Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Ice Flood Hazard Risk Assessment Based on Random Forest Method Under Simulation of Levee Breaches

Xiujie Wang
State Key Laboratory of Hydraulic Engineering Simulation and Safety
Zihua Qu (zihuashuyu@163.com)
State Key Laboratory of Hydraulic Engineering Simulation and Safety https://orcid.org/0000-0002-8826-2018
Fuchang Tian
State Key Laboratory of Hydraulic Engineering Simulation and Safety https://orcid.org/0000-0003-3769-5259
Yanpeng Wang
Yangtze Ecology and Environment Co.,Ltd.
Ximin Yuan
State Key Laboratory of Hydraulic Engineering Simulation and Safety
Kui Xu
State Key Laboratory of Hydraulic Engineering Simulation and Safety

Research Article

Keywords: The Yellow River's Inner Mongolia section, Ice flood hazard risk assessment, Roughness, Hazard-inducing factors, Random forest (RF) model

Posted Date: March 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1360466/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Ice flood hazard risk assessment based on random forest method under simulation of levee breaches

Xiujie Wang[a, b], Zihua Qu[a, b], Fuchang Tian[a, b, *], Yanpeng Wang[c], Ximin Yuan[a, b] , Kui Xu[a, b]

[a] *State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin 300072, China*

[b] *School of Civil Engineering, Tianjin University, Tianjin 300072, China.*

[c] *Yangtze Ecology and Environment Co., Ltd., Wuhan 430000, China*

Abstract

In areas with higher latitudes, ice sheets, ice jams or ice dams are often formed in rivers to overtop levees due to lower temperatures in winter, which can cause sizeable social and economic losses, as well as countless injuries and death. This study presents the typical ice flood area of the Inner Mongolia section of the Yellow River as the research object. On the basis of the hazard system theory and comprehensive analysis of ice flood risk factors, including the hazard-inducing factors and hazard-pregnant environments, as well as the vulnerability of the hazard-bearing bodies, we selected eight risk assessment indicators to construct an ice flood hazard risk assessment model based on random forest (RF) algorithm. The three hazard-inducing factors consist of: maximum submergence depth, maximum flood velocity and maximum submergence duration, which were derived from the river-flood ice flood backwater burst submergence coupling model; while the three hazard-pregnant environments are: topographic elevation, terrain gradient and the distance from the river. The two hazard-bearing bodies include: population density and GDP density. The modeling results show that compared with the risk assessment model of K-nearest neighbor (KNN) algorithm, both the index Precision (P) and the area under curve (AUC) of RF model are better in the ice flood

* *Corresponding author at: State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, China.*
E-mail address: tianfuchang@tju.edu.cn (F.C. Tian).

risk assessment of four study areas. This presents that RF has significant advantages in solving the classification and processing problem of multi-dimensional ice flood hazard data. It can provide support for the analysis and evaluation of the situation of ice flood prevention and mitigation.

Keywords

The Yellow River's Inner Mongolia section; Ice flood hazard risk assessment; Roughness; Hazard-inducing factors; Random forest (RF) model

1 Introduction

Major ice jams that often form when the flow of ice chunks stemming from upstream of the jamming site are arrested and accumulated along a river reach to cause damming of the river, which often block the high water level, overflow the embankment, submerge villages, farmland, roads, industrial and mining enterprises(Wu et al. 2014; Boucher et al. 2009), and pose a serious threat to river facilities and the lives and property of people on both sides of the river(Morse and Hicks 2005; Zbigniew et al. 2013; Pham et al. 2021; Ashton 1986). Therefore, the prevention and control of river ice flood hazard is also an important element related to people's livelihood(Kundzewicz et al. 2013; Bouwer et al. 2010). Affected by the cold air in Siberia and Mongolia in winter and spring, the Yellow River Basin has more northerly wind, cold climate and little rain. As a result, the main and tributaries of the Yellow River have different degrees of ice flood in winter, especially the upper reaches in the Ningxia Inner Mongolia (Ningmeng), which is one of the key ice flood controlling sections on the Yellow River. Considering its geographical location, river conditions and other factors, winter ice flood conditions there are more severe, thus causing devastating damage to the local area. To date, as a consequence of the air temperature, human activities, riverbed siltation and other changes in the basin environment, the threat of ice flood not only still exists, but the occurrence

of its frequency in the river section has also increased and the scope of impact is gradually expanding(Pham et al. 2021). According to statistics, during 1951-2010, a total of 30 years of ice flood hazards occurred in Inner Mongolia section of Yellow River in China, causing significant losses to the country as well as people's lives and properties. The ice flood hazard system is complex, which includes hazard-inducing factors, hazard-pregnant environments, and hazard-bearing bodies. It has the characteristics of high nonlinearity, spatial-temporal dynamics, and uncertainty, and coupling of various challenges in the system may produce extremely complex phenomena. In this regard, the study of ice flood risk assessment can provide technical support for the analysis of the situation of ice flood prevention and mitigation and the development of risk prevention and control programs, which is of great significance for the prevention and control of river ice flood hazards and the reduction of economic losses.

Ice flood risk assessment performs a comprehensive evaluation of natural and social attributes of ice floods with the aim of improving the accuracy of grasping the spatial distribution of ice flood risk, which is to divide the study area into different risk levels based on the degree of danger of the ice flood(Wu et al. 2014), through the form of maps to visually show the spatial distribution of the ice flood risk, which is an important basis for the formulation of the risk management and prevention transfer plan of the ice flood(Tian et al. 2021). Based on the principle of extreme value, Todorovic and Zelenhasic (1970) used the 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 indicators to analyze the rationality of the zoning results; Tan et al. (2004) comprehensively considered factors such as

flood submergence, socio-economic and hazard-pregnant environment to establish a zone County-level flood risk zoning model. Beltaos (2012) created a DFM method using the peak flow as the data source, and successfully carried out the risk assessment of ice blocking and flooding. Wang et al. (2013) established a flood risk model based on particle swarm optimization in the North River Basin of China; Wu et al. (2014) developed a comprehensive evaluation model for ice hazard levels in the Ningxia-Inner Mongolia section of the Yellow River using projection pursuit, fuzzy clustering and accelerated genetic algorithms. Based on on-the-spot investigations, Michael et al. (2017) used a prediction model for assessing ice flood risk by simulating the combination of ice and water in the Hay River and the delta based on the Hay River town located in the northwestern region of Canada. It can be seen that although a considerable number of studies have already been instrumental in providing site-specific understanding of ice flood risk indicators, the selection of ice flooding factors in previous studies is relatively single, or only one or two indicators are selected for expression, which cannot fully reflect the degree of ice flood. In view of this, this paper selecting the maximum submergence depth, maximum submergence flood velocity and maximum submergence duration under 100-year return period as hazard-inducing factors for research, which need to be obtained through numerical model simulation is proven necessary. Performing accurate ice flood risk assessment in China and other developing countries is of great significance, as it can guide stakeholders and government officials to focus on areas prone to ice flood hazards and improve regional management and planning efficiency.

Numerical simulation of ice flood is the basis of risk analysis, management and evaluation according to hydrodynamic methods, river ice dynamics methods and theoretical methods. To address this issue, Lal and Shen (1991) proposed the one-dimensional RICE river ice model, which

89 comprehensively considered the distribution characteristics of water temperature and ice
90 concentration. Based on hydraulics, thermodynamics and ice hydrodynamic theories, it simulates
91 the process of ice transportation, blockage and diving, which is widely acknowledged as a
92 pioneering research work in the numerical simulation of ice flood. Beltaos (1993) came up with the
93 RIVJAM river ice hydrodynamic model to simulate the water level changing process of the ice jam
94 in a wide shallow channel; Jon and Ettema (2000) employed a numerical model to simulate the
95 dynamic damage and reconstruction of the ice jams and simulated the changing process of under-
96 ice overcurrent based on the theory of hydrodynamics and ice transport; A and Andrew (2002)
97 applied Digital Elevation Model (DEM) to test a non-curved river ice state system affected by ice
98 bellicatters; Yang et al. (2002) analyzed and simulated the ice flood siltation and ice jam formation
99 process by studying the ice transport movement; Wang and Zhao (2008) combed the research
100 progress of river ice jam numerical simulation and ice hydraulics, and put forward a more systematic
101 river ice numerical model theory; Shen et al. (2010) realized the dynamic simulation of ice
102 transportation, blockage, and overcurrent under ice by studying the dynamic formation process of
103 ice sheet, and then established a full dynamic one-dimensional hydrodynamic model considering
104 ice effect, and introduced ice resistance and ice diffusion into the model, and applied it to the St.
105 John's River; Li (2015) used DynaRICE model to simulate the evolution of river ice formation and
106 evolution in Sanhuhekou. Lindenschmidt et al. (2016) using Monte-Carlo simulation and other
107 freezing numerical simulation methods, conducted a risk assessment of the ice flood hazard along
108 the Peace River in Canada, and analyzed the vulnerability of local cities and towns. However,
109 predecessors' research limitedly focused on the observation of freezing conditions, the evolution of
110 river ice sheets, river ice transport, and analysis of the cause of ice flood hazard to establish a series

111 of static and dynamic numerical river ice models such as RIVJAM (Beltaos and Wong 1986; Uzuner
112 and Kennedy 1974; Uzuner and Kennedy 1976), ICEJAM (Flato and Gerard 1986),
113 RIVER1D(Hicks et al. 1992) , DynaRICE (Shen et al. 2000), HECRAS (Daly and Vuyovich 2003),
114 RIVICE(Lindenschmidt et al. 2012) , RIVER2D (Brayall and Hicks 2012), and ICESIM, while only
115 a limited number of studies have resulted on the coupled combined numerical model of the ice flood
116 backwater burst submergence and the evolution process of the flooding area in the ice flood. For
117 example, the scholar (Feng 2014) used one-dimensional river gates and dams to block the water to
118 reflect the changing process of the high water level of the ice dam, but ignored the discharge capacity
119 of the river after the ice dam blocked, which is different from the actual ice dam water evolution.
120 However, large deviations and the backwater of the river ice dams have caused the cross-section
121 wet cycle and the roughness and flow resistance to increase, which have not been reflected in the
122 previous ice flood evolution models. Given the above concerns, this study particularly proposes a
123 comprehensive roughness optimization method of riverbed ice dam, to increase the roughness by
124 setting the ice block and ice dam section to simulate the stagnation process caused by the upstream
125 inflowing water blocked by the ice dam. In allusion to the characteristics of the ice flood evolution
126 in the two-dimensional flood area, a comprehensive roughness optimization method for the flow ice
127 surface layered is built to simulate the ice flood evolution process.

128 In the context of global warming change, due to the special geographical location, hydrological,
129 climatic conditions (Beltaos 2012) and river channel characteristics of the Yellow River, as well as
130 the constraints of meteorological prediction accuracy and forecast period, the current research on
131 ice situation cannot fully meet the needs of ice flood control of the Yellow River. Therefore, it is
132 urgent to carry out a systematic and in-depth research on the evolution of the ice flood risk

assessment of the Yellow River, which provides an opportunity to assess the usefulness of an existing ice flood protection system under future climatic conditions, particularly to evaluate whether it is adequate to accommodate future ice flood hazard risks. In this paper, the potential ice flood hazard risk assessment along the upper reaches of the Yellow River in the Ningxia Inner Mongolia was examined. The hazard-inducing factors selected maximum submergence depth, maximum flood velocity and maximum submergence duration derived from a physically based coupled 1D-2D hydrological model to evaluate ice flood scenarios and subsequent backwater level profiles. Additionally, topographic elevation, terrain gradient and the distance from the river were selected as the hazard-pregnant environments, and the population density and GDP density were selected as the hazard-bearing bodies to construct the ice flood hazard risk assessment model based on the RF model, and it is compared with KNN risk assessment model. The rest of the paper is organized as follows. Section 2 introduces the study area and data used; while Section 3 shows the methodology of the study. Section 4 displays detailed results of the methods. In Section 5 we present a series of conclusions on the implementation of the proposed methods. The developed framework of the current research is presented as a flowchart in Fig. 1. Some of the most important features relevant to this research work are the following:

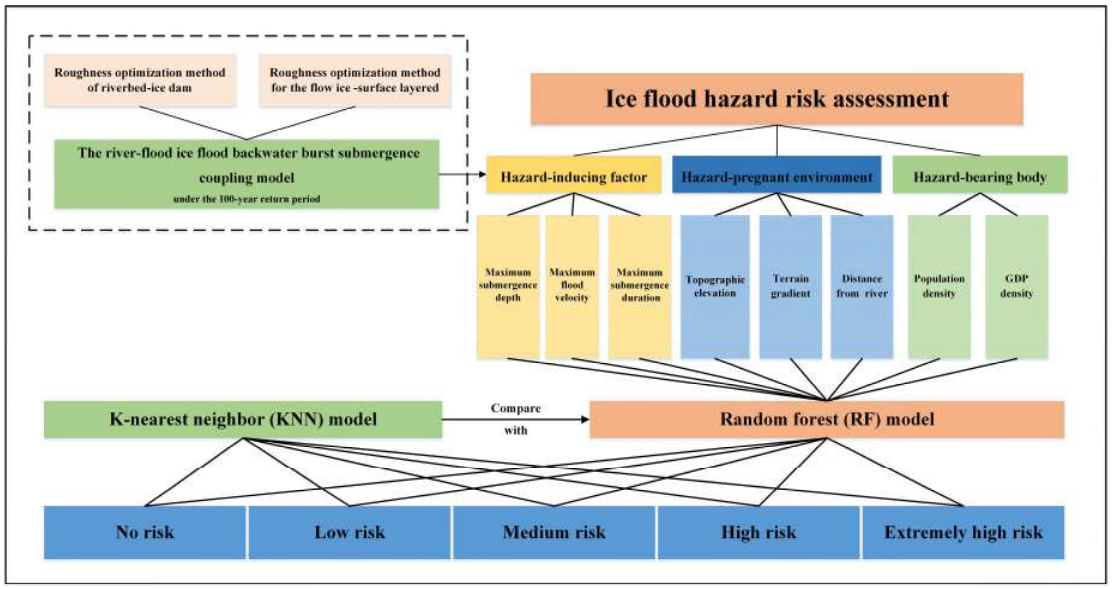


Fig. 1. The developed framework of the current research.

2. Study area and data

2.1 Study area

Inner Mongolia section of the Yellow River is located in the northernmost end of the Yellow River, west from Wuda District, east to Junger Banner, belongs to the winter ice dam prone river section. The total length of the river section is 843km. The river section is mainly distributed in Shizuishan, Bayangaole, Sanhuhekou, Toudaoguai and other hydrological stations. Among them, Bayangaole to the head of the river section in the Yellow River basin at the northernmost is one of the river sections with the most frequent ice flood hazard and the most serious hazard losses. Affected by geographical location and climatic characteristics, the river closure and opening sequence of the Inner Mongolia section of the Yellow River is opposite. During the opening of the Yellow River in Inner Mongolia, flowing ice was blocked and accumulated, and ice dams were most likely to form at the river bends. Due to the frozen dam, the overflow section is reduced and the upstream water level rises sharply. Generally, the water rise of ice dam varies from 0.5 to 6m. After

the water rises to a certain height, the ice dam was overwhelmed by water pressure. After the ice dam breaks, the water level drops rapidly, which can drop by 1.5m in one day. Affected by the temperature and the trend of the river course, the closure of the Inner Mongolia section of the Yellow River traces its source from bottom to top, and it is easy to form ice jams and high-water levels, and even flood dikes. On the contrary, the closure and opening sequence intensifies the ice flood hazard of the section. This paper selects this reach as the research area, which has important practical significance for non-engineering measures and hazard prevention and reduction management.

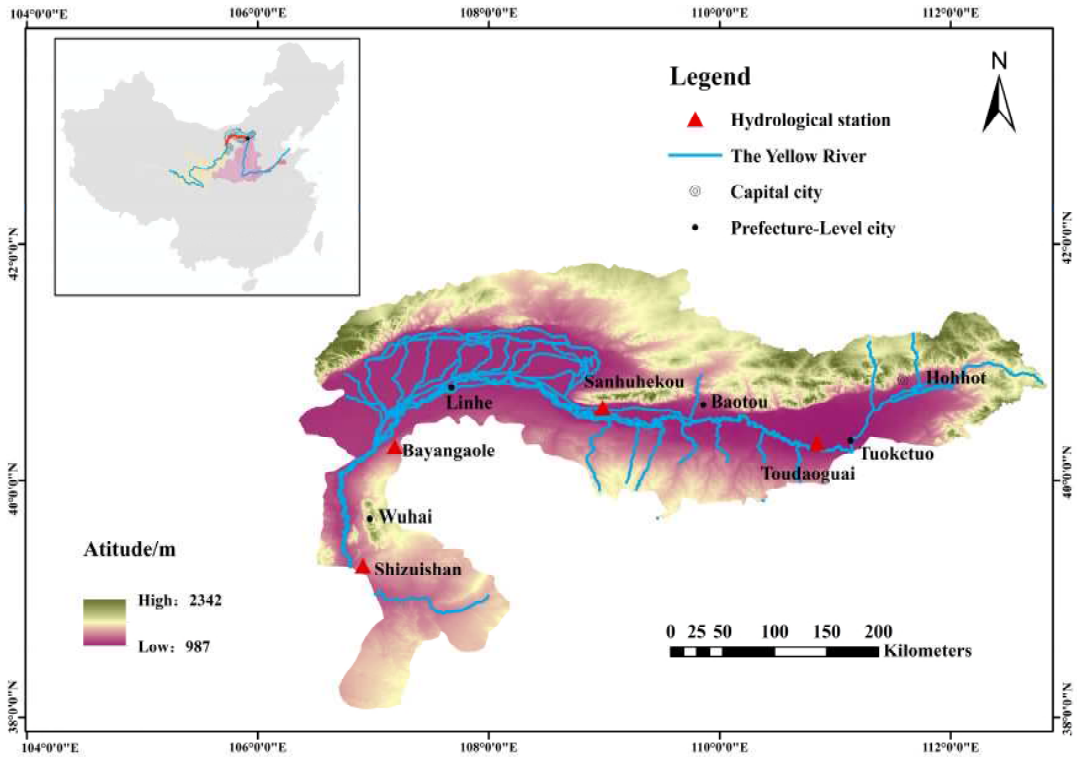


Fig.2. Location of study area, the Yellow River's Inner Mongolia section.

2.2 Data sets

The data in this study mainly include the data required by the coupled 1D-2D model established to obtain the ice flood hazard-inducing factors and the ice flood risk assessment indicators. River data and cross-section data correspond to the basic data of the coupled 1D-2D

177 model, which were obtained by Municipal Water Authority of the Inner Mongolia Section of the
178 Yellow River. Data on the boundary conditions were provided by the measured flood flow process
179 and water level flow relationship. The Shuttle Radar Topography Mission (SRTM) digital
180 elevation model (DEM), which has a 90-meter resolution, served as the basis for data on
181 topographic elevation, submergence depth, flood velocity, distance to river and the calculation of
182 slopes and aspects. Population density and GDP density data were collected from www.tjcn.org.

183

184 ## 3. Methodology

185 ### 3.1 River-flood ice flood backwater burst submergence coupling model

186 Based on the principle of hydrodynamics, considering the backwater characteristics of one-
187 dimensional river channel and the evolution characteristics of two-dimensional flood in floodplain,
188 the ice flood backwater burst submergence coupling model is constructed, so that it can calculate
189 the changes of hydraulic factors such as high water level change, dike break flow process, flood
190 submergence process, water depth and velocity.

191 (1) The control equation of one-dimensional river flood evolution

192 $$\frac{\partial Q}{\partial \mathrm{x}}+\frac{\partial A}{\partial t}=\mathrm{q} \tag{1}$$

193

194 $$\frac{\partial Q}{\partial t}+\frac{\partial(\alpha \frac{Q^2}{A})}{\partial x}+gA\frac{\partial Z}{\partial x}+\frac{gQ|Q|}{C^2AR}=0 \tag{2}$$

195 $$v=\frac{1}{n}R^{2/3}J^{1/2} \tag{3}$$

196 $$n=\left(\frac{\chi_b n_b{}^{3/2}+\chi_i n_i{}^{3/2}}{\chi_b+\chi_i}\right)^{2/3} \tag{4}$$

197 For natural rivers, 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 flow, m^3/s; 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 Chezy coefficient, $s/m^{1/3}$; R is the hydraulic radius, m; q is the unit width flow, m^2/s; Z is the ice flood variable water level, m; n is the comprehensive roughness; n_i is the roughness of the ice dam; n_b is river bed roughness; χ_i is the wet cycle of the ice dam, m; χ_b is the river bed wet cycle, m; α is the momentum correction coefficient. The roughness of the river section is increased by setting the ice dam to simulate the backwater process caused by the upstream water being blocked by the ice dam.

(2) The two-dimensional numerical simulation control equation of ice flood is as follows (Mao et al. 2003; Cao et al. 2018)：

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

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

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

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

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

where h is the water depth, m; z_b is the topographic elevation, m; τ_{ix} and τ_{iy} are the components of flow drag force in the x and y directions respectively; u and v are the velocity components in the x and y directions respectively, m/s; τ_{bx} and τ_{by} are the components of surface friction in the x and y directions respectively; n_I is the flow roughness; n_B is the surface roughness; q_L is the

source and sink terms. According to the characteristics of the evolution of the ice flood in the two-dimensional floodplain, the method of optimizing the comprehensive roughness of the flow ice-surface layered is adopted to simulate the evolution process of the ice flood in the two-dimensional floodplain.

The coupled 1D-2D hydrodynamic model realizes the dynamic simulation calculation of the whole process of ice flood backwater burst submergence through the coupling connection at the breach. In this paper, the NWS DAMBRK method is used to realize the real-time dynamic coupling calculation of the river-floodplain ice flood evolution model.

3.2 Risk assessment model of ice flood hazard

Random forest (RF) model

Random forest (RF) algorithm is a combination classification intelligent algorithm based on statistical learning theory (Leo 2001). At present, 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 converge to a lower generalization error, effectively improve the generalization ability of the algorithm, and the operation can be highly parallelized, thereby improving the computational efficiency of the model (Rodriguez-Galiano et al. 2012; Chen et al. 2017; Zhou et al. 2019).

The generation steps of RF algorithm are shown in Fig. 3: a. randomly select k sub-training sets S_1, S_2 ,..., S_K with bootstrap sampling method, and build K classification trees; b. At each node of the classification tree, randomly select m from n indicators, and select the optimal segmentation index for segmentation; c. Repeat the above steps until k classification trees are traversed; d. Cluster

K classification trees to construct the whole random forest.

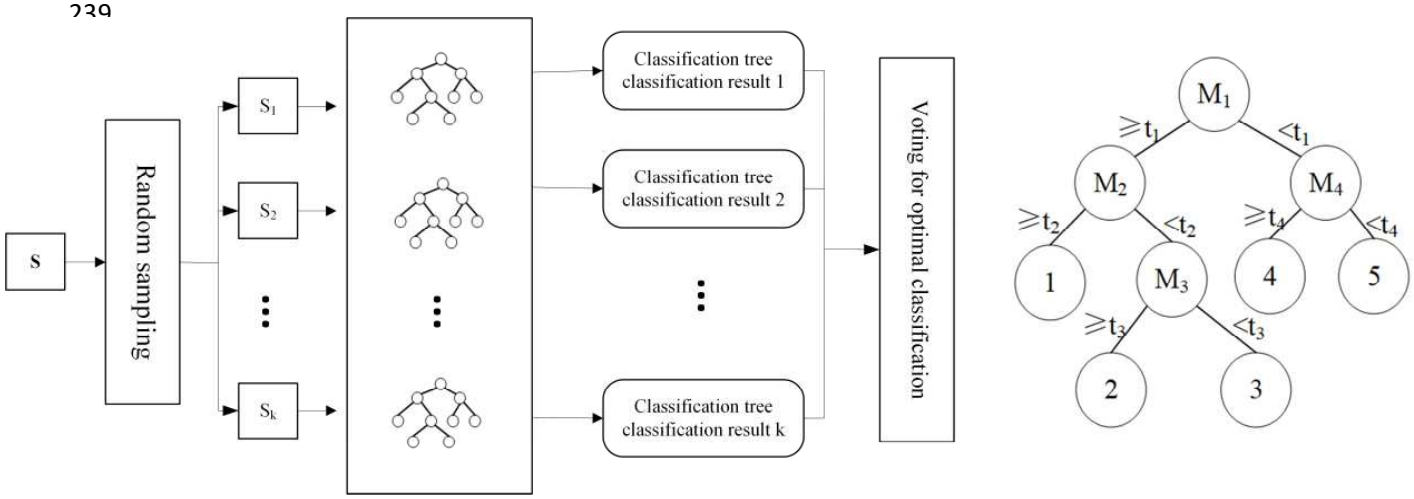


Fig. 3. Steps of RF decision tree generation.

When the RF algorithm is used to evaluate the risk level of the ice flood hazard, the sample set to be predicted needs to be brought into the trained RF classification tree. The risk level distributed on each leaf node is the result of the risk level division of the corresponding classification tree (Gounaridis et al. 2019). Perform data averaging on the risk level classification results of all classification trees to obtain the entire ice flood hazard risk zoning result based on RF algorithm.

$$p(c \mid v) = \sum_{t=i}^{T} P_{t}(c \mid v) \tag{11}$$

where T is the number of trees in RF; c is a certain risk level; $p(c \mid v)$ is the probability of ice flood hazard risk level c at leaf node v.

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 RF algorithm is guaranteed, and the model also has a high tolerance for sample abnormal values, which

avoids overfitting of data samples (Chen and Ishwaran 2012).

K-nearest neighbor (KNN) model

Additionally, to estimate the effects of the proposed RF method as-applied to ice flood hazard risk, another algorithm is compared to RF. The KNN algorithm is a non-parametric pattern recognition and classification algorithm based on statistics(Yang 2019). Due to the simplicity of its implementation process, it has been applied in many fields, such as text classification (Lan et al. 2016), short-term water demand forecasting (Oliveira and Boccelli 2017), annual average rainfall forecasting (Hu et al. 2013). The specific principle is shown in the literature and will not be repeated here.

Accuracy evaluation index

Statistical measures used to validate the models include 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 for the development of reliable ice flood susceptibility assessment.

In this paper, P and AUC are used to evaluate the accuracy of the ice flood hazard risk assessment of the RF algorithm and the KNN algorithm(Li and Mao 2013; Jiang et al. 2019). P is the ratio of positive samples predicted by the classifier to positive samples. AUC refers to the probability that the positive samples output by the classifier are positive than the probability that the negative samples output by the classifier are positive. P intuitively reflects the accuracy and feasibility of the classifier algorithm applied to the ice flood hazard risk assessment, and AUC accurately reflects the relationship between the true class rate and the false positive class rate of the

classifier, so as to evaluate and compare the classification quality of each classifier.

$$P = \frac{TP}{TP + FP} \times 100\% \tag{12}$$

$$AUC = \frac{\sum_{ins_i \in positiveclass} rank_{ins_i} - \frac{M \times (M+1)}{2}}{M \times N} \tag{13}$$

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; $rank_{ins_i}$ denotes the No. i sample; and M, N are the number of positive samples and negative samples respectively.

4. Results

According to the ice flood source, topography and historical ice flood hazard in the flood area on the North Bank of Inner Mongolia section, and comprehensively considering the characteristics of ice flood in Inner Mongolia section of the Yellow River, a relatively perfect ice flood risk evaluation index system is constructed based on hazard-inducing factors, hazard-pregnant environments and hazard-bearing bodies. By means of the ArcGIS platform and Python language, the RF model is used to carry out the risk assessment of ice flood hazard in the Inner Mongolia section of the Yellow River. P and AUC are employed to evaluate the accuracy of each model, and compared with the results obtained by KNN model to verify the applicability and accuracy of RF model for the ice flood risk assessment.

4.1 Hydrodynamic model calibration and ice flood hazard factors validation

The coupled 1D-2D hydrodynamic models under the condition of a 100-year return period from Sanhuhekou to Toudaoguai section extending along the Yellow River was adopted in this study

for the hazard-inducing factors selecting maximum submergence water depth, the maximum flood velocity, and the maximum submergence duration to develop the ice flood hazard risk assessment. The specific establishing process is as follows:

Step 1. The 1D river flow models

Using the measured succeeding cross-section data of Sanhuhekou-Toudaoguai section (The river is 316km long in total, with 88 measured cross sections.), the one-dimensional river flow hydrodynamic model was established. The inflow boundary is located at Sanhuhekou, which is the design ice flood discharge process; while the outflow boundary is located at Toudaoguai, which is the relationship between water level and flow.

Step 2. The 2D floodplain flow models

Taking the dangerous sections, historical dike break, high ice flood of this river section, the utilization of residential areas and the comprehensive consideration of experts into account, Sanhuhekou (A), Sanchakou (B), Xinhekou (C) and Shisifenzi (D) were selected as the study areas, all located on the left bank of the river. The setting of the break is shown in Fig. 4. The unstructured grid is used to divide the terrain of the study area. Topographic data comes from NASA (https://www.nasa.gov). The zoning roughness of residential land and dry land are set respectively.

311

312 **Fig. 4.** The distribution map of the ice flood risk analysis plan for the ice flood break

313 Calibration and analysis of model.

314 **Step 3.** The coupled 1D-2D models

315 The model realizes the dynamic connection between the 1D river channel and the 2D

316 floodplain of the ice flood through the real-time coupling connection of the ice flood at the breach.

317 Considering the most unfavorable conditions, the breach is instantaneous collapse to the end.

318 According to the historical ice flood hazards and expert opinions, the width of the breach is set to

319 100m. One breach is set in study areas A, B and C respectively, and two are set in study area D.

320 The four study areas all lack historical ice flood records, while the Kuisu area on the right bank

321 of the river experienced ice flood in 2008, and relevant data were recorded. Data showed that on

322 March 20, 2008, the Kuisu section of the Inner Mongolia section of the Yellow River was affected

323 by the rapidly rising water levels at the Sanhuhekou hydrological station. Two breaches occurred in

324 the morning. The east and west breaches were about 1km apart and occurred one after another. The

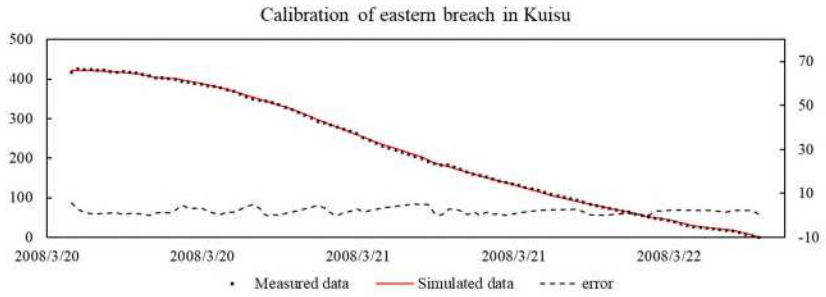
325 maximum width of the east and west breach is 100m and 60m separately. The ice flood caused the

326 submergence of towns such as Duguitala and Hangjinnaoer. The submergence area reached 126km^2,
327 and the direct economic losses reached 935 million RMB.

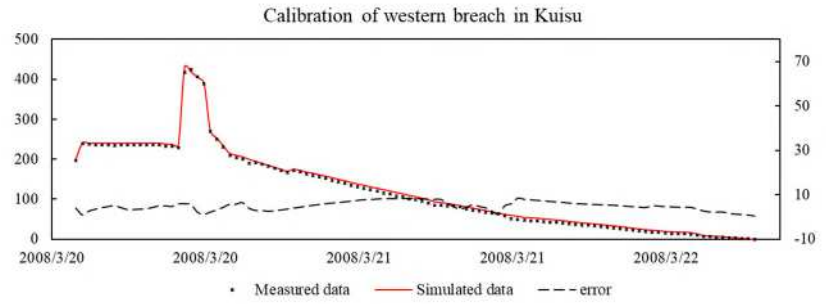
328 On the basis of the previous modeling steps, a coupled 1D-2D model of Kuisu area was
329 established. To reflect the dike burst submergence evolution process caused by ice flood backwater,
330 a comprehensive roughness optimization method of riverbed-ice dam was used to rate the river
331 channel roughness. Referring to the *Specification for ice flood computation SL428-2008*, combined
332 with the model roughness calibration and verification, the roughness of 10km downstream of the
333 breach in each study area was set to 0.0975. The width of the east and west breach was set to 100m
334 and 60m respectively according to the measured data. The starting time of the model was from
335 16:00:00 on March 18, 2008 to 16:00:00 on March 28, 2008. And the SRTM 90m DEM Terrain was
336 used to construct the two-dimensional ice flood submergence analysis model of Kuisu, and the
337 unstructured grid was used to divide the terrain of the study area. The maximum grid area was no
338 more than 0.01km^2, with a total of 33,500 grids. The zoning roughness of residential land and dry
339 land was set to reflect the impact of the actual evolution process of ice flood on residential land,
340 which was validated as 0.08 and 0.04 respectively. Set the initial water depth of the grid to 0.01m,
341 dry water depth to 0.005m, and wet water depth to 0.1m.

342 In accordance with the Kuisu ice flood coupled calculation model, the dynamic flow process
343 at the east and west breach of the model and the two-dimensional maximum submergence range in
344 the Kuisu area were extracted, and compared with the actual flow process of the breach and the
345 actual submergence range (Fig. 5). The results shows that the error between the calculated results
346 of the flowmeter at the east and west breach of Kuisu and the measured values is less than 5%. And
347 the calculated submergence range of the ice flood coupling model basically covers the historical

submergence area, and it is consistent with the historical data of the hazard-stricken area including Duguitala Town and Hangjinnaoer Township and other townships. The total submergence area is 111.51km^2 (Fig. 6). It can be seen that the coupled 1D-2D simulation model of ice flood established in this paper has high calculation accuracy and can meet the needs of ice flood numerical simulation and ice flood risk assessment.



a)



b)

Fig. 5. Validation results of one-dimensional river ice flood numerical model of Sanhuhekou-Toudaoguai: a) Eastern breach in Kuisu; b) Western breach in Kuisu.

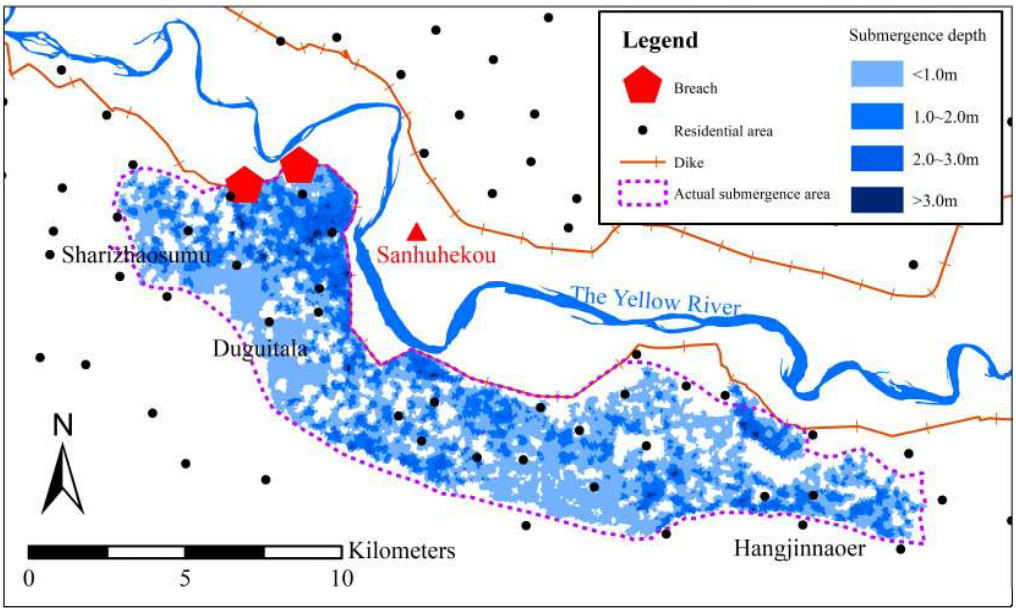


Fig. 6. Numerical simulation verification results of two-dimensional submergence area in Kuisu.

Ice flood hazard-inducing factor validation

According to the ice flood coupled model parameters calibrated in the Kuisu section, the 100-year return period ice flood in the four study areas of A, B, C and D were simulated and calculated, and the maximum submergence water depth, maximum flood velocity and maximum submergence duration were extracted as shown in Fig. 7, Fig. 8, Fig. 9 and Fig. 10 respectively.

According to the analysis of three ice flood hazard-inducing factors, areas with the maximum submergence depth in area A is concentrated in the river bend and low-lying downstream areas, while the submergence range is distributed along the river trend, and the submergence distribution range is relatively narrow and long. Contrastively, the maximum submergence depth in area B is concentrated at the breach and near the river channel; in area C at the breach and the low-lying area downstream and in area D in the low and flat area in the middle. In addition, due to the flat terrain of the flood area on the north bank, affected by the terrain and the surface roughness of the flow, the maximum flood velocity in area A is concentrated in the southern river bend and the upstream area,

and the corresponding maximum submergence duration is long, indicating that this area should be the focus of the management of ice flood hazard prevention. In contrast, the maximum flood velocity in area B is distributed in the area close to the river channel, along the river trend, and the submergence duration is also long, which indicates that the distance from the river is one of the important indexes to evaluate the ice flood risk. Nevertheless, in area C, due to the fact that the terrain in the central area is slightly higher than that in the northeast, which is easy to lead to large flood submergence range and large submergence depth in the northeast, the maximum flood velocity is concentrated in the central area, conveying that the ice flood risk is greatly affected by the terrain as well. The maximum flood velocity and maximum submergence depth in area D are distributed in the central region, in which the low and flat terrain and large resident population play an important part. Once the ice flood breaks the dike, the life and property safety of the resident population will be largely threatened.

Ice flood hazard-pregnant environment validation

Hazard-pregnant environment refers to the external environmental conditions of hazard-inducing factors and hazard-bearing body, such as topography, water system, vegetation distribution, etc. From the analysis of the background and mechanism of ice flood formation, the hazard-pregnant environment mainly considers the comprehensive impact of terrain and water system on the formation of ice flood. The lower the elevation is and the smaller the topographic relief gets, the more prone it is to trigger ice floods. The denser the river network is and the closer it is to the water body, the higher the risk of ice flood becomes (Bhuiyan and Baky 2014). Therefore, in this study, topographic elevation, terrain gradient and the distance from the river (Penning-Rowsell et al. 2005)

395 are selected as hazard-pregnant environment in this study, as shown in Fig. 7, Fig. 8, Fig. 9 and Fig.

396 10 respectively.

397 ***Ice flood hazard-bearing body validation***

398 The risk of hazard-inducing factors only reflects the possible harm caused by ice flood, and the

399 actual degree of harm is also related to the situation of the hazard-bearing body. Ice flood with the

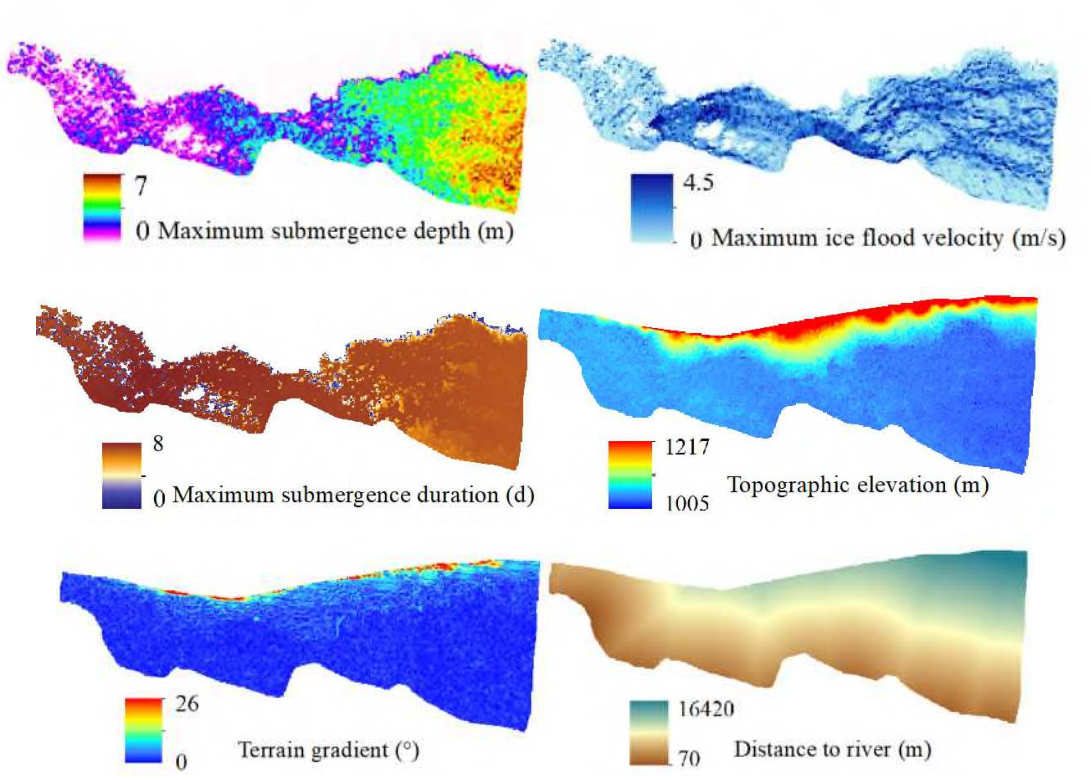
400 same intensity occurs in densely populated (Zou et al. 2012) and economically developed areas, and

401 the losses are often much greater than those in sparsely populated and economically backward areas

402 (Winsemius et al. 2015). Therefore, this study selects population density and GDP density as the

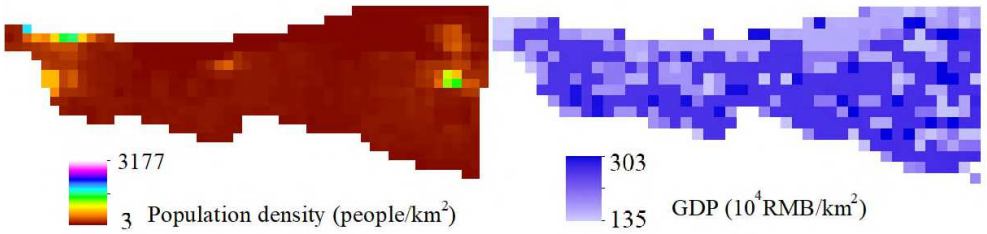
403 hazard-bearing body, as shown in Fig. 7, Fig. 8, Fig. 9 and Fig. 10 respectively.

404



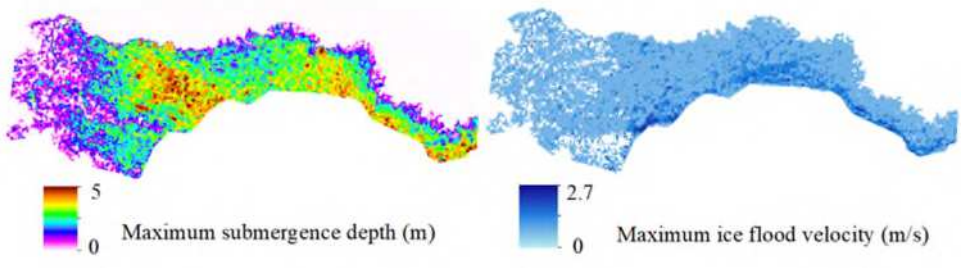
405

406

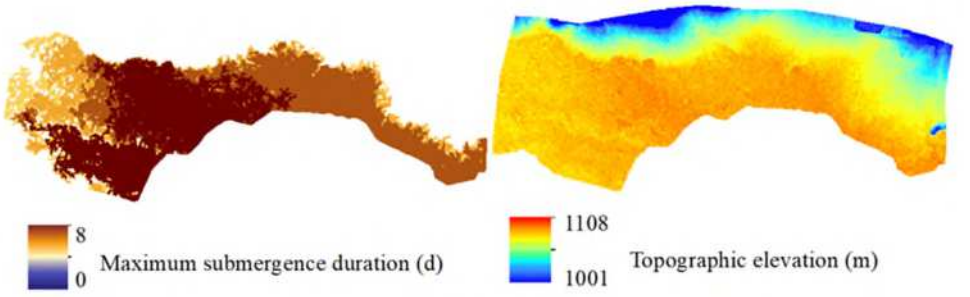


407

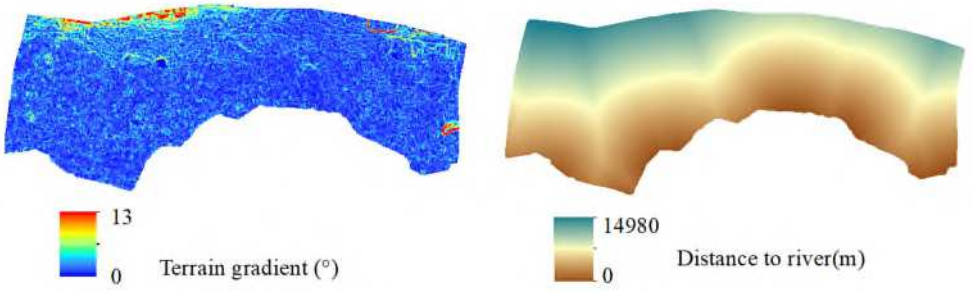
408 **Fig. 7.** Assessment indicators for hazard risk in area A.



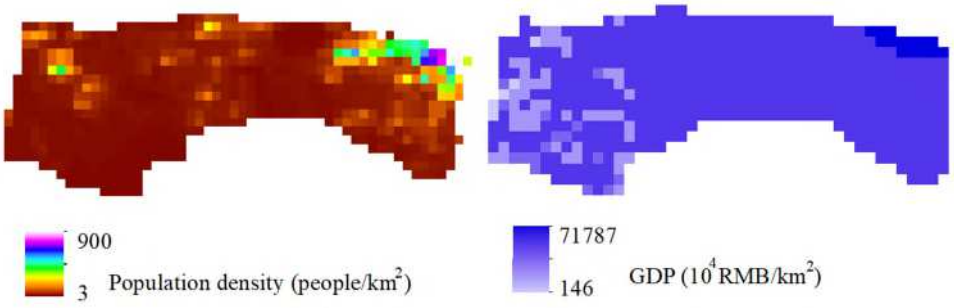
409



410



411



412

413 **Fig. 8.** Assessment indicators for hazard risk in area B.

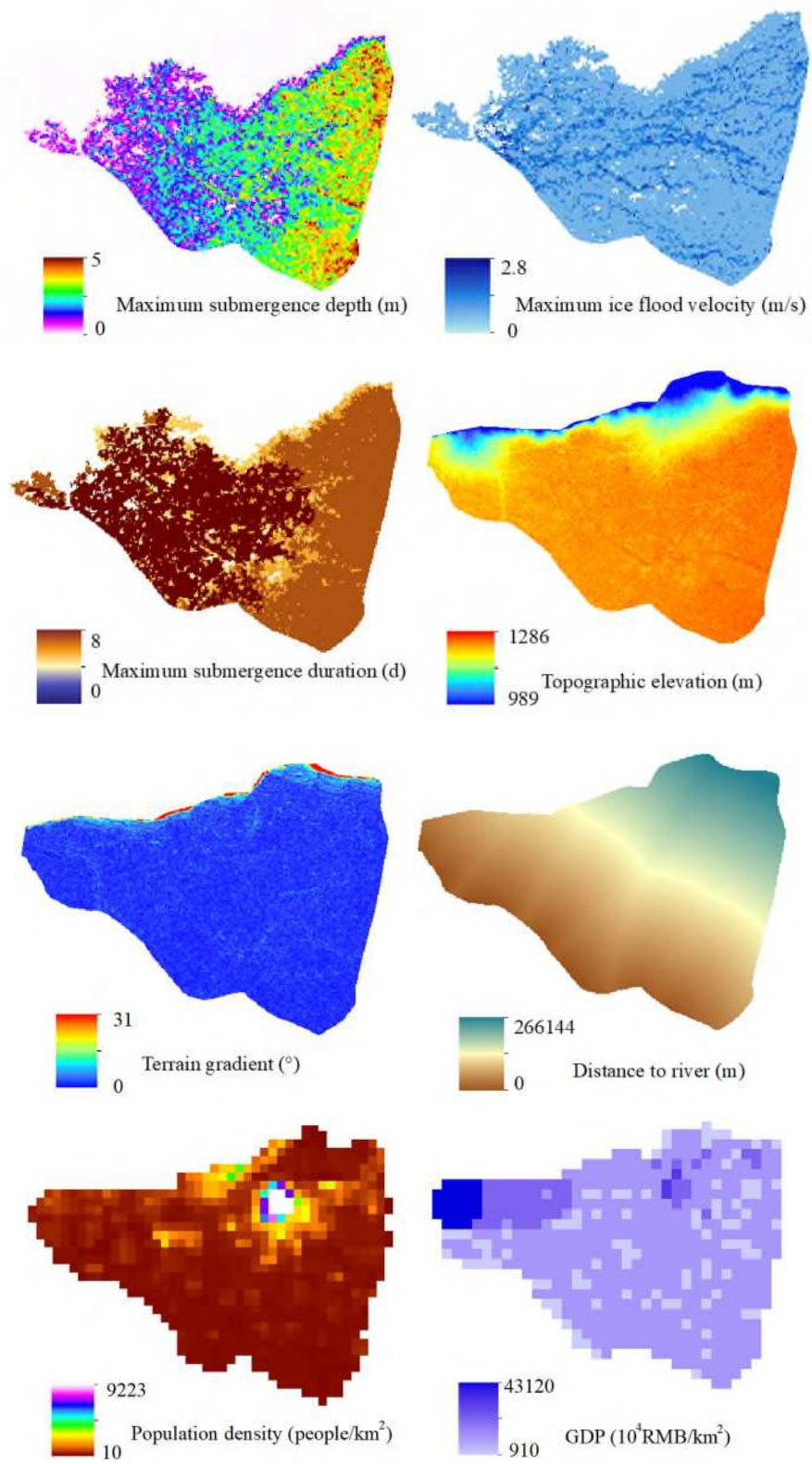


Fig. 9. Assessment indicators for hazard risk in area C.

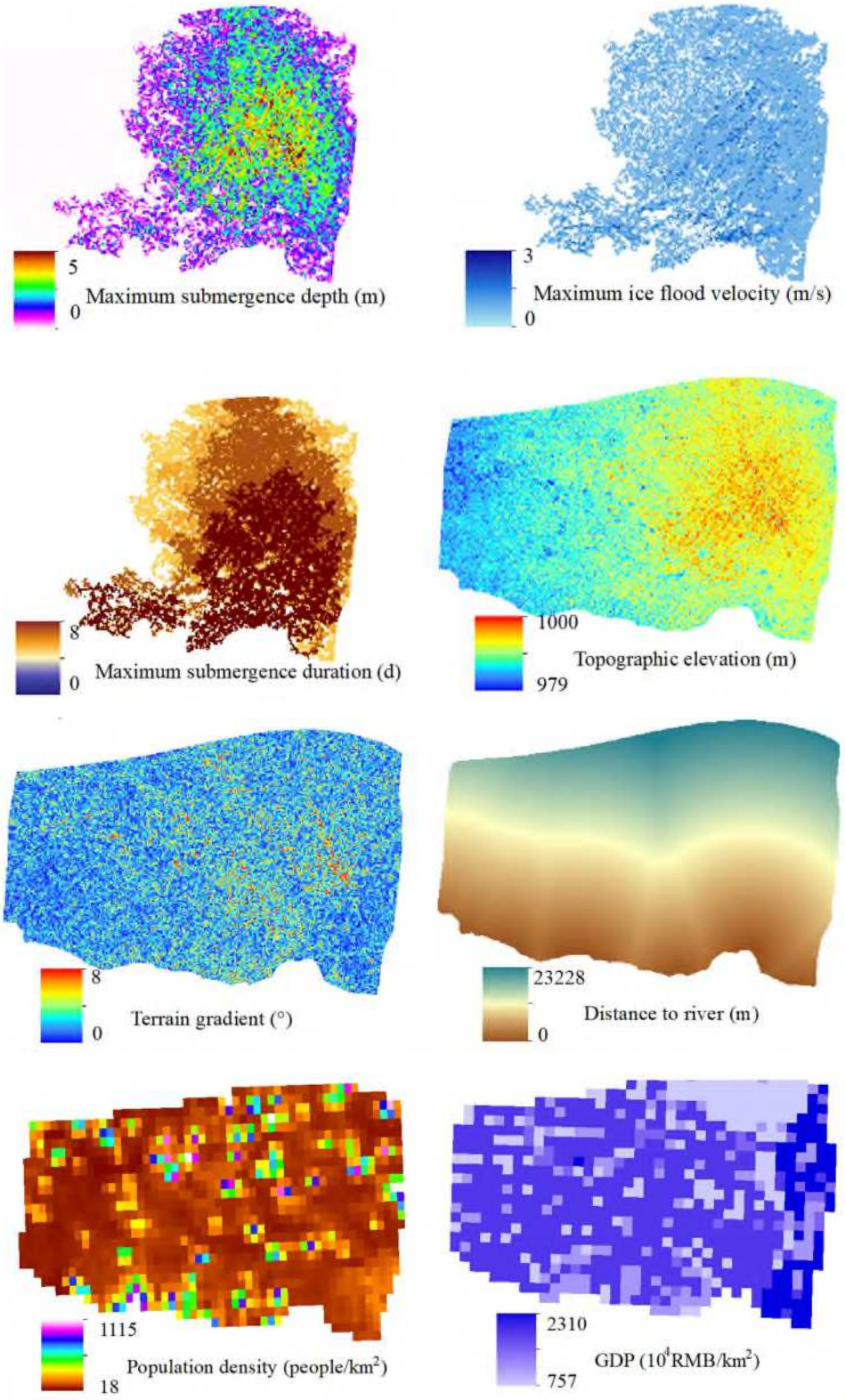


Fig. 10. Assessment indicators for hazard risk in area D.

4.2 Ice flood risk assessment

Ice flood risk assessment accuracy analysis

To implement the RF model, the first and vital step is to select a proper training dataset, which significantly affects the quality and effectiveness of the model. The historical information about past flood events can be utilized to select the sample dataset. In this study, we select 2,600 samples in total, and manually identify and analyze the sample risk levels based on historical data records, namely no risk, low risk, medium risk, high risk, extremely high risk, and divide the samples into training data sets and test data sets. The Bootstrap method is used to randomly select 70% as the training data sets and 30% as the test data sets to test the model to verify the accuracy of it. The parameter settings of 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, it is determined that the number of RF classification trees n is 50, and that of node bifurcations m is 3. In order to compare and verify the feasibility and accuracy of the RF model, the KNN model is selected for comparison. The model measures the Euclidean distance between the sample to be tested and the known sample to realize the prediction of ice flood hazard risk level. P and AUC are used to evaluate the accuracy of the sample.

In this study, two evaluation indexes of P and AUC are selected to evaluate the accuracy of the following RF algorithm and KNN algorithm in four areas A, B, C and D. According to the analysis of Table 1, the accuracy of RF algorithm model and KNN algorithm in study areas A, B, C and D are more than 80%, indicating that RF algorithm and KNN algorithm have significant advantages in solving the problem of multi-dimensional ice flood hazard data classification and processing.

While ensuring the accuracy of the algorithm, it can capture the relationship between the multi-dimensional feature index and the risk category. RF algorithm model and KNN algorithm model are both applicable to the ice flood risk assessment of the northern bank of the Inner Mongolia section of the Yellow River. Both P and AUC draw the same conclusion when evaluating accuracy, that is RF > KNN. The maximum disparities in AUC do arise in area C. The differential between the two algorithm is 6%, and it increases from 83% to 89% (Table 1). It further shows that RF algorithm has significant advantages in solving the classification and processing of multi-dimensional ice flood hazard data. Compared with the KNN algorithm, it has better performance for ice flood hazard risk assessment in the northern bank of the Inner Mongolia section of the Yellow River.

Table 1

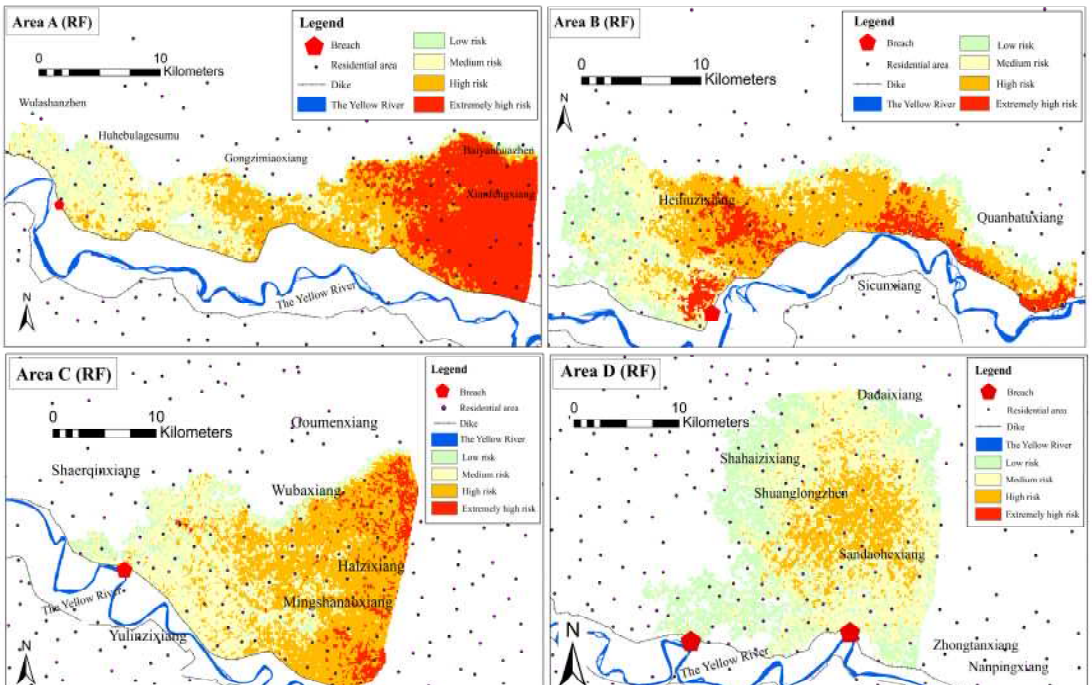
Comparison of accuracy of ice flood hazard risk assessment model.

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%

Ice flood risk assessment distribution analysis

Based on the RF algorithm, draw the hazard risk assessment map of the four areas A, B, C, and D, and compare them with the results obtained by the KNN algorithm. The results are shown in Fig. 11 and Fig. 12 respectively. The ice flood risk of areas A, B, C and D on the North Bank of the Inner Mongolia section of the Yellow River is classified, and the risk level areas of each study area are

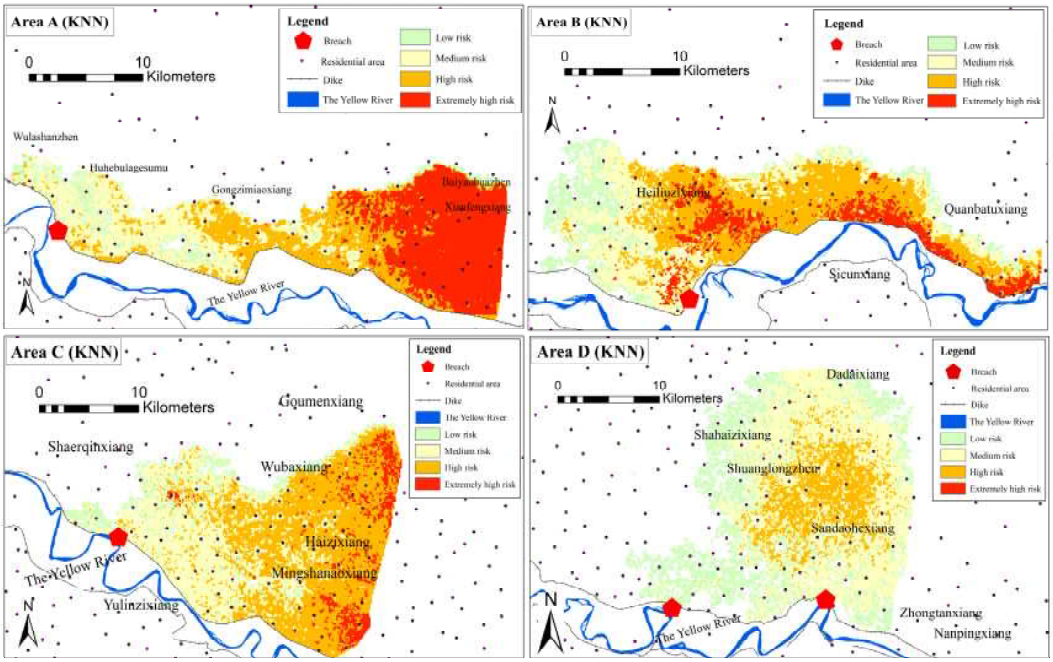
461 counted. The results are shown in Table 2. The following four areas are analyzed separately:



462

463

464 **Fig. 11.** Risk assessment map of ice flood hazard in area A, B, C and D based on RF model.



465

466 **Fig.12.** Risk assessment map of ice flood hazard in area A, B, C and D based on KNN model.

467

Table 2

Statistical table of different risk level areas (km^2) of ice flood hazard risk assessment in each study area.

Study area / Risk level areas (km^2)	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

As we can see from Table 2, for study area A, high and extremely high risks are concentrated in areas with maximum submergence depth and high population density, distributed in Xianfeng Township and Gongzimiao Township, which are consistent with historical ice flood hazard data. They should also be given priority in ice flood hazard management and risk prevention. In the results of RF model, the areas with no risk, low risk, medium risk, high risk and extremely high risk account for 29.8%, 5.7%, 19.1%, 17.6% and 27.7% respectively, and KNN model 30.6%, 6.4%, 18.9%, 16.6% and 27.6% respectively. It is obvious that the ratio of the two is basically the same, and the maximum differential occurs in high risk areas, by no more than 1%.

For study area B, the high risk and extremely high risk areas decreased compared with area A, mainly distributed in Heiliuzi Township and Quanbatu Township. In the results of RF model, the areas with no risk, low risk, medium risk, high risk and extremely high risk account for 44.8%, 9.4%,

15.5%, 22.5% and 7.8% respectively, and KNN model 45.5%, 8.6%, 16.3%, 21.7% and 7.9% respectively. It is obvious that the area proportions obtained by the two are basically the same, and the maximum differential occurs in low risk, medium risk and high risk areas, by 0.8%.

For study area C, the terrain is relatively flat, which provides a hazard-pregnant environment for ice flood hazard. In addition, the large population density in this area contributes to the increase of area above high risk in area C. High risk and extremely high risk are mainly distributed in Haizi Township and Mingshanao Township. In the results of RF model, the areas with no risk, low risk, medium risk, high risk and extremely high risk account for 30.7%, 6.1%, 23.7%, 34.3% and 5.2% respectively, and KNN model 31.3%, 5.7%, 22.8%, 34.6% and 5.6% respectively. It can be seen that the area proportions obtained by the two are basically the same, and the maximum differential occurs in the medium risk area, by 0.9%.

For study area D, the scope of breach submergence in this area is small, the evolution rate of flood is slow, and the population density and GDP density of this area are low, resulting in the reduction of high risk and extremely high risk in this area compared with the above three areas. The high risk and extremely high risk areas are mainly distributed in Sandaohe Township. In the results of RF model, the areas with no risk, low risk, medium risk, high risk and extremely high risk account for 47.1%, 18.9%, 24.3%, 9.7% and 0.0% respectively, and KNN model 49.2%, 16.9%, 24.3%, 9.4% and 0.1% respectively. It can be seen that the area proportions obtained by the two are basically the same, and the maximum differential occurs in the risk-free area, by 2.1%.

The results show that the differential between RF model and KNN model is controlled within 5%. Therefore, KNN model is also reasonable as an index to characterize the hazard characteristics of ice flood.

505 ## 5. Discussion

506 In this study, an assessment model based on RF was adopted to evaluate regional ice flood
507 hazard risk. The proposed ice flood hazard risk assessment method was implemented in the Inner
508 Mongolia section of the Yellow River. Eight risk indices were selected and 2,600 samples were
509 created for training and testing. The K-nearest neighbor (KNN) was used for risk assessment as a
510 comparison.

511 Some relevant researches (Stefanidis and Stathis 2013; Wang et al. 2015; Lai et al. 2016) also
512 focused on the assessment of flood hazards based on the analytic hierarchy process (AHP), the
513 support vector machine (SVM), the Ant-Miner and so on. While comparing with these previous
514 researches, there are some major advantages between this study and them. Following are the
515 deficiencies discussed of other studies in three aspects:

516 For example, index weights must be calculated by manual intervention before the AHP method
517 can be used, leading to high subjectivity. Additionally, the results come out from flood hazard
518 indexes were classified into four classes depending on the likelihood of flood hazard (low, medium,
519 high and very high). The division was achieved according to the Natural Breaks method.
520 Consequently, there are more qualitative than quantitative data in AHP evaluation; so that the results
521 are less than satisfactory.

522 Additionally, the zoning map for the flood hazard risk only reflects the probability of being
523 impacted by floods, the higher the risk level is, the easier the areas would be flooded. When socio-
524 economic factors were taken into account, some areas in high flood risk levels could have lower
525 risk levels. The reason is that if fewer people, property and crops are in the area with high level of
526 flood hazard risk, then the casualties and property losses would be lower despite the probability of

being flooded is high. 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 zoning map is more representative of the risk of the whole basin.

What's more, although the Ant-Miner has significant advantages in terms of both implementation step-reducing and computing time-saving. However, there are still some limitations of using this method to discover rules. The rules generated by the Ant-Miner have a large number of boxes in the feature space. This is because the condition part of the rule only contains the term using the logical operator AND.

In this case, risk assessment system for ice and flood hazards has been established, and the random forest (RF) algorithm has been used to carry out the risk assessment of the ice flood hazard in the north bank of the Yellow River Inner Mongolia. This study proved that despite a handful of drawbacks, application of the RF model to flood hazard risk shows significant potential.

6. Conclusion

Ice flood risk assessment has scientific and practical values in the context of basin-scale water resource and ice flood hazard management. In order to analyze and assess the ice flood risk, this study presents the typical ice flood area of the Inner Mongolia section of the Yellow River as the research object. Based on comprehensive consideration of ice flood hazard-inducing factors, hazard-pregnant environment and hazard-bearing bodies, this study selects eight risk assessment indicators,

549 constructs the ice flood hazard risk assessment model based on RF algorithm, and compares it with
550 the risk assessment model of KNN algorithm. The maximum submergence depth, maximum flood
551 velocity and maximum submergence duration are selected as the ice flood hazard-inducing factors,
552 which are obtained by constructing the river-flood ice flood backwater burst submergence coupling
553 model. The hazard-pregnant environment selects three indicators of topographic elevation, terrain
554 gradient and the distance from the river, while the hazard-bearing body selects two indicators of
555 population density and GDP density. Finally, P and AUC are used to evaluate the ice flood risk
556 assessment. The main conclusions of this paper are as follows：
557 (1) Compared with the traditional coupled 1D-2D hydrodynamic model, the river-flood ice
558 flood backwater burst submergence coupling model method takes the variation of the roughness of
559 the ice dam and the evolution characteristics of the ice flood into account, which provides a more
560 feasible and simplified algorithm for the numerical simulation of the ice flood evolution.
561 (2) Based on the calculation results of ice flood coupled 1D-2D model, three hazard-inducing
562 factors are selected together with hazard-pregnant environments and hazard-bearing bodies to
563 develop the ice flood hazard risk assessment based on RF algorithm model. By contrasting with
564 previous research both on the methods and the selection of indicators, it draws conclusion that the
565 RF algorithm exhibits higher accuracy in solving the classification and processing of multi-
566 dimensional ice flood hazard data. Compared with the KNN algorithm, it is more suitable for ice
567 flood hazard risk assessment in the northern bank of the Inner Mongolia section of the Yellow River.
568 (3) In the ice flood risk assessment of the Inner Mongolia section of the Yellow River, the
569 overall risk of the study areas is Sanhuhekou>Xinhekou>Sanchakou>Shisifenzi. The high risk areas
570 are mainly located in the vicinity of the breach and Xianfeng Township, Gongzimiao Township,

Heiliuzi Township, Haizi Township, Sandaohe Township and other areas. The above areas should be the focus areas for ice flood prevention and mitigation.

Highlights

• Risk assessment system for ice and flood hazards has been established, and the random forest (RF) algorithm has been used to carry out the risk assessment of the ice flood hazard in the north bank of the Yellow River Inner Mongolia.

• A coupled 1D-2D model is established to simulate the ice flood of dike break, and the maximum submergence depth, maximum flood velocity and maximum submergence duration were extracted as the hazard-inducing factors.

• In this model, a comprehensive roughness optimization method of riverbed ice dam and a method for the flow ice surface layered were used to replace the actual ice dam and ice water.

Acknowledgements

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].

Ethics 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

A HM, Andrew MT (2002) Ice Boom Simulations and Experiments. Journal of Cold Regions Engineering 16(3).

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

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

Beltaos S (1993) Numerical computation of river ice jams. Canadian Journal of Civil Engineering 20(1), 88-89.

Beltaos S (2012) Distributed function analysis of ice jam flood frequency. Cold Regions Science and Technology 71, 1-10. https://doi.org/10.1016/j.coldregions.2011.10.011

Beltaos S, Wong J (1986) Downstream transition of river ice jams. ASCE Journal of Hydraulic Engineering 112(2), 91-110. https://doi.org/10.1061/(ASCE)0733-9429(1986)112:2(91)

Bhuiyan S, Baky AA (2014) Digital elevation based flood hazard and vulnerability study at various return periods in Sirajganj Sadar Upazila,Bangladesh. International Journal of Disaster Risk Reduction 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. Global Environmental Change 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. Canadian Journal of Civil Engineering 39(6), 701-712.

https://doi.org/10.1139/l2012-056

Cao Y, Ye Y, Liang L et al. (2018) Uncertainty analysis of two-dimensional hydrodynamic model parameters and boundary conditions Journal of Hydroelectric Engineering 37(6), 47-61. https://doi.org/10.11660/slfdxb.20180606

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

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

Daly S, Vuyovich C (2003) Modeling river ice with HEC-RAS.In Proceedings of the 12th CGU-HS CRIPE Workshop on River Ice, Edmonton, Alta, 11 pp

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.

Flato G, Gerard R (1986) Calculation of Ice Jam Thickness Profiles.Proceedings of the 4th CGU-HS CRIPE Workshop on River Ice, Montreal,Que, pp

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. Transactions of the Institute of Measurement and Control 40(8), 2681-2693.

641 Hicks F, Steffler P, Gerard R (1992) Finite element modeling of surge propagation and an application to
642 the Hay River, N.W.T. Can. J. Civ. Eng 19, 454-462.
643 Hu J, Liu J, Liu Y et al. (2013) EMD-KNN Model for Annual Average Rainfall Forecasting. Journal of
644 Hydrologic Engineering 18(11), 1450-1457.
645 Jiang Z, Liu H, Fu B et al. (2019) Decomposition Theories of Generalization Error and AUC in Ensemble
646 Learning with Application in Weight Optimization. Chinese Journal of Computers 42(1), 1-15.
647 https://doi.org/10.118971/SP.J.1016.2019.00001
648 Jon EZ, Ettema R (2000) Fully Coupled Model of Ice-Jam Dynamics. Journal of Cold Regions
649 Engineering 14(1), 24-41.
650 Kundzewicz ZW, Kanae S, Seneviratne IS et al. (2013) Flood risk and climate change: global and
651 regional perspectives. Hydrological Sciences Journal 59(1), 1-28.
652 https://doi.org/10.1080/02626667.2013.857411
653 Lai C, Shao Q, Chen X et al. (2016) Flood risk zoning using a rule mining based on ant colony algorithm.
654 Journal of Hydrology 542, 268-280. https://doi.org/10.1016/j.jhydrol.2016.09.003
655 Lal AW, Shen HT (1991) Mathematical Model for River Ice Processes. Journal of Hydraulic Engineering
656 117(7), 851-867.
657 Lan Q, Li W, Liu W (2016) Performance and Choice of Chinese Text Classification Models in Different
658 Situations. Journal of Hunan University(Natural Sciences) 43(4), 141-146.
659 https://doi.org/10.16339/j.cnki.hdxbzkb.2016.04.019
660 Leo B (2001) Random Forests. Machine Learning 45(1), 5-32.
661 https://doi.org/10.1023/A:1010933404324
662 Li C (2015) Study on Characteristics of River Ice Evolution and Numerical Simulation of the Yellow

River (Inner Mongolia reach). Inner Mongolia Agricultural University.

Li Q, Mao Y (2013) AUC Optimization Boosting Based on Data Rebalance. Acta geographica sinica 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 Regions Science Technology 82, 36-48.

Lindenschmidt K-E, Das A, Rokaya P et al. (2016) Ice-jam flood risk assessment and mapping. Hydrological Processes 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. Advances in Water Science 14(6), 700-705. https://doi.org/10.14042/j.cnki.32.1309.2003.06.005

Michael DC, 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. Canadian Journal of Civil Engineering 44(8).

Mihăilescu DM, Gui V, Toma CI et al. (2013) Computer aided diagnosis method for steatosis rating in ultrasound images using random forests. Medical ultrasonography 15(3), 184.

Morse B, Hicks F (2005) Advances in river ice hydrology 1999-2003. Hydrological Processes 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

Penning-Rowsell E, Floyd P, Ramsbottom D et al. (2005) Estimating injury and loss of life in floods: a deterministic framework. Natural Hazards 36(1), 43-64. https://doi.org/10.1007/s11069-004-4538-7

685 Pham BT, Luu C, Dao DV et al. (2021) Flood risk assessment using deep learning integrated with multi-
686 criteria decision analysis. Knowledge-Based Systems 219.
687 https://doi.org/10.1016/j.knosys.2021.106899
688 Pham BT, Luu C, Phong TV et al. (2021) Flood risk assessment using hybrid artificial intelligence models
689 integrated with multi-criteria decision analysis in Quang Nam Province, Vietnam. Journal of
690 Hydrology 592. https://doi.org/10.1016/j.jhydrol.2020.125815
691 Rodriguez-Galiano V, Ghimire B, Rogan J et al. (2012) An assessment of the effectiveness of a random
692 forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote
693 Sensing 67(1), 93-104. https://doi.org/10.1016/j.isprsjprs.2011.11.002
694 Shen HT, Su J, Liu L (2000) SPH simulation of River Ice Dynamics. Journal of Computational Physics
695 165(2), 752-770. https://doi.org/10.1006/jcph.2000.6639
696 Shen HT, Shen H, Tsai SM (2010) Dynamic transport of river ice. Journal of Hydraulic Research 28(6),
697 659-671.
698 Stefanidis S, Stathis D (2013) Assessment of flood hazard based on natural and anthropogenic factors
699 using analytic hierarchy process (AHP). Natural Hazards 68, 569-585.
700 https://doi.org/10.1007/s11069-013-0639-5
701 Tan X, Zhang W, Ma J et al. (2004) Research on regional assessment of flood risk and regionalization
702 mapping in China. Journal of China Institute of Water Resources and Hydropower Research 2(1),
703 54-64. https://doi.org/10.13244/j.cnki.jiwhr.2004.01.008
704 Tian F, Wang Y, Wang X (2021) Classification and Distribution Characteristics of Ice Flood Risk in the
705 Ningxia-Inner Mongolia Reach of the Yellow River. Yellow River 43(2), 49-52+69.
706 https://doi.org/10.3969/j.issn.1000-1379.2021.02.010

Todorovic P, Zelenhasic E (1970) A Stochastic Model for Flood Analysis. John Wiley & Sons, Ltd 6(6), 1641-1648. https://doi.org/10.1029/WR006i006p01641

Uzuner M, Kennedy J (1974) Hydraulics and Mechanics of River Ice Jams. IOWA INST OF HYDRAULIC RESEARCH IOWA CITY.

Uzuner M, Kennedy J (1976) Theoretical model of river ice jams. Journal of Hydraulic Engineering 102(9), 1365-1383. https://doi.org/10.1061/JYCEAJ.0004618

Wang J, Zhao H (2008) Recent development in simulation of river ice jam. Advances in Water Science 19(4), 597-604. https://doi.org/10.14042/j.cnki.32.1309.2008.04.004

Wang Z, Chen X, Lai C et al. (2013) Flood risk assessment model based on particle swarm optimization rule mining algorithm. Systems Engineeering — Theory & Practice 33(6), 1615-1621.

Wang Z, Lai C, Chen X et al. (2015) Flood hazard risk assessment model based on random forest. Journal of Hydrology 527, 1130-1141. https://doi.org/10.1016/j.jhydrol.2015.06.008

Winsemius H, Aerts J, Beek LV et al. (2015) Global drivers of future river flood risk. Nature Climate Change 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. Natural Hazards 75(S2), 179-197. https://doi.org/10.1007/s11069-014-1308-z

Yang J (2019) ROC Analysis Method Based on K-Nearest Neighbor Classifier. A Dissertation Submitted to Guangdong University of Technology for the Degree of Master,Guangdong University of Technology.

Yang K, Liu Z, Li G et al. (2002) Simulation of ice jam channel. Water Resources and Hydropower Engineering 33(10), 40-47. https://doi.org/10.13928/j.cnki.wrahe.2002.10.013

729 Zbigniew WK, Pińskwar I, Brakenridge GR (2013) Large floods in Europe, 1985–2009. Hydrological
730 Sciences Journal 58(1), 1-7. https://doi.org/10.1080/02626667.2012.745082
731 Zhou C, Wan Q, Huang S et al. (2000) A GIS-based Approach to Flood Risk Zonation. Acta geographica
732 sinica 55(1), 15-24.
733 Zhou Y, Li S, Zhou C et al. (2019) Intelligent Approach Based on Random Forest for Safety Risk
734 Prediction of Deep Foundation Pit in Subway Stations. Journal of Computing in Civil Engineering
735 33(1). https://doi.org/10.1061/(asce)cp.1943-5487.0000796
736 Zou Q, Zhou J, Zhou C et al. (2012) Comprehensive flood risk assessment based on set pair analysis-
737 variable fuzzy sets model and fuzzy AHP. Stochastic Environmental Research and Risk Assessment
738 27, 525-546.
739
740