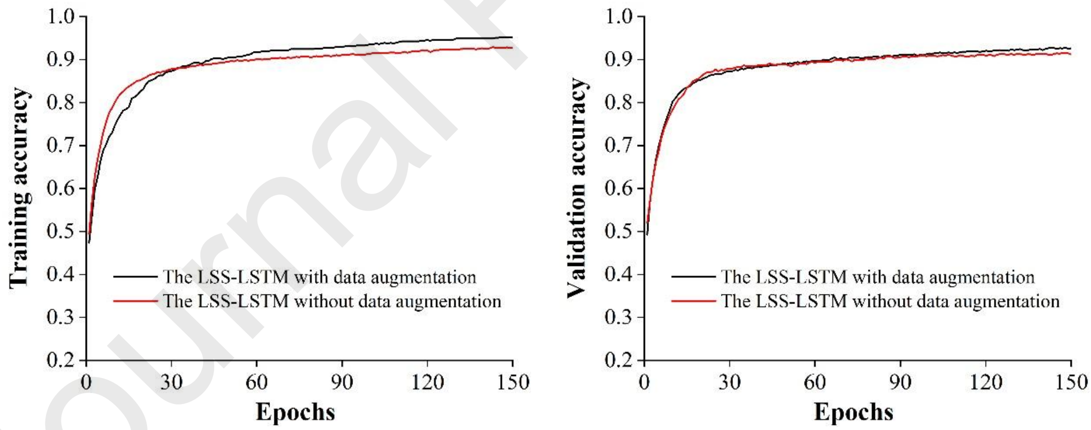


Fig. 15 Loss variation using the LSS-LSTM models with and without data augmentation.

In the third experiment, we analyzed the parameter of window size mentioned in section 3.2. It is an important hyperparameter in the LSS-LSTM modelling process. **Fig. 16** presents the AUC values of the proposed method as the number of window size is increased during the training process. The LSS-LSTM method obtained the highest AUC value when this parameter is set to 3. Meanwhile, since the larger window size may bring a lot of irrelevant and redundant information, the AUC value decreases as this parameter is increased from 5 to 11, which has a negative impact on flood susceptibility prediction.

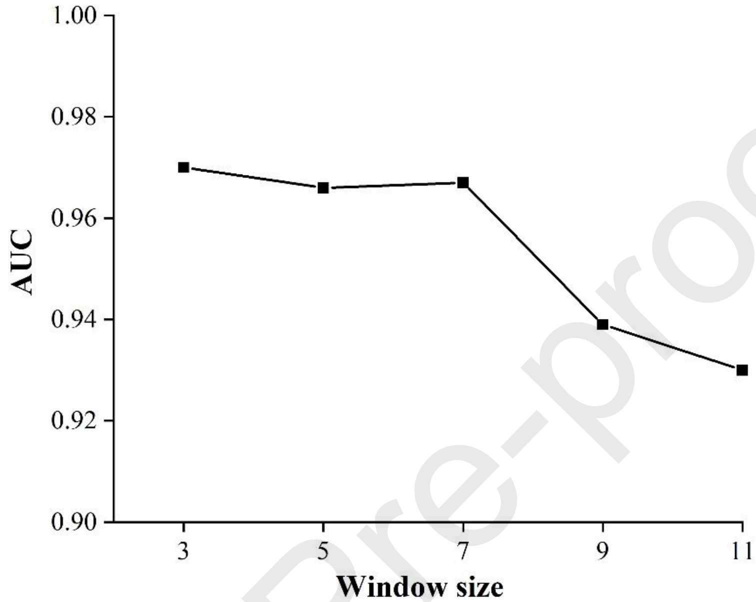


Fig. 16. AUC variation using the LSS-LSTM method with different window sizes.

4.6. Comparison with state-of-the-art techniques

To demonstrate the effectiveness of the LSS-LSTM method, we selected three benchmark deep learning techniques for comparison: regular deep neural network (DNN), one-dimensional convolutional neural network (1D-CNN) and three-dimensional convolutional neural network (3D-CNN). The input form of DNN and 1D-CNN used a commonly used one-dimensional vector-based method. The 3D-CNN extract factors information and local spatial information from window patches. The implementation details of DNN, 1D-CNN, and 3D-CNN can refer to several previous publications (Bui et al., 2020; Bui et al., 2019d; Wang et al., 2020a).

Note that the data augmentation method in Section 3.5.1 was used in 3D-CNN modelling process. **Table 8** presents the prediction accuracies of the four methods. The LSS-LSTM method achieved the highest accuracy value of 93.75%, followed by 3D-CNN (92.19%), 1D-CNN (90.63%), and DNN (89.06%). Moreover, 1D-CNN obtained the highest sensitivity value of 100%, followed by LSS-LSTM (96.67%) 3D-CNN (96.55%), and DNN (96.43%).

Table 8 Prediction accuracies of different methods.

Method	Accuracy	Sensitivity	Specificity
LSS-LSTM	93.75%	96.67%	91.18%
DNN	89.06%	96.43%	86.11%
1D-CNN	90.63%	100%	84.21%
3D-CNN	92.19%	96.55%	88.57%

The ROC curves of the three methods using the test set is shown in **Fig. 17**. The LSS-LSTM method had the highest AUC value of 0.965, followed by 3D-CNN (0.956), 1D-CNN (0.929), and DNN (0.917), indicating that the proposed method is superior to the other methods for flood prediction. In fact, to the best of our knowledge, the prediction ability of any model has its limitations due to different morphological and hydrological conditions of a certain study area, we cannot be sure that LSS-LSTM can always maintain its superiority in various inundated regions. However, the application mode of LSTM in flood susceptibility analysis can help other researchers to some extent. Moreover, the reasons why the LSS-LSTM method has the potential to portray exciting performance can be explained from the following three points. First, as mentioned in Section 3.3, LSTM is one of the powerful deep learning technique that has achieved reliable results in many fields (Graves and Jaitly, 2014; Mou et al., 2017; Sundermeyer et al., 2015; Zhang et al., 2018). Second, the occurrence of floods is not spatially independent and is closely related to neighboring terrain units. Third, the forget gate in LSTM can effectively filter out the useless information from the input data, which can improve the prediction capability of the model. Therefore, the LSS-LSTM method not only captures hidden information in flood conditioning factors, but also considers local spatial information in a specific sequence perspective. Furthermore, the proper selection of window size associated with the batch normalization and data augmentation techniques can further improve the prediction performance of the LSS-LSTM method, as mentioned in Section 4.4.

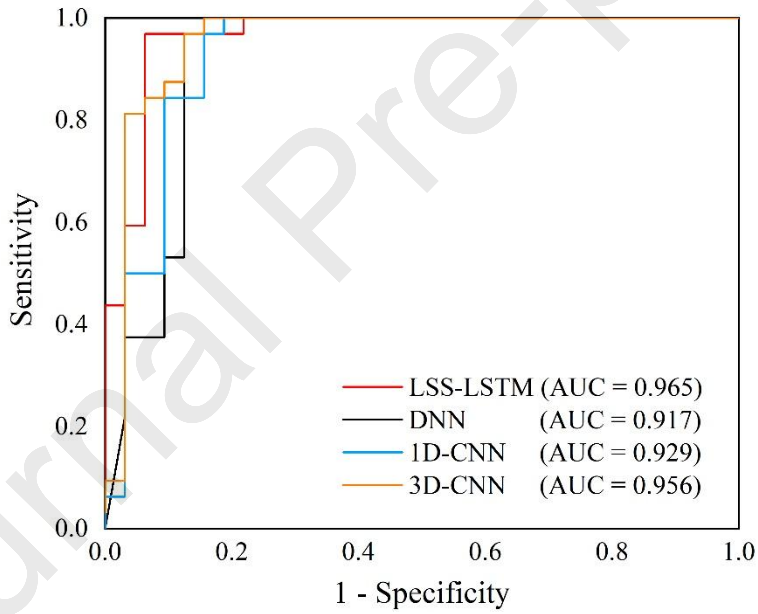


Fig. 17. The ROC curve of different methods.

Analyzing the impact of different training sample sizes is very important for measuring model performance (Schulz et al., 2020; Yang et al., 2020). To further demonstrate the superiority of the LSS-LSTM method to other methods, we conducted a sample sensitivity analysis using different training sample sizes for comparison. Different proportions ranging from 10% to 90% of the flood historical locations with a step size of 20% were randomly selected as training samples. **Fig. 18** plots the impact of different training samples on AUC for the study area. The LSS-LSTM method obtained a higher AUC value than the other three methods when the percentage of training data increases from 30% to 70%. The reason for this observation is that LSS-LSTM method can use sufficient data to learn the best fit function. Moreover, the optimization techniques of data augmentation and batch normalization can help avoid over-fitting and improve prediction capability. Note that the LSS-LSTM method cannot achieve higher performance than other three methods with 10% and 90% training sets. This is because 10% training samples cannot provide enough flood information for modelling. When the training samples occupy 90% of the total samples, the test sample size is too small. Therefore, that two scenarios cannot reliably and accurately reflect the prediction performance of models.

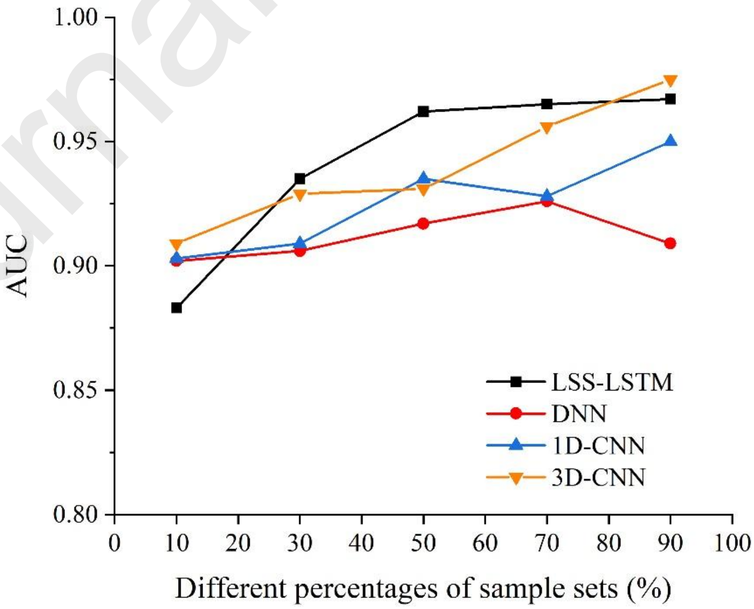


Fig. 18. AUC variation with different percentages of the training set.

5. Conclusions

In this study, we proposed a new LSS-LSTM method to obtain a reliable and accurate flood susceptibility map by integrating an appropriate feature engineering technique with the LSTM. The proposed method can retain the superior sequence modelling ability of LSTM and capture the local spatial information of floods. The main conclusions based on the experimental results can be summarized as follows. First, the proposed LSS-LSTM method achieved a satisfactory prediction performance with the accuracy and AUC values of 93.75% and 0.965, respectively. Second, the LSS-LSTM method is not sensitive to the randomness of training/test sets splitting process and the factors stacking order. Third, the LSS-LSTM method achieved better results than the benchmark methods of DNN, 1D-CNN, and 3D-CNN with several evaluation criteria. Finally, the prediction accuracy of LSS-LSTM can be effectively improved through the two manipulations of data augmentation and batch normalization. As a conclusion, the proposed LSS-LSTM method can be an inspiring alternative for decision-makers to prevent and mitigate flood hazards. In the future, our research will explore more representative feature engineering methods that accurately portray flood information for other state-of-the-art models.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (61271408, 41602362), the International Partnership Program of Chinese Academy of

Sciences (115242KYSB20170022) and the China Scholarship Council (201906860029). The authors would also like to thank the handling editors and the anonymous reviewers for their valuable comments and suggestions, which significantly improved the quality of this paper.

Appendix

Table A1 Detailed information of FR values of all the flood conditioning factors.

Factor	Class	No. of floods	Percentage of floods	No. of pixels in domain	Percentage of domain	FR
Altitude (m)	< 300	97	89.81	642,724	37.49	2.40
	300–600	10	9.26	691,735	40.35	0.23
	600–900	1	0.93	239,346	13.96	0.07
	900–1,200	0	0.00	101,286	5.91	0.00
	> 1,200	0	0.00	39,328	2.29	0.00
Aspect	Flat	2	1.85	8,969	0.52	3.54
	North	12	11.11	197,258	11.51	0.97
	Northeast	23	21.30	209,774	12.24	1.74
	East	15	13.89	224,526	13.10	1.06
	Southeast	12	11.11	267,607	15.61	0.71
	South	13	12.04	220,593	12.87	0.94
	Southwest	9	8.33	198,148	11.56	0.72
	West	10	9.26	185,559	10.82	0.86
	Northwest	12	11.11	201,985	11.78	0.94
Curvature	< -1	23	21.30	414,286	24.16	0.88
	-1–1	80	74.07	928,713	54.17	1.37
	1–3	5	4.63	312,336	18.22	0.25
	> 3	0	0.00	59,084	3.45	0.00
Distance to rivers (m)	0–200	100	92.59	491,345	28.66	3.23
	200–500	5	4.63	595,256	34.72	0.13
	500–800	2	1.85	401,932	23.44	0.08
	> 800	1	0.93	225,886	13.18	0.07
Land use	Water	0	0.00	38,042	2.22	0.00
	Farmland	1	0.93	108,600	6.33	0.15

	Forest	0	0.00	590,942	34.47	0.00
	Bare	12	11.11	10,103	0.59	18.85
	Residential	5	4.63	12,300	0.72	6.45
	Grass	90	83.33	954,432	55.67	1.50
Lithology	A	0	0.00	752	0.04	0.00
	B	1	0.93	98,015	5.72	0.16
	C	11	10.19	431,040	25.14	0.41
	D	12	11.11	99,104	5.78	1.92
	E	1	0.93	99,094	5.78	0.16
	F	2	1.85	161,751	9.43	0.20
	G	12	11.11	62,405	3.64	3.05
	H	2	1.85	99,490	5.80	0.32
	I	67	62.04	662,768	38.66	1.60
NDVI	< 0.1	73	67.59	191,936	11.20	6.04
	0.1–0.2	29	26.85	355,224	20.72	1.30
	0.2–0.3	6	5.56	510,347	29.77	0.19
	0.3–0.4	0	0.00	453,945	26.48	0.00
	> 0.4	0	0.00	202,967	11.84	0.00
Rainfall	< 1,500	0	0.00	11,799	0.69	0.00
	1,500–1,600	25	23.15	211,303	12.33	1.88
	1,600–1,700	37	34.26	511,137	29.81	1.15
	1,700–1,800	39	36.11	919,713	53.65	0.67
	> 1,800	7	6.48	60,467	3.53	1.84
Slope (°)	< 10	102	94.44	435,411	25.40	3.72
	10–20	6	5.56	631,952	36.86	0.15
	20–30	0	0.00	441,683	25.76	0.00
	> 30	0	0.00	205,373	11.98	0.00
Soil	ATc	27	25.00	109,518	6.39	3.91
	WR	0	0.00	31,973	1.86	0.00
	ACu	3	2.78	351,468	20.50	0.14
	ALh	0	0.00	132,386	7.72	0.00
	Ach	77	71.30	1,036,970	60.49	1.18
	CMo	0	0.00	450	0.03	0.00
	LVh	0	0.00	24,647	1.44	0.00
	RGc	1	0.93	27,007	1.58	0.59
SPI	0–25	66	61.11	867,514	50.60	1.21
	20–50	6	5.56	295,701	17.25	0.32
	50–100	12	11.11	211,872	12.36	0.90
	100–200	4	3.70	128,630	7.50	0.49
	> 200	20	18.52	210,702	12.29	1.51
STI	< 10	79	73.15	838,580	48.91	1.50
	10–30	16	14.81	594,586	34.68	0.43
	30–50	2	1.85	134,566	7.85	0.24
	50–70	2	1.85	49,108	2.86	0.65

	> 70	9	8.33	97,579	5.69	1.46
TWI	< 4	0	0.00	56,562	3.30	0.00
	4–6	13	12.04	1,016,532	59.29	0.20
	6–8	41	37.96	401,754	23.43	1.62
	8–10	28	25.93	134,892	7.87	3.30
	> 10	26	24.07	104,679	6.11	3.94

References

Ahmadlou, M., Karimi, M., Alizadeh, S., Shirzadi, A., Parvinnejhad, D., Shahabi, H., Panahi, M., 2018. Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA). Geocarto Int. 34, 1252-1272.

Arabameri, A., Rezaei, K., Cerdà, A., Conoscenti, C., Kalantari, Z., 2019. A comparison of statistical methods and multi-criteria decision making to map flood hazard susceptibility in Northern Iran. Sci. Total. Environ. 660, 443-458.

Bengio, Y., Simard, P., Frasconi, P., 1994. Learning long-term dependencies with gradient descent is difficult. IEEE. T. Neural. Netw. 5, 157-166.

BEVEN, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrol Sci J. 24, 43-69.

Billi, P., 2011. Flash flood sediment transport in a steep sand-bed ephemeral stream. Int. J. Sediment Res. 26, 193-209.

Bui, D.T., Hoang, N.-D., Martínez-Álvarez, F., Ngo, P.-T.T., Hoa, P.V., Pham, T.D., Samui, P., Costache, R., 2020. A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area. Sci. Total. Environ. 701, 134413.

Bui, D.T., Hoang, N.-D., Pham, T.-D., Ngo, P.-T.T., Hoa, P.V., Minh, N.Q., Tran, X.-T., Samui, P., 2019a. A new intelligence approach based on GIS-based Multivariate Adaptive Regression Splines and metaheuristic optimization for predicting flash flood susceptible areas at high-frequency tropical typhoon area. J. Hydrol. 575, 314-326.

Bui, D.T., Ngo, P.-T.T., Pham, T.D., Jaafari, A., Minh, N.Q., Hoa, P.V., Samui, P., 2019b. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. Catena. 179, 184-196.

Bui, D.T., Tsangaratos, P., Ngo, P.-T.T., Pham, T.D., Pham, B.T., 2019c. Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based

ensemble methods. Sci. Total. Environ. 668, 1038-1054.
Bui, Q.-T., Nguyen, Q.-H., Nguyen, X.L., Pham, V.D., Nguyen, H.D., Pham, V.-M., 2019d. Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping. J. Hydrol., 124379.
Campolo, M., Soldati, A., Andreussi, P., 2003. Artificial neural network approach to flood forecasting in the River Arno. Hydrol Sci J. 48, 381-398.
Caprario, J., Finotti, A.R., 2019. Socio-technological tool for mapping susceptibility to urban flooding. J. Hydrol. 574, 1152-1163.
Chapi, K., Singh, V.P., Shirzadi, A., Shahabi, H., Bui, D.T., Pham, B.T., Khosravi, K., 2017. A novel hybrid artificial intelligence approach for flood susceptibility assessment. Environ. Modell. Softw. 95, 229-245.
Chen, W., Hong, H., Li, S., Shahabi, H., Wang, Y., Wang, X., Ahmad, B.B., 2019. Flood susceptibility modelling using novel hybrid approach of Reduced-error pruning trees with Bagging and Random subspace ensembles. J. Hydrol. 575, 564-873.
Choi, K., Fazekas, G., Sandler, M., Cho, K., 2017. Convolutional recurrent neural networks for music classification, 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 2392-2396.
Choubin, B., Moradi, E., Golshan, M., Adamowski, J., Sajedi-Hosseini, F., Mosavi, A., 2019. An Ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. Sci. Total. Environ. 651, 2087-2096.
Collobert, R., Weston, J., 2008. A unified architecture for natural language processing: Deep neural networks with multitask learning, Proceedings of the 25th international conference on Machine learning. ACM, pp. 160-167.
Costache, R., Pham, Q.B., Avand, M., Linh, N.T.T., Vojtek, M., Vojteková, J., Lee, S., Khoi, D.N., Nhi, P.T.T., Dung, T.D., 2020a. Novel hybrid models between bivariate statistics, artificial neural networks and boosting algorithms for flood susceptibility assessment. J. Environ. Manage. 265, 110485.
Costache, R., Popa, M.C., Bui, D.T., Diaconu, D.C., Ciubotaru, N., Minea, G., Pham, Q.B., 2020b. Spatial predicting of flood potential areas using novel hybridizations of fuzzy decision-making, bivariate statistics, and machine learning. J. Hydrol. 585, 124808.
Diakakis, M., Mavroulis, S., Deligiannakis, G., 2012. Floods in Greece, a statistical and spatial approach. Nat. Hazards. 62, 485-500.
Dodangeh, E., Choubin, B., Eigdir, A.N., Nabipour, N., Panahi, M., Shamshirband, S., Mosavi, A., 2020. Integrated machine learning methods with resampling algorithms for flood susceptibility prediction. Sci. Total. Environ. 705, 135983.
Fang, Z., Wang, Y., Peng, L., Hong, H., 2020a. A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping. Int. J. Geogr. Inf. Sci., 1-27.
Fang, Z., Wang, Y., Peng, L., Hong, H., 2020b. Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. Comput. Geosci. 139, 104470.
Fuller, I.C., 2008. Geomorphic impacts of a 100-year flood : Kiwitea Stream, Manawatu catchment, New Zealand. Geomorphology. 98, 84-95.

Gebrehiwot, A., Hashemi-Beni, L., Thompson, G., Kordjamshidi, P., Langan, T.E., 2019. Deep Convolutional Neural Network for Flood Extent Mapping Using Unmanned Aerial Vehicles Data. Sensors. 19, 1486.

Giovannettone, J., Copenhaver, T., Burns, M., Choquette, S., 2018. A Statistical Approach to Mapping Flood Susceptibility in the Lower Connecticut River Valley Region. Water Resour. Res. 54, 7603-7618.

González-Arqueros, M.L., Mendoza, M.E., Bocco, G., Castillo, B.S., 2018. Flood susceptibility in rural settlements in remote zones: The case of a mountainous basin in the Sierra-Costa region of Michoacán, Mexico. J. Environ. Manage. 223, 685-693.

Graves, A., Jaitly, N., 2014. Towards end-to-end speech recognition with recurrent neural networks, International conference on machine learning, pp. 1764-1772.

Graves, A., Mohamed, A.-r., Hinton, G., 2013. Speech recognition with deep recurrent neural networks, 2013 IEEE international conference on acoustics, speech and signal processing. IEEE, pp. 6645-6649.

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., Galli, M., 2006. Estimating the quality of landslide susceptibility models. Geomorphology. 81, 166-184.

Han, D., Liu, Q., Fan, W., 2018. A new image classification method using CNN transfer learning and web data augmentation. Expert. Syst. Appl. 95, 43-56.

Heitmuller, F.T., Hudson, P.F., Asquith, W.H., 2015. Lithologic and hydrologic controls of mixed alluvial–bedrock channels in flood-prone fluvial systems: Bankfull and macrochannels in the Llano River watershed, central Texas, USA. Geomorphology. 232, 1-19.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural. Comput. 9, 1735-1780.

Hong, H., Panahi, M., Shirzadi, A., Ma, T., Liu, J., Zhu, A.-X., Chen, W., Kougias, I., Kazakis, N., 2018a. Flood susceptibility assessment in Hengfeng area coupling adaptive neuro-fuzzy inference system with genetic algorithm and differential evolution. Sci. Total. Environ. 621, 1124-1141.

Hong, H., Tsangaratos, P., Ilia, I., Liu, J., Zhu, A.-X., Chen, W., 2018b. Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China. Sci. Total. Environ. 625, 575-588.

Hu, G., Yang, Y., Yi, D., Kittler, J., Christmas, W., Li, S.Z., Hospedales, T., 2015. When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition, Proceedings of the IEEE international conference on computer vision workshops, pp. 142-150.

Huang, F., Zhang, J., Zhou, C., Wang, Y., Huang, J., Zhu, L., 2019. A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction. Landslides. 17, 1-13.

Huang, S., Li, J., Xu, M., 2012. Water surface variations monitoring and flood hazard analysis in Dongting Lake area using long-term Terra/MODIS data time series. Nat. Hazards. 62, 93-100.

Hudson, P.F., Kesel, R.H., 2000. Channel migration and meander-bend curvature in the lower Mississippi River prior to major human modification. Geology. 28, 531-534.

Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167.

Iosub, M., Minea, I., Chelariu, O.E., Ursu, A., 2020. Assessment of flash flood susceptibility potential in Moldavian Plain (Romania). J. Flood Risk Manag., e12588.

Kanani-Sadat, Y., Arabsheibani, R., Karimipour, F., Nasseri, M., 2019. A new approach to flood susceptibility assessment in data-scarce and ungauged regions based on GIS-based hybrid multi criteria decision-making method. J. Hydrol. 572, 17-31.

Khosravi, K., Pham, B.T., Chapi, K., Shirzadi, A., Shahabi, H., Revhaug, I., Prakash, I., Bui, D.T., 2018. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. Sci. Total. Environ. 627, 744-755.

Khosravi, K., Shahabi, H., Pham, B.T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H.-B., Gróf, G., Ho, H.L., 2019. A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. J. Hydrol. 573, 311-323.

Le, X.-H., Ho, H.V., Lee, G., Jung, S., 2019. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. Water. 11, 1387.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature. 521, 436.

Li, K., Wu, S., Dai, E., Xu, Z., 2012. Flood loss analysis and quantitative risk assessment in China. Nat. Hazards. 63, 737-760.

Liu, Z., Xu, W., Feng, J., Palaiahnakote, S., Lu, T., 2018. Context-aware attention LSTM network for flood prediction, 2018 24th International Conference on Pattern Recognition (ICPR). IEEE, pp. 1301-1306.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. Isprs J. Photogramm. 152, 166-177.

Ma, X., Tao, Z., Wang, Y., Yu, H., Wang, Y., 2015. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transport. Res. C-emer. 54, 187-197.

Mahmoud, S.H., Gan, T.Y., 2018. Urbanization and climate change implications in flood risk management: Developing an efficient decision support system for flood susceptibility mapping. Sci. Total. Environ. 636, 152-167.

Moore, I.D., Gessler, P., Nielsen, G., Peterson, G., 1993. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 57, 443-452.

Moore, I.D., Wilson, J.P., 1992. Length-slope factors for the Revised Universal Soil Loss Equation: Simplified method of estimation. J. Soil Water Conserv. 47, 423-428.

Mou, L., Ghamisi, P., Zhu, X.X., 2017. Deep recurrent neural networks for hyperspectral image classification. IEEE. T. Geosci. Remote. 55, 3639-3655.

Paola, J.D., Schowengerdt, R.A., 1995. A detailed comparison of backpropagation neural network and maximum-likelihood classifiers for urban land use classification. IEEE. T. Geosci. Remote. 33, 981-996.

Peng, L., Niu, R., Huang, B., Wu, X., Zhao, Y., Ye, R., 2014. Landslide susceptibility mapping based on rough set theory and support vector machines: A case of the Three Gorges area, China. Geomorphology. 204, 287-301.

Rahmati, O., Pourghasemi, H.R., Zeinivand, H., 2016. Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. Geocarto Int. 31, 42-70.

Renda, A., Barsacchi, M., Bechini, A., Marcelloni, F., 2019. Comparing ensemble strategies for deep learning: An application to facial expression recognition. Expert. Syst. Appl. 136, 1-11.

Rijal, S., Rimal, B., Sloan, S., 2018. Flood Hazard Mapping of a Rapidly Urbanizing City in the Foothills (Birendranagar, Surkhet) of Nepal. Land. 7, 60.

Roy, P., Pal, S.C., Chakrabortty, R., Chowdhuri, I., Malik, S., Das, B., 2020. Threats of climate and land use change on future flood susceptibility. J. Clean Prod., 122757.

Sak, H., Senior, A.W., Beaufays, F., 2014. Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling, Proceedings of the Annual Conference of International Speech Communication Association pp. 338-342.

Salvati, P., Bianchi, C., Rossi, M., Guzzetti, F., 2010. Societal landslide and flood risk in Italy. Nat. Hazards Earth Syst. Sci. 10, 465-483.

Sameen, M.I., Pradhan, B., Lee, S., 2019. Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. Catena. 186, 104249.

Sampson, C.C., Smith, A.M., Bates, P.D., Neal, J.C., Alfieri, L., Freer, J.E., 2015. A high-resolution global flood hazard model. Water Resour. Res. 51, 7358-7381.

Santos, P.P., Reis, E., Pereira, S., Santos, M., 2019. A flood susceptibility model at the national scale based on multicriteria analysis. Sci. Total. Environ. 667, 325-337.

Schmidhuber, J., 2015. Deep learning in neural networks: An overview. Neural Networks. 61, 85-117.

Schulz, M.-A., Yeo, B.T., Vogelstein, J.T., Mourao-Miranada, J., Kather, J.N., Kording, K., Richards, B., Bzdok, D., 2020. Different scaling of linear models and deep learning in UKBiobank brain images versus machine-learning datasets. Nat. Commun. 11, 1-15.

Shafizadeh-Moghadam, H., Valavi, R., Shahabi, H., Chapi, K., Shirzadi, A., 2018. Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. J. Environ. Manage. 217, 1-11.

Sundermeyer, M., Ney, H., Schlüter, R., 2015. From feedforward to recurrent LSTM neural networks for language modeling. IEEE-ACM Trans. Audio Speech Lang. 23, 517-529.

Tehrany, M.S., Jones, S., Shabani, F., 2019. Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. Catena. 175, 174-192.

Tehrany, M.S., Pradhan, B., Jebur, M.N., 2014. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. J. Hydrol. 512, 332-343.

Tehrany, M.S., Pradhan, B., Mansor, S., Ahmad, N., 2015. Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. Catena. 125, 91-101.

Termeh, S.V.R., Kornejady, A., Pourghasemi, H.R., Keesstra, S., 2018. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. Sci. Total. Environ. 615, 438-451.

Tien Bui, D., Khosravi, K., Shahabi, H., Daggupati, P., Adamowski, J.F., Melesse, A.M., Thai Pham, B., Pourghasemi, H.R., Mahmoudi, M., Bahrami, S., 2019. Flood spatial modeling in northern Iran using remote sensing and gis: A comparison between evidential belief functions and its ensemble with a multivariate logistic regression model. Remote Sens. 11, 1589.

Tiwari, V., Jain, P.K., Tandon, P., 2016. Product design concept evaluation using rough sets and VIKOR method. Adv. Eng. Inform. 30, 16-25.

Turner, C.R., Fuggetta, A., Lavazza, L., Wolf, A.L., 1999. A conceptual basis for feature engineering. J. Syst. Software. 49, 3-15.

Wang, Y., Fang, Z., Hong, H., 2019a. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. Sci. Total. Environ. 666, 975-993.

Wang, Y., Fang, Z., Hong, H., Peng, L., 2020a. Flood susceptibility mapping using convolutional neural network frameworks. J. Hydrol. 582, 124482.

Wang, Y., Fang, Z., Wang, M., Peng, L., Hong, H., 2020b. Comparative study of landslide susceptibility mapping with different recurrent neural networks. Comput. Geosci. 138, 104445.

Wang, Y., Hong, H., Chen, W., Li, S., Pamučar, D., Gigović, L., Drobnjak, S., Bui, D.T., Duan, H., 2019b. A Hybrid GIS Multi-Criteria Decision-Making Method for Flood Susceptibility Mapping at Shangyou, China. Remote Sens. 11, 62.

Wang, Y., Hong, H., Chen, W., Li, S., Panahi, M., Khosravi, K., Shirzadi, A., Shahabi, H., Panahi, S., Costache, R., 2019c. Flood susceptibility mapping in Dingnan County (China) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm. J. Environ. Manage. 247, 712-729.

Werner, M., Hunter, N., Bates, P., 2005. Identifiability of distributed floodplain roughness values in flood extent estimation. J. Hydrol. 314, 139-157.

Yang, F., Wanik, D.W., Cerrai, D., Bhuiyan, M.A.E., Anagnostou, E.N., 2020. Quantifying uncertainty in machine learning-based power outage prediction model training: A tool for sustainable storm restoration. Sustainability. 12, 1525.

Zazo, S., Rodríguez-Gonzálvez, P., Molina, J.-L., González-Aguilera, D., Agudelo-Ruiz, C., Hernández-López, D., 2018. Flood Hazard Assessment Supported by Reduced Cost Aerial Precision Photogrammetry. Remote Sens. 10, 1566.

Zhang, G., Wang, M., Liu, K., 2019. Forest Fire Susceptibility Modeling Using a Convolutional Neural Network for Yunnan Province of China. Int. J. Disaster Risk Sci. 10, 386-403.

Zhang, X., Sun, Y., Jiang, K., Li, C., Jiao, L., Zhou, H., 2018. Spatial sequential recurrent neural network for hyperspectral image classification. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 11, 4141-4155.

Zhao, G., Pang, B., Xu, Z., Peng, D., Xu, L., 2019. Assessment of urban flood susceptibility using semi-supervised machine learning model. Sci. Total. Environ. 659, 940-949.

Zhao, G., Pang, B., Xu, Z., Yue, J., Tu, T., 2018. Mapping flood susceptibility in mountainous areas on a national scale in China. Sci. Total. Environ. 615, 1133-1142.

Zheng, A., Casari, A., 2018. Feature engineering for machine learning: principles and techniques for data scientists. O'Reilly Media, Inc.

Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., Oliva, A., 2014. Learning deep features for scene recognition using places database, Advances in neural information processing systems, pp. 487-495.

Credit Author Statement

Zhice Fang, Data curation, Methodology, Validation; Visualization;

Roles/Writing - original draft

Yi Wang, Conceptualization; Formal analysis; Funding acquisition; Supervision; Writing - review & editing

Haoyuan Hong, Investigation; Resources

Ling Peng, Funding acquisition; Writing - review & editing

Highlights

- LSTM is considered for flood susceptibility prediction in a sequence perspective.
- An appropriate feature engineering method is integrated with the LSTM network.
- A reliable flood susceptibility map can be obtained by using the LSS-LSTM method.
- The proposed method can achieve better performance than benchmark methods.