Natural Hazards (2023) 115:331–355
https://doi.org/10.1007/s11069-022-05557-8

ORIGINAL PAPER

Ice-jam flood hazard risk assessment under simulated levee breaches using the random forest algorithm

Xiujie Wang[1,2] · **Zihua Qu**[1,2] · **Fuchang Tian**[1,2] · **Yanpeng Wang**[3] · **Ximin Yuan**[1,2] · **Kui Xu**[1,2]

Received: 15 February 2022 / Accepted: 4 August 2022 / Published online: 2 November 2022

Abstract
At higher latitudes, lower winter temperatures can cause ice jams to form in rivers, leading to levee breaches and significant economic losses, injuries, and deaths. This study examines the portion of the Inner Mongolia section of the Yellow River where ice-jam floods typically occur. Drawing on hazard system theory and a comprehensive analysis of hazard risk indices, including hazard-inducing factors and hazard-pregnant environments, as well as the vulnerability of the hazard-bearing bodies, we selected eight risk assessment indices to construct an ice-jam flood hazard risk assessment model that uses the random forest (RF) algorithm. The three hazard-inducing factors consisted of max water depth, max overbank flow velocity, and max inundation time, which were derived from the ice-jam flood backwater-levee break-inundation coupling model. The three hazard-vulnerable environment indices were elevation, terrain slope, and the distance to the river channel. The two hazard-bearing body indices were population density and Gross Domestic Product density. The modeling results show that compared with the K-nearest neighbor algorithm, the RF model performed better on both the Precision (P) and the area under the curve in assessing the four study areas. The RF model has significant advantages in classifying multi-dimensional ice-jam flood hazard data. It can provide support for ice-jam flood hazard prevention and mitigation.

Keywords The Yellow River's Inner Mongolia section · Ice-jam flood hazard risk assessment · Roughness · Hazard-inducing factors · Random forest model

✉ Fuchang Tian
tianfuchang@tju.edu.cn

1 State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin 300072, China

2 School of Civil Engineering, Tianjin University, Tianjin 300072, China

3 Production Management Department, Yangtze Ecology and Environment Co., Limited, Wuhan 430000, China

1 Introduction

Ice-jam floods commonly occur during the transitional periods of freezeup and breakup, which often extend for many kilometers along a river and can attain aggregate thicknesses of several meters. This can cause the river to exceed the high water level, overflowing levees and submerging villages, farmland, roads, and industrial and mining sites (Boucher et al. 2009; Wu et al. 2014). These events pose a serious threat to river facilities and the lives and property of people near the river (Ashton 1986; Kundzewicz et al. 2013a, b; Morse and Hicks 2005; Pham et al. 2021a). Therefore, managing the ice-jam flood hazard is important for securing people's livelihoods (Bouwer et al. 2010; Kundzewicz et al. 2013a, b). In this regard, the study of risk assessment of ice-jam flood hazards can inform the development of programs to prevent and mitigate ice-jam floods to reduce economic losses.

The ice-jam flood hazard system is complex and includes hazard-inducing factors, hazard-pregnant environments, and hazard-bearing bodies. It features high nonlinearity, spatial–temporal dynamics, and uncertainty, and the coupling of various challenges in the system may produce extremely complex phenomena (Tingsanchali and Karim 2010). Ice-jam flood hazard risk assessment involves a comprehensive evaluation of natural and social attributes to accurately capture the spatial distribution of ice-jam flood hazard risk (Osei et al. 2021), to identify variation in the degree of danger over different areas (Wu et al. 2014). Maps can be used to visualize the spatial distribution of the ice-jam flood hazard risk and to inform risk management, prevention, and transfer plans (Tian et al. 2021). Based on the principle of extreme value, Todorovic and Zelenhasic (1970) used the peaks over threshold (POT) model to illustrate the seasonal changes in flood risk. Anselmo et al. (1996) used hydraulic and hydrological coupling models to select a flood-prone area in Italy to assess flood risk. Zhou et al. (2000) put forward a model method that integrates rainfall, GDP, terrain, and other multi-index factors, and compared them with individual risk indices to analyze the rationality of the zoning results. Tan et al. (2004) considered factors such as flood inundation, socio-economic conditions, and hazard-pregnant environment to establish a county-level flood risk zoning model. Beltaos (2012) utilized the distributed function method (DFM) with peak flow data to assess the risks of ice blocking and flooding. Wang et al. (2013) established a flood risk model based on Particle Swarm Optimization (PSO) for the North River Basin of China. Based on onsite investigations, De Coste et al. (2017) simulated the combination of ice and water in the Hay River and its delta in the northwestern region of Canada to assess ice-jam flood risk. Due to the Yellow River's special geographical location (Beltaos 2012), river channel characteristics, the constraints of meteorological prediction accuracy as well as forecast period, current research cannot adequately address ice-jam flood hazards control there.

Although a considerable number of studies have provided a site-specific understanding of ice-jam flood risk indices, most of this research has focused only on one or two factors, which do not fully reflect the risks of ice-jam flood hazards. Given this, this article considers both the natural and the socio-economic indices as the ice-jam flood risk indices to ensure the results more comprehensive and reasonable, of which the hazard-inducing factors must be obtained through numerical simulations.

Numerical simulation of ice-jam flood hazards is the basis of risk analysis, management, and evaluation according to hydrodynamic methods, river ice dynamics methods, and theoretical methods. Lal and Shen (1991) proposed the one-dimensional river ice (RICE) model, which considered the distribution characteristics of water temperature and ice concentration.

Based on hydraulics, thermodynamics, and ice hydrodynamic theories, it simulates the process of ice transportation, blockage, and diving, and it is widely acknowledged as a pioneering accomplishment in the numerical simulation of ice-jam floods. Beltaos (1993) developed the RIVJAM river ice hydrodynamic model to simulate the process of water level changes during ice jams in wide, shallow channels. Jon and Ettema (2000) employed a numerical model to simulate the dynamic damage and reconstruction of ice jams and simulated the changing process of under-ice overcurrent based on the theory of hydrodynamics and ice transport. Hopkins and Tuthill (2002) applied the Digital Elevation Model (DEM) to test a non-curved river ice state system affected by a segmented ice boom. Yang et al. (2002) simulated the ice-jam flood congestion formation process and analyzed ice transport. Studying the dynamic formation process of ice cover, Shen et al. (2010) provided a dynamic simulation of ice transportation, blockage, and overcurrent under the ice, and then applied a full dynamic one-dimensional hydrodynamic model, including ice resistance and ice diffusion, to the St. John's River. Lindenschmidt et al. (2016) used Monte-Carlo simulation and other freezing numerical simulation methods to conduct a risk assessment of the ice-jam flood hazard and identify the vulnerability of cities and towns along the Peace River in Canada.

Existing research has been limited to a focus on freezing conditions, the evolution of ice transport, and the analysis of causes of ice-jam flood hazards. This work has established a series of static and dynamic river ice models, such as ICEJAM (Flato and Gerard 1986), RIVER1D (Hicks et al. 1992), DynaRICE (Shen et al. 2000), RIVICE (Lindenschmidt et al. 2012), RIVER2D (Brayall and Hicks 2012), and ICESIM. However, only a limited number of studies have considered backwater-levee break-inundation in a numerical model of ice-jam flooding and how it evolves in a floodplain. For example, Feng (2014) used one-dimensional river gates to block water to assess changes in backwater levels with ice jams but ignored the discharge capacity of the river after the ice-jam-prone, which is largely different from the actual ice-jam water evolution. Meanwhile, the backwater of the river ice jams have caused the cross-section wetted perimeter, the roughness, and the flow resistance to increase, which have not been reflected in previous ice-jam flood evolution models. Given these concerns, this study particularly proposes a comprehensive roughness optimization method for riverbeds. This method increases the roughness by setting the ice jam section to simulate the stagnation process caused by the upstream inflowing water blocked by the ice jam. To capture the characteristics of ice-jam flood evolution in a two-dimensional flood area, a comprehensive roughness optimization method for the flow ice surface layers is used to simulate the ice-jam flood evolution process.

Therefore, the main objectives of this study are: (1) to develop comprehensive and systematic indices for ice-jam flood risk assessment using RF of the Inner Mongolia section of the Yellow River. (2) to prove that RF is a suitable and reasonable method of ice-jam flood hazard risk assessment. (3) to analyze the ice-jam flood hazard risk distribution of the study basin. This will provide a novel opportunity to assess the usefulness of an existing ice-jam flood protection system, showing significant scientific and practical merits in terms of ice-jam flood hazard risk management and reduction of hazard in the Inner Mongolia section of the Yellow River and beyond.

2 Study area and data

2.1 Study area

The Yellow River is well-known for its high sediment load, frequent flooding, levee breaching, and channel migration (Fan et al. 2012). It originates from the Tibetan Plateau, flows eastwards through the Loess Plateau and the North China Plain, and flows into the Bohai Sea of the Pacific Ocean. The river can be divided into the upper, middle, and lower reaches based on distinctive geomorphologic and climatic conditions. The Inner Mongolia section of the Yellow River is located in the upper reach of the river in China, which stretches from the Wuda District in the west to Jungar Banner in the east. The total length of this section of the river section is 843 km. The river stages can be observed in Shizuishan, Bayangaole, Sanhuhekou, Toudaoguai. Among these locations, the section of the river running from Bayangaole to the northernmost head of the basin has the most frequent ice-jam flooding and suffers the most serious losses. As a result of its geographical location, the Inner Mongolia section of the Yellow River has the opposite stage of river closure and opening sequence. During the opening of the Yellow River in Inner Mongolia, flowing ice is blocked and accumulates, and ice jams are most likely to form at the river bends. Due to the ice jams, the overflow section is reduced and the upstream water level rises sharply. Generally, the water rises between 0.5 and 6 m due to the ice jam. After the water rises to a certain height, the ice jam succumbs to water pressure. When the ice jam breaks, the water level drops rapidly, sometimes as much as 1.5 m in 1 day. Due to the cold air in Siberia and Mongolia in winter and spring, the upper Yellow River Basin has a prevailing northerly wind, cold climate, and little rain. As a result, the mainstream and tributaries of the Yellow River have various degrees of ice-jam flooding in winter, especially the upper reaches in Ningxia Inner Mongolia (Ningmeng), which is a key ice-jam flood hazard controlling sections of the river. Considering its geographical location and the increase in rainfall intensity and frequency, winter ice-jam flood hazard conditions there are more severe and floods cause devastating damage to the local area (Das et al. 2020). To date, as a consequence of the air temperature and anthropogenic activities, such as deforestation, land-use pattern changes, migration, and riverbed siltation in the basin environment, the ice-jam flooding not only still occurs but has increased in frequency and severity (Pham et al. 2021b). From 1951 to 2010, ice-jam flooding occurred in the Inner Mongolia section of the Yellow River, causing significant losses of life and damage to property. This article focuses on this reach as the research area, because it has important practical significance for non-engineering measures, hazard prevention, and reduction management (Fig. 1).

2.2 Data sets

This study uses the data required by the coupled 1D–2D model to obtain the ice-jam flood hazard-inducing factors and the ice-jam flood risk assessment indices. This includes river data and cross-section data, which were obtained from the Municipal Water Authority of the Inner Mongolia Section of the Yellow River. Data on the boundary conditions included the upstream discharge and the downstream water level. The Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM), which has a 90 m resolution, provided data on elevation, water depth, overbank flow velocity, distance to the river channel, and

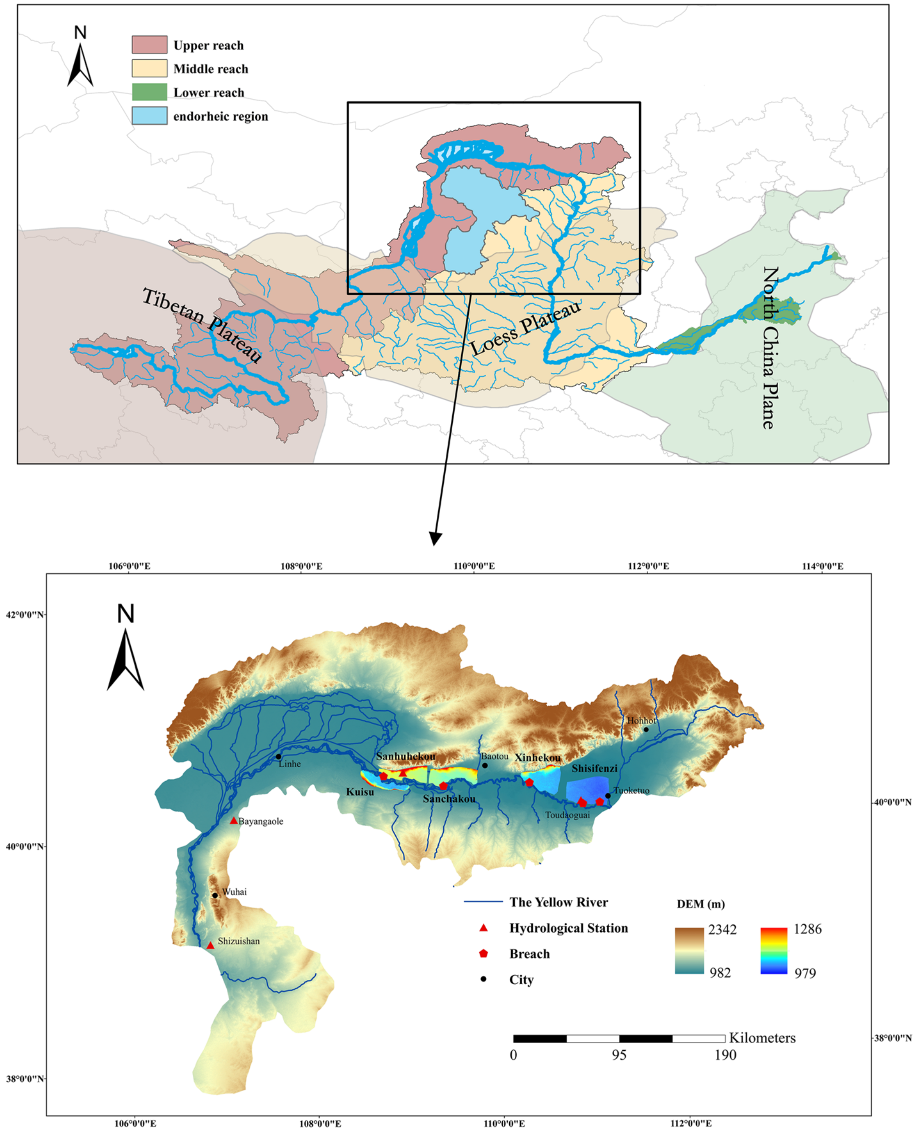

Fig. 1 Location of the Yellow River's Inner Mongolia section

the calculation of slopes and aspects. Population density and GDP density data were collected from China Statistical Information Network (2010).

3 Methodology and model construction

The framework developed in this study is presented as a flowchart in Fig. 2. Some of the most important features relevant to this research work are the following:

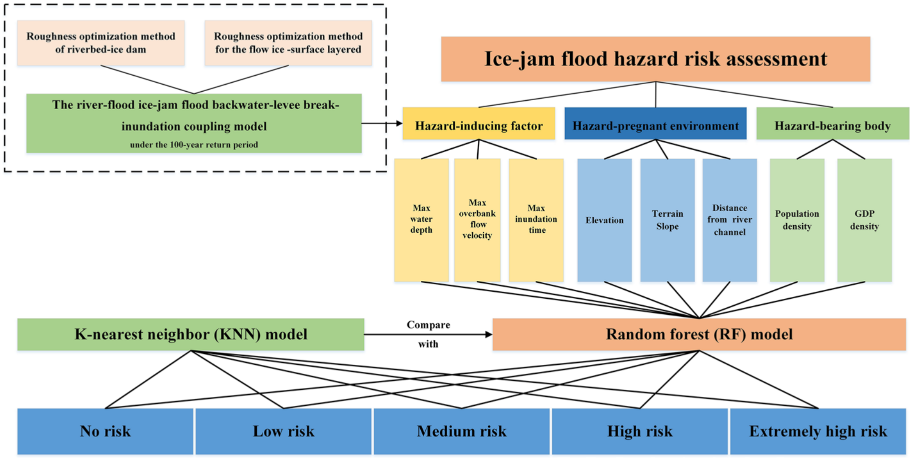

Fig. 2 Methodological flowchart adopted in the Yellow River's Inner Mongolia section

1. The ice-jam flood hazard risk assessment of the Inner Mongolia section of the Yellow River is conducted by the RF model, and KNN model is used for a comparison.
2. The hazard-inducing factors considered include the max water depth, max overbank flow velocity and max inundation time, which are derived from a physically based coupled 1D-2D hydrological model with a comprehensive roughness optimization method under the 100-year return period.
3. Elevation, terrain slope, and the distance to the river channel are selected as indices of the hazard-pregnant environment, while population density and GDP density as the socio-economic indices are selected as the hazard-bearing bodies.
4. The ice-jam flood hazard risk distribution of the study basin involves five hazard levels, including no risk, low risk, medium risk, high risk and extremely high risk, respectively.

Therefore, the overall framework of model construction in this study are: (1) obtaining ice-jam flood hazard-inducing factors by ice-jam flood backwater-levee break-inundation coupling model, (2) choosing comprehensive indices of hazard-pregnant environment and hazard-bearing body, and (3) developing a systematic procedure for ice-jam flood hazard risk assessment using RF, with KNN as a comparison. A detailed description of the above steps is given below.

3.1 Ice-jam flood backwater-levee break-inundation coupling model

The ice-jam flood backwater-levee break-inundation coupling model draws on the principles of hydrodynamics and considers the backwater characteristics of a one-dimensional river channel and the evolution characteristics of a two-dimensional ice-jam in a floodplain.

3.1.1 The 1D river flow models

The 1D river flow model is used for computing unsteady flow, discharge and water level in rivers and channels. It uses a one-dimensional, implicit, finite difference scheme for the numerical solution of the Saint–Venant equations and can be formulated as follows:

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = q \tag{1}$$

$$\frac{\partial Q}{\partial t} + \frac{\partial\left(\alpha \frac{Q^2}{A}\right)}{\partial x} + gA\frac{\partial Z}{\partial x} + \frac{gQ|Q|}{C^2AR} = 0 \tag{2}$$

$$v = \frac{1}{n}R^{2/3}J^{1/2} \tag{3}$$

$$n = \left(\frac{\chi_b n_b^{3/2} + \chi_i n_i^{3/2}}{\chi_b + \chi_i}\right)^{2/3} \tag{4}$$

For natural rivers, where generally $\chi_b \approx \chi_i$, it is defined as follows:

$$n = \left(\frac{n_b^{3/2} + n_i^{3/2}}{2}\right)^{2/3} \tag{5}$$

where Q is the discharge ($m^3 \cdot s^{-1}$); A is the cross-sectional area of the water (m^2); x is the distance along the river channel (m); t is the time (s); C is the Chezy coefficient ($s \cdot m^{-1/3}$); R is the hydraulic radius (m); q is the unit width discharge ($m^2 \cdot s^{-1}$); Z is the ice-jam flood variable water level (m); n is the comprehensive roughness; n_i is the roughness of the ice jam; n_b is river bed roughness; χ_i is the wetted perimeter of the ice jam (m); χ_b is the wetted perimeter of riverbed (m); α is the momentum correction coefficient. The roughness of the river section is increased by setting the ice jam to simulate the backwater process caused by the upstream water being blocked by the ice jam.

The computational grid consists of alternating Q-points and h points along the river (i.e., points where the discharge, Q, and water level, h, are computed at each time step). The model automatically generates a computational grid on the basis of the maximum distance, dx, defined as the distance between two adjacent h-points.

3.1.2 The 2D floodplain flow models

The floodplain flow model is two-dimensional mathematical model for the simulation of ice-jam flood flow and sediment transport in a floodplain. The hydrodynamic part of the models solves the vertically integrated Saint–Venant equations (continuity and conservation of momentum) in two directions. Considering the motion characteristics of ice-jam flood and the variation characteristics of surface topography in the floodplain, the ice-jam flood drag force and surface friction force are calculated by the optimization method of ice-jam flood and surface roughness. The two-dimensional numerical simulation control equation for the ice-jam flood is as follows (Cao et al. 2018; Mao et al. 2003):

$$\frac{\partial h}{\partial t}+\frac{\partial (hu)}{\partial x}+\frac{\partial (hv)}{\partial y}=hS \tag{6}$$

$$\frac{\partial u}{\partial t}+u\frac{\partial (u)}{\partial x}+v\frac{\partial (u)}{\partial y}+g\frac{\partial h}{\partial x}+g\frac{\partial z_b}{\partial x}+\frac{\tau_{ix}+\tau_{bx}}{\rho h}=0 \tag{7}$$

$$\frac{\partial v}{\partial t}+u\frac{\partial (v)}{\partial x}+v\frac{\partial (v)}{\partial y}+g\frac{\partial h}{\partial y}+g\frac{\partial z_b}{\partial y}+\frac{\tau_{iy}+\tau_{by}}{\rho h}=0 \tag{8}$$

$$\tau_{ix}+\tau_{bx}=\frac{\rho g\left(n_I^2+n_B^2\right)\sqrt{u^2+v^2}}{h^{1/3}}u \tag{9}$$

$$\tau_{iy}+\tau_{by}=\frac{\rho g\left(n_I^2+n_B^2\right)\sqrt{u^2+v^2}}{h^{1/3}}v \tag{10}$$

where h is the water depth (m); z_b is the elevation (m a.s.l.); τ_{ix} and τ_{iy} are the components of flow drag force in the x and y directions (Pa), respectively; u and v are the velocity components in the x and y directions, respectively (m·s^{-1}); τ_{bx} and τ_{by} are the components of surface friction in the x and y directions (Pa), respectively; n_I is the flow roughness; n_B is the surface roughness; and S is the magnitude of discharge of the point source (kg).

3.1.3 The 1D–2D coupling models

The simulation of ice-jam flood evolution is a dynamic coupling process of ice-jam flood backwater, levee breaking and inundation. In this paper, the real-time dynamic coupling of one-dimensional backwater model of river channel and two-dimensional numerical model of ice-jam flood evolution in floodplain is realized by connecting the side buildings of levee breaking. At any time of model calculation, the exchange of flow direction, discharge and momentum is determined by water level comparison between connecting grids, realizing the simulation of flow pattern change at levee breaking and inundation. The NWS DAMBRK method (Fread 1980) is used to calculate the real-time dynamic coupling of the river-floodplain ice-jam flood evolution model. The equations are as follows:

$$Q'=c_v k_s\left[c_w b\sqrt{g\left(h-h_b\right)}\left(h-h_b\right)+c_s S\sqrt{g\left(h-h_b\right)}\left(h-h_b\right)^2\right] \tag{11}$$

$$c_v = 1 + \frac{c_B Q_p^2}{gW_R^2 (h - h_d)^2 (h - h_b)} \tag{12}$$

$$k_s = \max\left[1 - 27.8\left(\frac{(h_{ds} - h_b)}{(h - h_b)} - 0.67\right)^3, 0\right] \tag{13}$$

where Q' is the bypass discharge of breach ($m^3 \cdot s^{-1}$); c_v is correction coefficient of inflow loss; k_s is inundation correction coefficient; b is the width of the breach (m); h is the water level (m) inside the breach; c_w is the horizontal weir coefficient of breach; h_b is the bottom elevation (m) of the breach; c_s is the breach slope weir coefficient; S is the breach slope; W_R is river width (m) at breach; h_d is final breach elevation (m); Q_p is the breach flow of the previous iteration ($m^3 \cdot s^{-1}$); h_{ds} is the water level (m) of the floodplain outside the breach; generally $c_w \approx 0.55$; $c_s \approx 0.43$; $c_B \approx 0.74$.

3.2 Risk assessment model of ice-jam flood hazard

In this study, a risk assessment model based on RF is adopted to evaluate regional ice-jam flood hazard, and the KNN model is used for risk assessment as a comparison, to prove RF a considerable advantage in solving ice-jam flood risk assessment.

3.2.1 Random forest (RF) model

The random forest (RF) algorithm (Leo 2001) is a commonly-used machine learning algorithm developed by Leo Breiman and Adele Cutler, which combines the output of multiple decision trees to reach a single result. The algorithm has been widely used in classification, regression, and unsupervised learning (Han et al. 2018). For multi-classification problems, using random sampling to form multiple classifiers can reduce errors and improve the ability to generalize from the algorithm. The operation can be highly parallelized, thereby

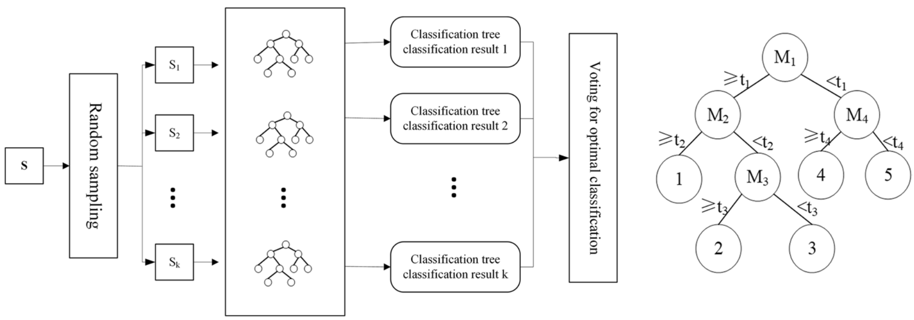

Fig. 3 Steps of RF decision tree generation

improving the computational efficiency of the model (Chen et al. 2017; Rodriguez-Galiano et al. 2012; Zhou et al. 2019).

The steps for generating the RF algorithm are shown in Fig. 3. First, k sub-training sets, $S_1, S_2, \ldots, S_K$, are randomly selected with bootstrap sampling to build K classification trees. At each node of the classification tree, m is randomly selected from n indices and the optimal segmentation index is chosen. These steps are repeated until k classification trees are traversed. K classification trees are clustered to construct the whole random forest.

When the RF algorithm is used to evaluate the risk level of the ice-jam flood hazards, the sample set to be predicted needs to be introduced into the trained RF classification tree. The risk level distributed on each leaf node is a result of the risk level division for the corresponding classification tree (Gounaridis et al. 2019). Data averaging is performed on the risk level classification results of all classification trees to obtain the entire ice-jam flood hazard risk zoning results using the RF algorithm.

$$p(c|\upsilon) = \sum_{t=i}^{T} P_t(c|\upsilon) \tag{14}$$

where T is the number of trees in the RF algorithm; c is a certain risk level; $p(c|\upsilon)$ is the probability of ice-jam flood hazard risk level c at the leaf node υ .

The RF algorithm has significant advantages in dealing with multi-index variable problems. The algorithm does not need to set index weights and does not perform pruning operations on classification trees (Mihăilescu et al. 2013). By gathering the voting results of multiple classification trees, multiple weak classifiers are combined to form a strong classifier. The accuracy of the RF algorithm is guaranteed, and the model also has a high tolerance for abnormal sample values, which avoids overfitting (Chen and Ishwaran 2012).

3.2.2 K-nearest neighbor (KNN) model

We also compared the application of the RF method to ice-jam flood hazard risk to another algorithm. The KNN model is a non-parametric pattern recognition and classification algorithm (Yang 2019). Due to the simplicity of its implementation, it has been applied in many fields, such as text classification (Lan et al. 2016), short-term water demand forecasting (Oliveira and Boccelli 2017), and annual average rainfall forecasting (Hu et al. 2013). The model measures the Euclidean distance between the sample to be tested and the known sample to predict the ice-jam flood hazard risk.

3.3 Accuracy evaluation index

We used the following statistical measures to validate the models: True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN), Accuracy (ACC), Kappa (K), Root Mean Squared Error (RMSE), and Reciever Operating Characteristic (ROC) curve. These methods were used to evaluate the performance of models to develop a reliable ice-jam flood susceptibility assessment.

Precision (P) and area under the curve (AUC) were used to evaluate the accuracy of the ice-jam flood hazard risk assessments using the RF algorithm and the KNN algorithm (Jiang et al. 2019; Li and Mao 2013). P is the ratio of positive samples predicted by the classifier to positive samples observed. AUC refers to the probability that the positive samples identified by the classifier are positive is greater than the probability that the negative samples identified by the classifier are positive. P reflects the accuracy and feasibility of the classifier algorithm applied to the ice-jam flood hazard risk assessment, and AUC reflects the relationship between the true classification rate and the false positive classification rate, to evaluate each classifier.

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \times 100\% \tag{15}$$

$$\mathrm{AUC} = \frac{\sum_{\mathrm{ins}_i \in \text{positive class}} \mathrm{rank}_{\mathrm{ins}_i} - \frac{M \times (M+1)}{2}}{M \times N} \tag{16}$$

where P refers to Precision; TP refers to the number of positive samples correctly predicted by the classifier; FP refers to the number of negative samples that are wrongly predicted as positive by the classifier; $\mathrm{rank}_{\mathrm{ins}_i}$ denotes the ith sample; M and N are the number of positive samples and negative samples, respectively.

3.4 Case model construction

3.4.1 Ice-jam flood numerical simulation construction

A 100-year return period for the Sanhuhekou to Toudaoguai section of the Yellow River is assumed. The model allows the dynamic simulation of the whole process of ice-jam flood backwater-levee break-inundation through the coupling connection at the breach. It calculates changes in hydraulic factors, such as the high water level change, outflow through the levee breach, ice flood inundation process, water depth, and overbank flow velocity.

(1) The 1D river flow models construction

Using data for 88 measured cross sections of the 316 km Sanhuhekou-Toudaoguai stretch of the river, the one-dimensional river flow hydrodynamic model was estimated. The inflow boundary, which is the designed ice-jam flood discharge process, is located at Sanhuhekou; while the outflow boundary, which is the relationship between water level and flow, is located at Toudaoguai.

(2) The 2D floodplain flow models construction

Taking the dangerous sections, historical levee breaches, and the presence of residential areas into account, Sanhuhekou (A), Sanchakou (B), Xinhekou (C), and Shisifenzi (D) were selected as study areas; each of these is located on the left bank of the river. The locations of the levee breaches are shown in Fig. 1. An unstructured grid is used to divide the terrain of the study area. Topographic data comes from National Aeronautics and Space Administration (NASA).

(3) The coupled 1D-2D models construction

The model captures the dynamic connection between the 1D river channel and the 2D floodplain of the ice-jam flood through the real-time coupling connection of the ice-jam flood at the breach. According to the historical ice-jam flood hazards, the width of the breach is set to 100 m. One breach is set in each of the study areas A, B, and C and two are set in study area D.

While the four study areas all lack historical ice-jam flood records, the Kuisu area on the right bank of the river experienced an ice-jam flood in 2008, and relevant data were recorded. On March 20, 2008, the Kuisu section of the Yellow River was affected by rapidly rising water levels at the Sanhuhekou hydrological station. Two breaches occurred in the morning. The east and west breaches were about 1 km apart and occurred one after another. The maximum width of the east and west breaches was 100 m and 60 m, respectively. The ice-jam flood caused the inundation of Duguitala and Hangjinnaoer towns. The inundation area totaled 126km^2 in size, and the direct economic losses reached RMB 935 million.

Following the previously described steps, a coupled 1D-2D model of the Kuisu area was created. To reflect the backwater-levee break-inundation evolution process caused by the ice-jam flood, a comprehensive roughness optimization method for riverbed ice jams was used to rate the river channel roughness. Referring to the *Specification for ice-jam flood computation SL428-2008*, the roughness of the river 10 km downstream of the breach in each study area was set to 0.0975. The widths of the east and west breaches were set to 100 m and 60 m, respectively, based on the data. The starting time for the model was set at 16:00:00 on March 18, 2008, and the ending time was set to 16:00:00 on March 28, 2008 in China. SRTM 90 m DEM Terrain was used to construct the two-dimensional ice-jam flood inundation analysis model of Kuisu, and an unstructured grid was used to divide the terrain of the study area. The maximum grid area was set at 0.01 km^2, with a total of 33,500 grids. The zoning roughness values for residential land and dry land were set at 0.08 and 0.04, respectively, to reflect the impact of the actual evolution process of the ice-jam flood. The

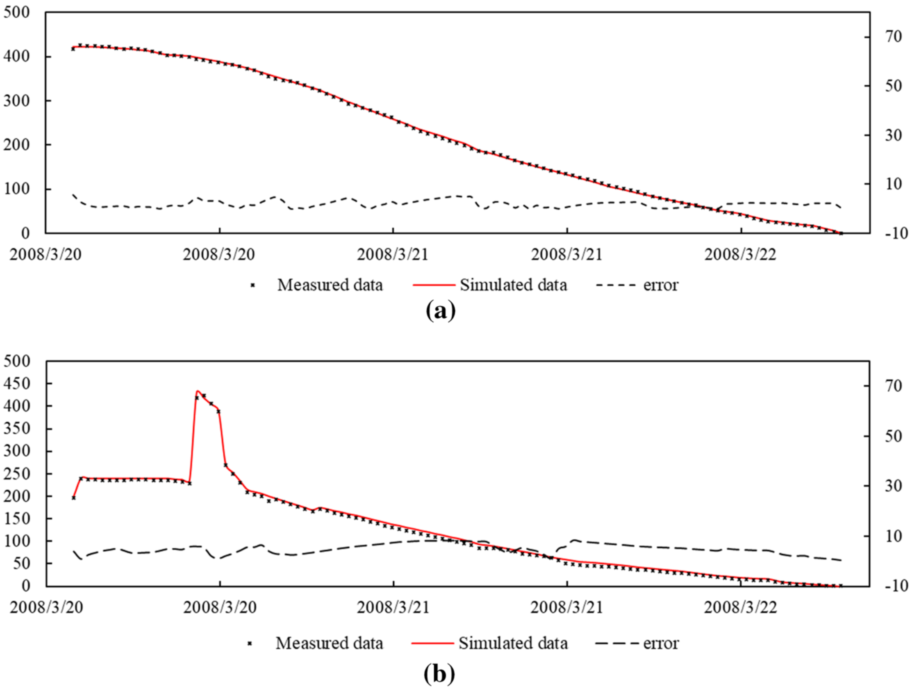

Fig. 4 Validation results of one-dimensional river ice-jam flood numerical model of Sanhuhekou-Toudaoguai: a) eastern breach in Kuisu; b) western breach in Kuisu

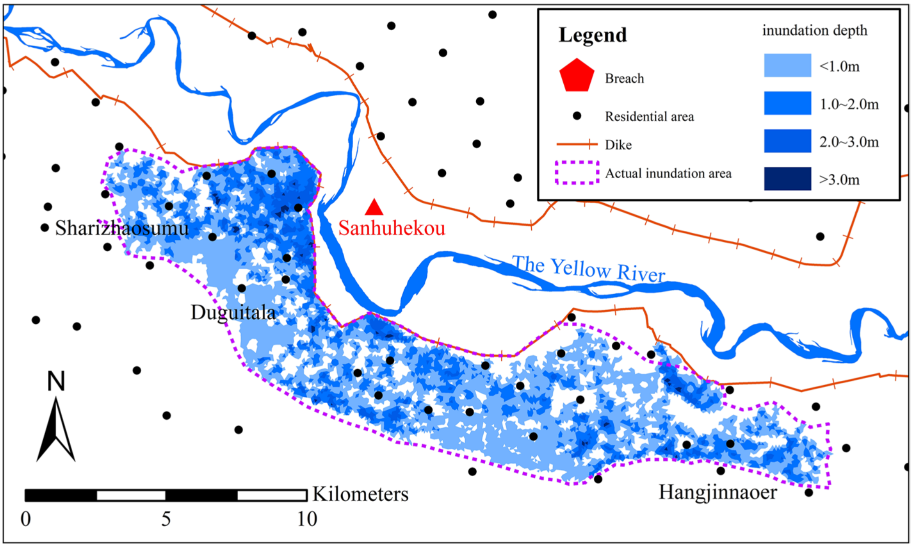

Fig. 5 Numerical simulation verification results of two-dimensional inundation area in Kuisu

initial water depth of the grid was set at 0.01 m; the dry water depth was set at 0.005 m; the wet water depth was set at 0.1 m.

In accordance with the Kuisu ice-jam flood coupled calculation model, the dynamic flow process at the east and west breaches and the two-dimensional maximum inundation area in the Kuisu area were extracted, and compared with the actual flow process of the breach and the actual inundation area (Fig. 4). The results show that the flow values estimated for the east and west breach of Kuisu deviated from the observed values by less than 5%. The inundation area predicted by the model matched the actual inundation area, and it was consistent with the historical data of the high risk areas, including Duguitala, Hangjinnaoer, and other townships. The total inundation area was 111.51km^2 (Fig. 5). The coupled 1D–2D simulation model of ice-jam flood generated accurate estimates, making it a useful simulation for hazard risk assessment.

3.4.2 Ice-jam flood hazards risk assessment indices system construction

Ice-jam flood hazard risk assessment involves a comprehensive evaluation of natural and social attributes to accurately capture the spatial distribution of ice-jam flood hazard risk. The ice-jam flood hazard system is complex and includes hazard-inducing factor, hazard-pregnant environment, and hazard-bearing body. After considering the actual conditions of ice-jam floods and relevant characteristics in the study area and reviewing recommendations provided by previous research, three ice-jam flood hazard-inducing factors, three ice-jam flood hazard-pregnant environment indices and two ice-jam flood hazard-bearing body indices are selected as follows.

(1) Ice-jam flood hazard-inducing factor.

Using the model parameters estimated for the Kuisu section, the 100-year return period ice-jam flood in the four study areas of A, B, C, and D were simulated and

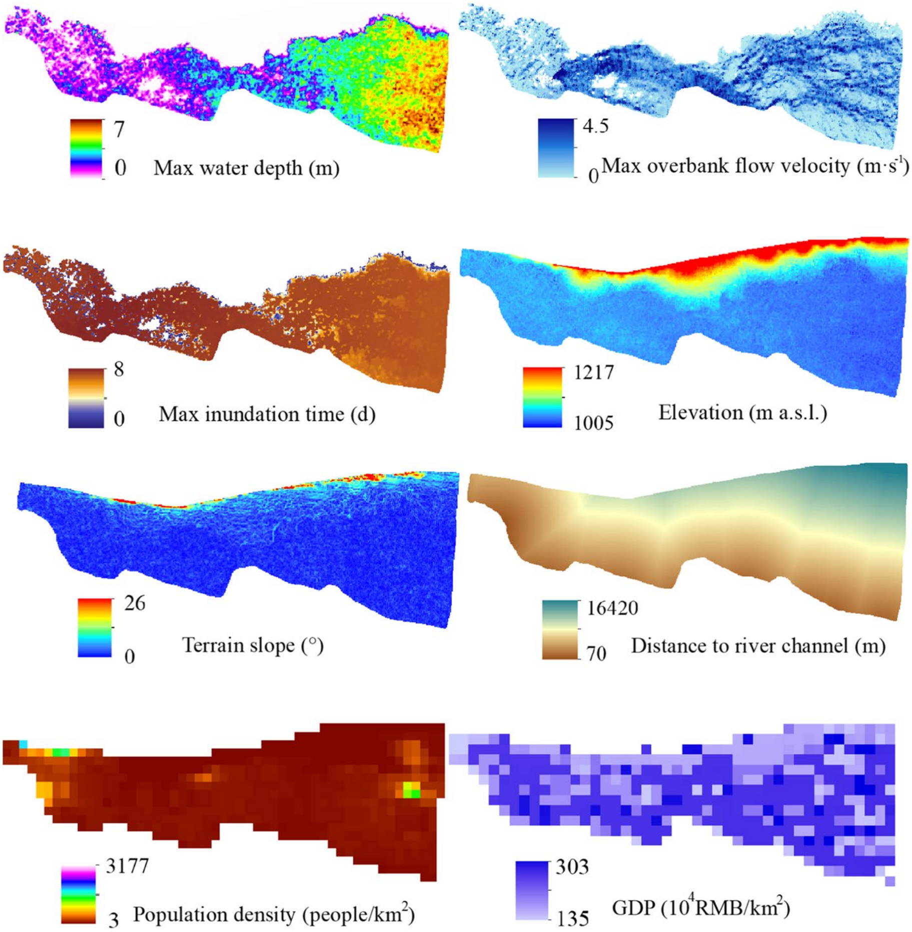

Fig. 6 Assessment indices for hazard risk in area A

calculated, and the max water depth, max overbank flow velocity, and max inundation time were extracted, as shown in Figs. 6, 7, 8, and 9, respectively.

According to the analysis of three ice-jam flood hazard-inducing factors, the portion of area A with the max water depth is located in the river bend and low-lying downstream areas, while the inundation area is distributed along the potential of the stream, and the inundation distribution range is relatively narrow and long. By contrast, the max water depths are concentrated at the breach and near the river channel in the area, at the breach and the low-lying area downstream in area C, and in the low and flat area in the middle of area D. In addition, due to the flat terrain of the floodplain on the north bank and the surface roughness of the flow, the max overbank flow velocity in area A is concentrated in the southern river bend and the upstream area. The corresponding max inundation time is long, indicating that this area should be the focus of ice-jam flood hazard prevention. In contrast, the max overbank flow velocity in area B is located in the area close to the channel along the potential of the stream, and the inundation time is lengthy, which indicates that the distance to the river channel is an important

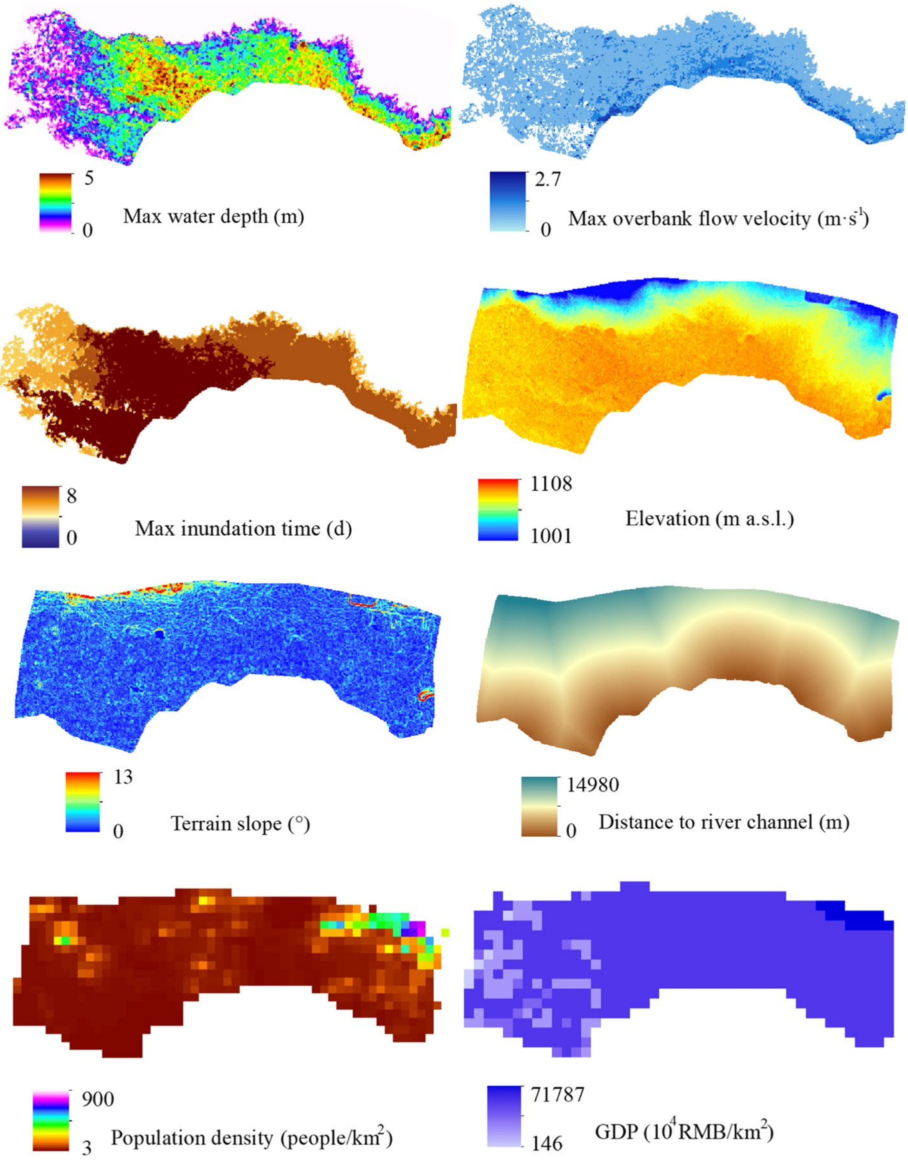

Fig. 7 Assessment indices for hazard risk in area B

determinant of ice-jam flood hazard risk. Compared to the northeast, which has a larger inundation area and greater water depths, area C features slightly higher terrain and the max overbank flow velocity is concentrated in the central area; this indicates that terrain level is an important factor in ice-jam flood risk. The max overbank flow velocity and max water depth in area D are located in the central region, where the terrain is low and flat. Because of the large resident population, the threat to life and property due to ice-jam floods is significant.

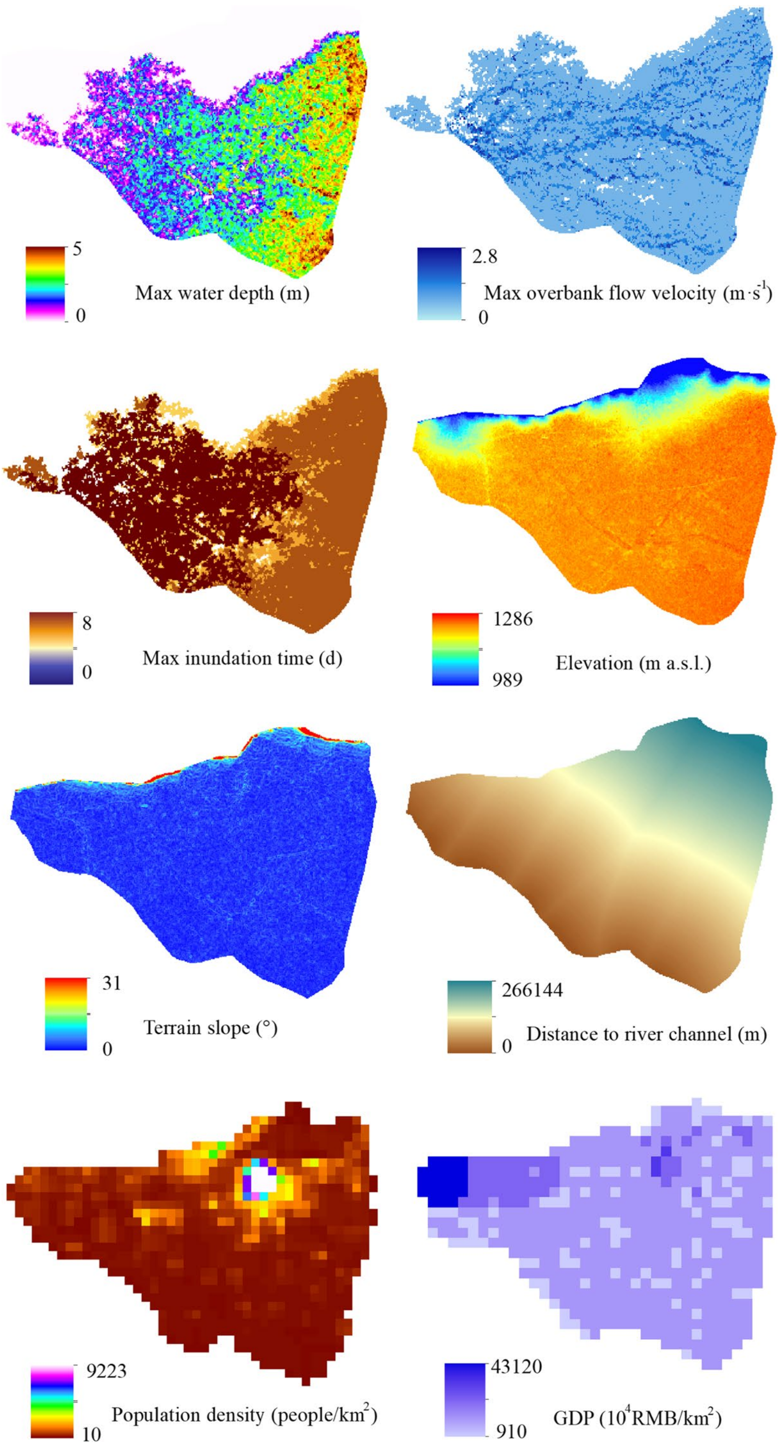

Fig. 8 Assessment indices for hazard risk in area C

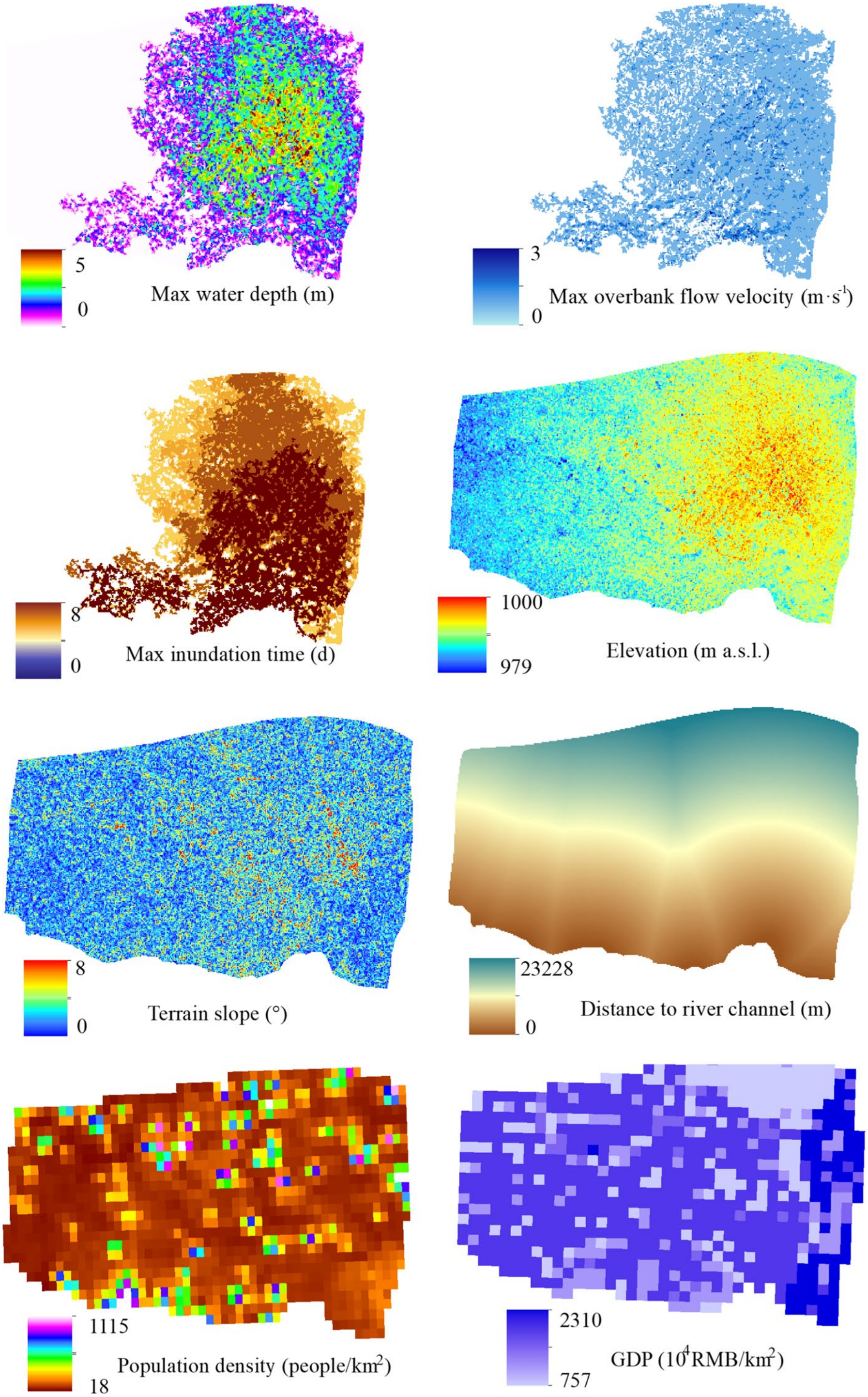

Fig. 9 Assessment indices for hazard risk in area D

(2) Ice-jam flood hazard-pregnant environment.

The term hazard-pregnant environment refers to the external environmental conditions, such as topography, the water system, and vegetation distribution. The hazard-pregnant environment concept mainly captures the impact of terrain and the water system on the formation of ice-jam floods. Areas with lower elevations and smaller topographic reliefs are more prone to ice-jam flood hazards. Areas with dense river networks where are closer to the water body also face greater risks of ice-jam flood hazards (Bhuiyan and Baky 2014). Therefore, in this study, elevation, terrain slope, and distance to the river channel (Penning-Rowsell et al. 2005) are selected as indices of a hazard-pregnant environment, as shown in Figs. 6, 7, 8, and 9.

(3) Ice-jam flood hazard-bearing body.

The risk associated with hazard-inducing factors reflects the potential damage caused by ice-jam floods. The actual severity of hazard is also related to the characteristics of the hazard-bearing body. Ice-jam floods of similar intensity have stronger effects in densely populated (Zou et al. 2012) and economically developed areas than in sparsely populated and economically backward areas (Winsemius et al. 2015). Therefore, this study uses population density and GDP density as indices of the hazard-bearing bodies, as shown in Figs. 6, 7, 8, and 9.

(4) Flood hazard risk assessment based on RF

The model accounts for the ice-jam flooding center, topography, and historical ice-jam flood hazard along the north bank of Inner Mongolia section. As a result, a relatively comprehensive ice-jam flood risk assessment indices system is constructed based on the hazard-inducing factors, hazard-pregnant environment, and hazard-bearing bodies. Using the ArcGIS platform and Python language, the RF model is used to carry out the risk assessment of ice-jam flood hazards for the Inner Mongolia section of the Yellow River. The first step is to select a proper training dataset, which is vital for generating accurate predictions. Historical information about past flood events can be used as a sample dataset. In this study, we selected 2,600 samples and used historical data to assign one of five risk levels: no risk, low risk, medium risk, high risk, and extremely high risk. We then divided the samples into training data sets and test data sets. The Bootstrap method (Cao 1999) was used to randomly select 70% of the data as training sets and 30% as test sets to verify the accuracy of the model. The parameter settings for the number of classification trees *n* and the number of node bifurcations *m* in the RF algorithm directly affect the accuracy of the model. After repeatedly adjusting the parameters, the number of RF classification trees was set to 50 the number of node bifurcations was set to 3.

Table 1 Comparison of accuracy of ice-jam flood hazard risk assessment model

Evaluation index	Study area			
	P		AUC	
	RF (%)	KNN (%)	RF (%)	KNN (%)
A	91	87	85	82
B	91	86	90	85
C	89	86	89	83
D	90	83	88	83

4 Results

4.1 Assessing the accuracy of the ice-jam flood hazard risk analysis

P and AUC were selected to evaluate the accuracy of the RF algorithm and KNN algorithm in the four study areas: A, B, C, and D. As shown in Table 1, the accuracy ratings of the RF algorithm model and KNN algorithm in the study areas were more than 80%, indicating that the two algorithms can effectively classify and process multi-dimensional ice-jam flood hazard data. While ensuring the accuracy of the algorithm, they can capture the relationship between the multi-dimensional feature index and the risk category. The RF and KNN algorithm models were both applicable to the ice-jam flood hazard risk assessment of the northern bank of the Inner Mongolia section of the Yellow River. The P and AUC statistics both indicated that the RF model was more accurate than the KNN model. The largest difference in accuracy between the two algorithms according to AUC was observed in area C, which the RF model predicted with 89% accuracy and the KNN model predicted with 83% accuracy. This further shows that the RF algorithm has significant advantages in classifying and processing multi-dimensional ice-jam flood hazard data. The RF model was more accurate than the KNN model in predicting ice-jam flood hazard risk along the northern bank of the Inner Mongolia section of the Yellow River.

4.2 Ice-jam flood hazard risk assessment distribution analysis

Using the results of the RF algorithm, we developed hazard risk assessment maps of the four areas and compared them with the results obtained with the KNN algorithm. The results are shown in Figs. 10 and 11. The ice-jam flood hazard risks of the four areas were

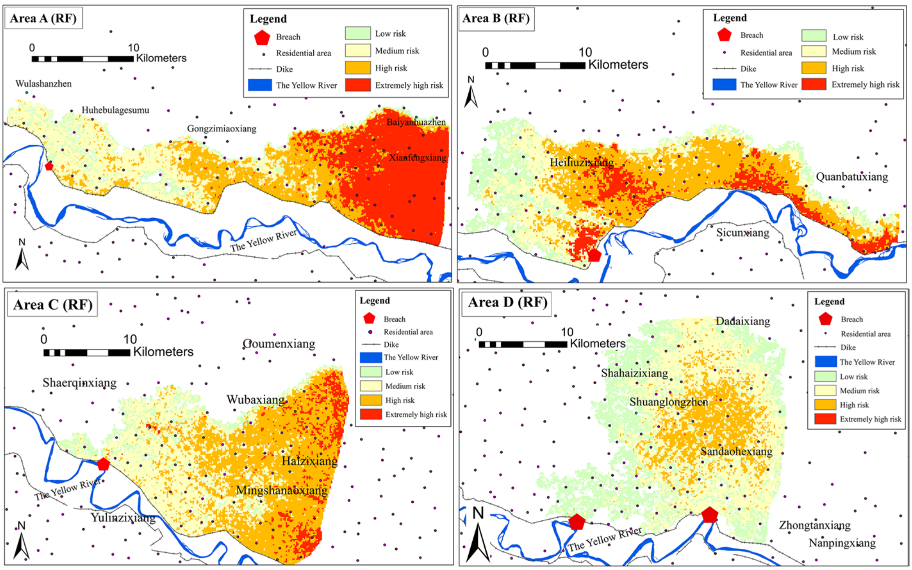

Fig. 10 Risk assessment map of ice-jam flood hazard in areas A, B, C, and D based on the RF model

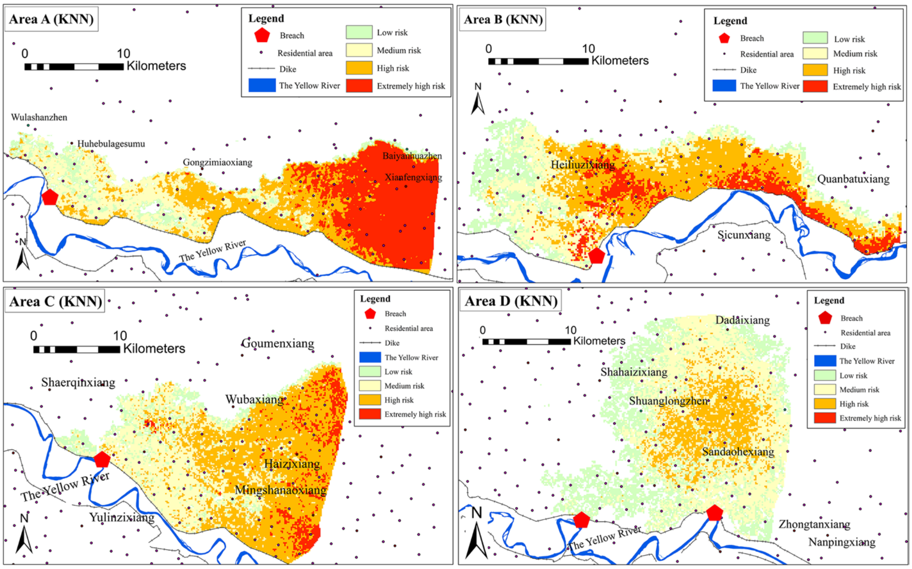

Fig. 11 Risk assessment map of ice-jam flood hazard in areas A, B, C, and D based on the KNN model

Table 2 Different risk level areas (km^2) in each study area

Risk level areas (km^2)	Study area							
	A		B		C		D	
	RF	KNN	RF	KNN	RF	KNN	RF	KNN
No risk	122.2	125.2	238.5	242.1	182.2	185.7	383.7	400.9
Low risk	23.2	26.1	50.3	45.9	36.5	34.2	153.6	137.9
Medium risk	78.3	77.4	82.4	86.6	140.8	135.7	197.8	198.1
High risk	72.2	67.8	119.8	115.8	203.7	205.6	78.9	76.8
Extremely high risk	113.7	113.1	41.4	42.0	31.0	33.0	0.2	0.5

classified, and the risk level areas of each area were counted. The results are shown in Table 2.

As shown in Table 2, for the study area A, the high and extremely high risk levels were concentrated in areas with max water depth and high population density in Xianfeng Township and Gongzimiao Township, which is consistent with historical ice-jam flood hazard data. These locations should be given priority in ice-jam flood hazard management and risk prevention. In the results of the RF model, areas with no risk, low risk, medium risk, high risk, and extremely high risk accounted for 29.8%, 5.7%, 19.1%, 17.6%, and 27.7% of the total area, respectively. In the results of the KNN model, areas with no risk, low risk, medium risk, high risk, and extremely high risk accounted for 30.6%, 6.4%, 18.9%, 16.6%, and 27.6%, respectively. The distributions of the territory among the risk levels predicted by the two models were the same. The maximum differential of roughly 1% occurred among high-risk areas.

The high risk and extremely high risk areas were smaller in area B than in area A and were mainly distributed in Heiliuzi Township and Quanbatu Township. In the results of the RF model, areas with no risk, low risk, medium risk, high risk, and extremely high risk accounted for 44.8%, 9.4%, 15.5%, 22.5%, and 7.8%, respectively. In the results of the KNN model, areas with no risk, low risk, medium risk, high risk, and extremely high risk accounted for 45.5%, 8.6%, 16.3%, 21.7%, and 7.9%, respectively. The distributions of risk levels predicted by the two models were the same. The maximum differential of 0.8% occurred in low risk, medium risk, and high-risk areas.

In study area C, the terrain is relatively flat, which provides a hazard-pregnant environment for ice-jam flooding. In addition, the large population density in this area contributed to the increase in high-risk areas compared to areas A and B. The high risk and extremely high-risk areas were mainly distributed in Haizi Township and Mingshanao Township. In the results of the RF model, the areas with no risk, low risk, medium risk, high risk, and extremely high risk accounted for 30.7%, 6.1%, 23.7%, 34.3%, and 5.2%, respectively. In the results of the KNN model, the areas with no risk, low risk, medium risk, high risk, and extremely high risk accounted for 31.3%, 5.7%, 22.8%, 34.6%, and 5.6%, respectively. The distributions of risk levels predicted by the two models were the same. The maximum differential of 0.9% occurred in the medium-risk area.

In study area D, the breach inundation was narrow in scope; the evolution rate of flooding was slow; the population density and GDP density area were low. This resulted in fewer areas being classified as high risk or extremely high risk compared with areas A, B, and C. The high-risk and extremely high-risk areas were mainly distributed in Sandaohe Township. In the results of the RF model, the areas with no risk, low risk, medium risk, high risk, and extremely high risk accounted for 47.1%, 18.9%, 24.3%, 9.7%, and 0.0%, respectively. In the KNN model, the areas with no risk, low risk, medium risk, high risk, and extremely high risk accounted for 49.2%, 16.9%, 24.3%, 9.4%, and 0.1%, respectively. The distributions of risk levels predicted by the two models were the same. The maximum differential of 2.1% occurred in the risk-free area.

The results show the differences in areas classified by the RF model and KNN model were less than 5%. Therefore, the KNN model performs well in classifying the ice-jam flood risks.

5 Discussions and conclusions

5.1 Innovation and limitation

This study analyzed the typical ice-jam flood area of the Inner Mongolia section of the Yellow River. While prior studies have tended to focus on a single ice-jam flood hazard risk indice, this study has advanced the field by selecting several ice-jam flood hazard risk indices, including ice-jam flood hazard-inducing factor, hazard-pregnant environment, and hazard-bearing body. These improvements have contributed to overcoming the incompleteness in some recent studies that undertook the ice-jam flood hazard risk assessment. For example, Wang et al. (2015) used an assessment model based on RF to evaluate regional flood hazard risk. The proposed flood hazard risk assessment method was implemented in Dongjiang River Basin, China. Eleven risk indices including ice-jam flood hazard-inducing factors, hazard-pregnant environment were selected. The support vector machine (SVM) was used for risk assessment as a comparison, as well as an analysis of index importance degree. However, it neglects the influence of socio-economic factors that do play an important role in flood control. When socio-economic factors were taken into account, some areas in high flood risk levels could have lower risk levels. In a conclusion, comprehensive flood risk is determined by natural conditions and social factors. Most of the high risk zones, exhibiting a significant threat to local residents, typically have adverse hazard-inducing factors and hazard-pregnant environments as well as a large number of hazard-bearing bodies. The hazard-bearing bodies such as population, property, cultivated land and other vulnerable factors should be considered to delineate the comprehensive flood risk. By considering the hazard and vulnerability, the integrated flood risk assessment map is more representative of the risk of the whole basin. Cai et al. (2019) developed a multi-index fuzzy comprehensive evaluation model (MFCE model) for flood disaster risk in the area of Yifeng, Jiangxi Province. The MFCE model contains three input indicators: the hazard factor, the exposure factor and the vulnerability factor, which solved the question that neglects the influence of hazard-bearing bodies. Although the fuzzy comprehensive evaluation method is an improved method of the AHP, it still has certain subjectivity in weight calculation. Further research on the sensitivity of subjective weight to risk analysis is suggested. Therefore, this study is more comprehensive in terms of indices selection and risk assessment methods selection than the previous studies.

While doing constitute progress, the study does have the following limitations. First, due to data constraints, the validation of the model cannot fully reflect the accuracy of the results and may cause certain errors. As a result, more works need to be done in the further study once the solid information are available. Second, there is no systematic system of ice-jam flood hazards risk assessment indices in the Inner Mongolia section of the Yellow River. Better and more reasonable results can be obtained if more indices are used, like the influences of dam and reservoir.

5.2 Conclusion and future research

In this study, we selected eight indices including the hazard-inducing factor, hazard-pregnant environment, and hazard-bearing body to construct the ice-jam flood hazard risk assessment model using the RF algorithm, and it was compared with a KNN risk assessment model in the Inner Mongolia section of the Yellow River. The hazard-inducing factors considered were derived from a physically based coupled 1D–2D hydrological model with a comprehensive roughness optimization method. The following conclusions are drawn:

1. In accordance with the Kuisu ice-jam flood coupled calculation model, the results show that the flow values estimated for the east and west breach of Kuisu deviated from the observed values by less than 5%. The inundation area predicted by the model matched the actual inundation area, and it was consistent with the historical data of the high risk areas, including Duguitala, Hangjinnaoer, and other townships. It shows that the coupled 1D–2D simulation model of ice-jam flood generated accurate estimates, making it a useful and accurate simulation for hazard risk assessment.
2. Both P and AUC accuracy ratings of RF model and KNN model in each level, RF was higher than KNN, indicating that the RF algorithm has significant advantages in classifying and processing multi-dimensional ice-jam flood hazard data.
3. The risk levels of the areas along the Inner Mongolia section of the Yellow River were identified such that they can be classified in descending order of risk as follows:

Sanhuhekou, Xinhekou, Sanchakou, and Shisifenzi. The high risk areas were mainly located in the vicinity of the breaches, in areas such as Xianfeng Township, Gongzimiao Township, Heiliuzi Township, Haizi Township, and Sandaohe Township. These areas should be the focus of ice-jam flood prevention and mitigation efforts in the region.

This study proved that despite a handful of drawbacks, application of the RF model to ice-jam flood hazard risk shows significant potential. Evaluation results provided a reference for ice-jam flood hazard risk management, prevention, and reduction in the upper Yellow River Basin. However, a dynamic assessment of ice-jam flood risk should also be used to provide technical support for the formulation of risk prevention and transfer policy for the future research. In addition, although RF method and KNN method to assess ice-flood hazard risk along the Inner Mongolia section of the Yellow River obtained classification results, more measured data should be collected to verify the rationality and advantages of the model. Moreover, the RF model should be expanded to include other watershed risks and compare the results among different methods to obtain a most appropriate assessment method.

Acknowledgements The authors would like to thank the faculty of Tianjin University for their valuable feedback and comments on a draft of the manuscript. Additionally, we extend our gratitude to the editor and two anonymous reviewers for their helpful comments and suggestions, which improved the quality of this manuscript.

Funding The study was supported by National Key R&D Program of China [Grant Number 2018YFC1508403], the Science Fund for Creative Research Groups of the National Natural Science Foundation of China [Grant Number 51621092], the Fund for Key Research Area Innovation Groups of China Ministry of Science and Technology [Grant Number 2014RA4031].

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Anselmo V, Galeati G, Palmieri S et al (1996) Flood risk assessment using an integrated hydrological and hydraulic modelling approach: a case study. J Hydrol 175(1–4):533–554

Ashton G (1986) River and Lake Ice Engineering. Water Resources Publications, Littleton, Colorado

Beltaos S (1993) Numerical computation of river ice jams. Can J Civ Eng 20(1):88–89

Beltaos S (2012) Distributed function analysis of ice jam flood frequency. Cold Reg Sci Technol 71:1–10. https://doi.org/10.1016/j.coldregions.2011.10.011

Bhuiyan S, Baky AA (2014) Digital elevation based flood hazard and vulnerability study at various return periods in Sirajganj Sadar Upazila, Bangladesh. Int J Disaster Risk Reduct 10:48–58

Boucher E, Bégin Y, Arseneault D (2009) Impacts of recurring ice jams on channel geometry and geomorphology in a small high-boreal watershed. Geomorphology 108:273–281. https://doi.org/10.1016/j.geomorph.2009.02.014

Bouwer L, Bubeck P, Aerts J (2010) Changes in future flood-risk due to climate and development ina Dutchpolderarea. Glob Environ Chang 20:463–471

Brayall M, Hicks F (2012) Applicability of 2-D modeling for forecasting ice jam flood levels in the Hay River Delta Canada. Can J Civil Eng 39(6):701–712. https://doi.org/10.1139/l2012-056

Cai T, Li X, Ding X et al (2019) Flood risk assessment based on hydrodynamic model and fuzzy comprehensive evaluation with GIS technique. Int J Disaster Risk Reduct 35:1–12

Cao R (1999) An overview of bootstrap methods for estimating and predicting in time series. Estadística e Invest Oper 8(1):95–116
Cao Y, Ye Y, Liang L et al (2018) Uncertainty analysis of two-dimensional hydrodynamic model parameters and boundary conditions. J Hydroelectr Eng 37(6):47–61
Chen X, Ishwaran H (2012) Random forests for genomic data analysis. Genomics 99(6):323–329. https://doi.org/10.1016/j.ygeno.2012.04.003
Chen W, Xie X, Wang J et al (2017) A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. CATENA 151:147–160. https://doi.org/10.1016/j.catena.2016.11.032
China Statistical Information Network (2010) China Statistical Information Network. www.tjcn.org.
Das A, Rokaya P, Lindenschmidt K-E (2020) Ice-jam flood risk assessment and hazard mapping under future climate. J Water Resour Plan Manag 146(6):04020029. https://doi.org/10.1061/(asce)wr.1943-5452.0001178
De Coste M, She Y, Blackburn J (2017) Incorporating the effects of upstream ice jam releases in the prediction of flood levels in the Hay River delta Canada. Can J Civil Eng 44(8):643–651
Fan X, Shi C, Zhou Y et al (2012) Sediment rating curves in the Ningxia-Inner Mongolia reaches of the upper Yellow River and their implications. Quatern Int 282:152–162. https://doi.org/10.1016/j.quaint.2012.04.044
Feng G (2014) Research on ice regime forecast model and ice flood risk in the Inner Mongolia Reach of the Yellow River Tianjin University, China: 1–62
Flato G, Gerard R (1986) Calculation of ice jam thickness profiles. In: Proceedings of the 4th CGU-HS CRIPE workshop on river ice, Montreal, Que
Fread D (1980) DAMBRK: the NWS dam-break flood forecasting model. Hydrology technical note: 1–56
Gounaridis D, Chorianopoulos I, Symeonakis E et al (2019) A random forestcellular automata modeling approach to explore future land use/cover change in Attica (Greece), under different socio-economic realities and scales. Sci Total Environ 646:320–335. https://doi.org/10.1016/j.scitotenv.2018.07.302
Han T, Jiang D, Qi Z (2018) Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. Trans Inst Meas Control 40(8):2681–2693
Hicks F, Steffler P, Gerard R (1992) Finite element modeling of surge propagation and an application to the Hay River, N.W.T. Can J Civ Eng 19:454–462
Hopkins MA, Tuthill AM (2002) Ice boom simulations and experiments. J Cold Reg Eng 16(3):138–155
Hu J, Liu J, Liu Y et al (2013) EMD-KNN model for annual average rainfall forecasting. J Hydrol Eng 18(11):1450–1457
Jiang Z, Liu H, Fu B et al (2019) Decomposition theories of generalization error and AUC in ensemble learning with application in weight optimization. Chin J Comput 42(1):1–15
Jon EZ, Ettema R (2000) Fully coupled model of ice-jam dynamics. J Cold Reg Eng 14(1):24–41
Kundzewicz ZW, Kanae S, Seneviratne IS et al (2013a) Flood risk and climate change: global and regional perspectives. Hydrol Sci J 59(1):1–28. https://doi.org/10.1080/02626667.2013.857411
Kundzewicz ZW, Pińskwar I, Brakenridge GR (2013b) Large floods in Europe, 1985–2009. Hydrol Sci J 58(1):1–7. https://doi.org/10.1080/02626667.2012.745082
Lal AW, Shen HT (1991) Mathematical model for river ice processes. J Hydraul Eng 117(7):851–867
Lan Q, Li W, Liu W (2016) Performance and choice of chinese text classification models in different situations. J Hunan Univ (Nat Sci) 43(4):141–146
Leo B (2001) Random forests. Mach Learn 45(1):5–32. https://doi.org/10.1023/A:1010933404324
Li Q, Mao Y (2013) AUC optimization boosting based on data rebalance. Acta Geogr Sin 39(9):1467–1475. https://doi.org/10.3724/SP.J.1004.2013.01467
Lindenschmidt K-E, Sydor M, Carson R (2012) Modelling ice cover formation of a lake–river system with exceptionally high flows (Lake St. Martin and Dauphin River, Manitoba). Cold Reg Sci Technol 82:36–48
Lindenschmidt K-E, Das A, Rokaya P et al (2016) Ice-jam flood risk assessment and mapping. Hydrol Process 30(21):3754–3769. https://doi.org/10.1002/hyp.10853
Mao Z, Wu J, Zhang L et al (2003) Numerical simulation of river ice jam. Adv Water Sci 14(6):700–705
Mihăilescu DM, Gui V, Toma CI et al (2013) Computer aided diagnosis method for steatosis rating in ultrasound images using random forests. Med Ultrason 15(3):184
Morse B, Hicks F (2005) Advances in river ice hydrology 1999–2003. Hydrol Process 19(1):247–263. https://doi.org/10.1002/hyp.5768
Oliveira P, Boccelli D (2017) K-nearest neighbor for short term water demand forecasting. World Environmental and Water Resources Congress 2017. May 21–25, 2017. Sacramento, California
Osei BK, Ahenkorah I, Ewusi A et al (2021) Assessment of flood prone zones in the Tarkwa mining area of Ghana using a GIS-based approach environmental. Challenges 3:1–12

Penning-Rowsell E, Floyd P, Ramsbottom D et al (2005) Estimating injury and loss of life in floods: a deterministic framework. Nat Hazards 36(1):43–64. https://doi.org/10.1007/s11069-004-4538-7
Pham BT, Luu C, Dao DV et al (2021a) Flood risk assessment using deep learning integrated with multi-criteria decision analysis. Knowl-Based Syst 219:106899. https://doi.org/10.1016/j.knosys.2021.106899
Pham BT, Luu C, Phong TV et al (2021b) Flood risk assessment using hybrid artificial intelligence models integrated with multi-criteria decision analysis in Quang Nam Province. Vietnam J Hydrol 592:125815. https://doi.org/10.1016/j.jhydrol.2020.125815
Rodriguez-Galiano V, Ghimire B, Rogan J et al (2012) An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J Photogramm Remote Sens 67(1):93–104. https://doi.org/10.1016/j.isprsjprs.2011.11.002
Shen HT, Su J, Liu L (2000) SPH simulation of river ice dynamics. J Comput Phys 165(2):752–770. https://doi.org/10.1006/jcph.2000.6639
Shen HT, Shen H, Tsai SM (2010) Dynamic transport of river ice. J Hydraul Res 28(6):659–671
Tan X, Zhang W, Ma J et al (2004) Research on regional assessment of flood risk and regionalization mapping in China. J China Inst Water Resour Hydropower Res 2(1):54–64
Tian F, Wang Y, Wang X (2021) Classification and distribution characteristics of ice flood risk in the Ningxia-Inner Mongolia reach of the Yellow River. Yellow River 43(2):49–52. https://doi.org/10.3969/j.issn.1000-1379.2021.02.010
Tingsanchali T, Karim F (2010) SJFlood-hazard assessment and risk-based zoning of a tropical flood plain: case study of the Yom River, Thailand. Hydrol Sci J J des Sci Hydrol 55(2):145–161
Todorovic P, Zelenhasic E (1970) A stochastic model for flood analysis. Water Resour Res 6(6):1641–1648
Wang Z, Chen X, Lai C et al (2013) Flood risk assessment model based on particle swarm optimization rule mining algorithm. Syst Eng-Theory Pract 33(6):1615–1621
Wang Z, Lai C, Chen X et al (2015) Flood hazard risk assessment model based on random forest. J Hydrol 527:1130–1141
Winsemius H, Aerts J, Beek LV et al (2015) Global drivers of future river flood risk. Nat Clim Chang 6(4):381–385. https://doi.org/10.1038/nclimate2893
Wu C, Wei Y, Jin J et al (2014) Comprehensive evaluation of ice disaster risk of the Ningxia-Inner Mongolia Reach in the upper Yellow River. Nat Hazards 75(S2):179–197. https://doi.org/10.1007/s11069-014-1308-z
Yang K, Liu Z, Li G et al (2002) Simulation of ice jam channel. Water Resour Hydropower Eng 33(10):40–47
Yang J (2019) ROC analysis method based on K-nearest neighbor classifier a dissertation. Guangdong University of Technology, the Degree of Master: 1–63
Zhou C, Wan Q, Huang S et al (2000) A GIS-based approach to flood risk zonation. Acta Geogr Sin 55(1):15–24
Zhou Y, Li S, Zhou C et al (2019) Intelligent approach based on random forest for safety risk prediction of deep foundation pit in subway stations. J Comput Civil Eng 33(1):05018004. https://doi.org/10.1061/(asce)cp.1943-5487.0000796
Zou Q, Zhou J, Zhou C et al (2012) Comprehensive flood risk assessment based on set pair analysis-variable fuzzy sets model and fuzzy AHP. Stoch Env Res Risk Assess 27:525–546