

Article

A Multi-Source Data Fusion Method for Assessing the Tunnel Collapse Risk Based on the Improved Dempster–Shafer Theory

Bo Wu ¹, Jiajia Zeng ^{2,*}, Ruonan Zhu ², Weiqiang Zheng ² and Cong Liu ^{1,*}¹ School of Civil and Architecture Engineering, East China University of Technology, Nanchang 330013, China² School of Water Resources and Environmental Engineering, East China University of Technology, Nanchang 330013, China

* Correspondence: 13051139728@163.com (J.Z.); 202060055@ecut.edu.cn (C.L.)

Abstract: Collapse is the main engineering disaster in tunnel construction when using the drilling and blasting method, and risk assessment is one of the important means to significantly reduce engineering disasters. Aiming at the problems of random decision-making and misjudgment of single indices in traditional risk assessment, a multi-source data fusion method with high accuracy based on improved Dempster–Shafer evidence theory (D-S model) is proposed in this study, which can realize the accurate assessment of tunnel collapse risk value. The evidence conflict coefficient K is used as the identification index, and the credibility and importance are introduced. The weight coefficient is determined according to whether the conflicting evidence is divided into two situations. The advanced geological forecast data, on-site inspection data and instrument monitoring data are trained by Cloud Model (CM), Gradient Boosting Decision Tree (GBDT) and Support Vector Classification (SVC), respectively, to obtain the initial BPA value. Combined with the weight coefficient, the identified conflict evidence is adjusted, and then the evidence from different sources is fused to obtain the overall collapse risk value. Finally, the accuracy is selected to verify the proposed method. The proposed method has been successfully applied to Wenbishan Tunnel. The results show that the evaluation accuracy of the proposed multi-source information fusion method can reach 88%, which is 16% higher than that of the traditional D-S model and more than 20% higher than that of the single-source information method. The high-precision multi-source data fusion method proposed in this paper has good universality and effectiveness in tunnel collapse risk assessment.

Keywords: tunnel collapse; multi-source data fusion; collapse possibility; risk assessment; machine learning

Citation: Wu, B.; Zeng, J.; Zhu, R.; Zheng, W.; Liu, C. A Multi-Source Data Fusion Method for Assessing the Tunnel Collapse Risk Based on the Improved Dempster–Shafer Theory. *Appl. Sci.* **2023**, *13*, 5606. <https://doi.org/10.3390/app13095606>

Academic Editor: Asterios Bakolas

Received: 21 February 2023

Revised: 27 April 2023

Accepted: 28 April 2023

Published: 1 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Highways are extremely important infrastructure in most countries, ensuring connectivity and development between different regions, especially in mountainous and hilly areas. However, due to the long construction period of highway tunnels and the large disturbance to surrounding rock, there are unpredictable factors and huge safety risks. In the construction process, the drilling and blasting method is widely used in tunnel excavation due to its low cost and strong geological applicability. However, due to many risk factors and complex construction procedures, it is easy to cause collapse accidents [1]. Once the tunnel collapse occurs, it will cause delays in construction, economic losses, and even casualties. Therefore, in order to ensure safe construction during highway tunnel engineering, it is of great significance to accurately evaluate the collapse risk of tunnel construction.

Aiming at the problem of tunnel collapse risk assessment, the main research methods include single information source risk assessment methods such as analytic hierarchy process [2–4], fuzzy comprehensive evaluation method [5], risk matrix method [6,7], network analysis method [8], Bayesian network [9–11] and fault tree method [12,13], as well as multi-source information fusion assessment methods such as neural network [14,15], rough set [16], D-S evidence theory [17] and maximum entropy method [18]. Among them, the

results of using a single information source to analyze the risk of tunnel collapse are deviate slightly from the actual situation. This is because the single-source information cannot reflect the actual construction situation. The evaluation results are inaccurate and cannot provide accurate suggestions for decision makers. The fusion model can better understand the risk factors and greatly improve the accuracy of the prediction results. It is one of the most effective means to study the risk assessment of tunnel collapse. A multi-classifier information fusion model is proposed [19]. Support Vector Machine and D-S evidence theory are used in the model to evaluate the health risk of subway structure under uncertain conditions. It is proved that the fusion model has better robustness and accuracy than the single-classifier model. The safety risk of adjacent buildings during tunnel excavation is perceived through the cloud model and the improved D-S evidence theory. The reliability of the safety risk perception results was tested when the measurement factors were at different deviation levels [20]. The improved multi-source information fusion method is used to evaluate the risk of subsea tunnel excavation models. The results show that the prediction results of the improved algorithm are in good agreement with the actual water inrush phenomenon observed in the model test. Based on the previous research results, the size of the evidence conflict coefficient determines the degree of agreement between the evidence fusion results and the actual results [21]. Wu considered the on-site inspection of experts and the monitoring of on-site instruments. At the same time, a weighted average was used for non-conflict evidence fusion, and traditional D-S fusion was used for high-conflict evidence. However, the advanced geological forecast data are not considered. Due to the complex geological conditions of mountain tunnels, the advanced geological forecast data can directly describe the geological conditions before excavation, which has an advanced expected effect on the construction and can reflect the collapse risk to a certain extent. At the same time, the traditional D-S fusion is used for high conflict evidence, while the traditional D-S evidence theory is very insensitive to high conflict evidence, which leads to inconsistency between the fusion result and the actual result [22].

Through the above analysis, because the geological structure of the mountain tunnel is often more complex, the identification and adjustment of the evidence conflict between different information sources still need further study. In addition, the existing research results are based on multi-source information fusion methods, which fail to consider the combined effects of advanced geological forecast data, on-site inspection data and instrument detection data, and are different from the actual situation, and cannot fully and truly reflect the construction site situation. Based on the improved D-S model, a high-accuracy multi-source data fusion method is proposed relied on the Fujian Wenbisha tunnel project. In this method, the evidence conflict coefficient K is used as the identification index, and the credibility and importance are introduced. The weight coefficient is determined according to whether the conflicting evidence is divided into two situations. The advanced geological forecast data, on-site inspection data and instrument monitoring data are trained by CM (Cloud Model), GBDT (Gradient Boosting Decision Tree) and SVC (Support Vector Classification), respectively, to obtain the initial value of BPA (Basic Probability Assignment). The identified conflict evidence is adjusted by combining the weight coefficient. Finally, the evidence from different sources is fused to obtain the overall collapse risk value, so as to provide reference for the collapse risk assessment and control of mountain tunnel construction.

2. Methodology

In order to improve the accuracy of tunnel collapse risk assessment, a multi-source information fusion assessment method based on the Improved Dempster–Shafer Theory is proposed by combining artificial intelligence models. Figure 1 shows the flow of the method proposed in this paper for assessing tunnel collapse risk. Cloud Models (CM), Gradient Boosting Decision Tree (GBDT) and Support Vector Machines (SVM) are used for the advance geological forecast information, site inspection information and instrument monitoring information, respectively, so as to obtain the tunnel collapse failure probability

from a single information source. Then, the weight coefficient between each piece of evidence is calculated according to whether the conflict coefficient K is greater than the conflict threshold ($\zeta = 0.9$), and the weight coefficient is adjusted and fused to obtain the overall collapse risk assessment result.

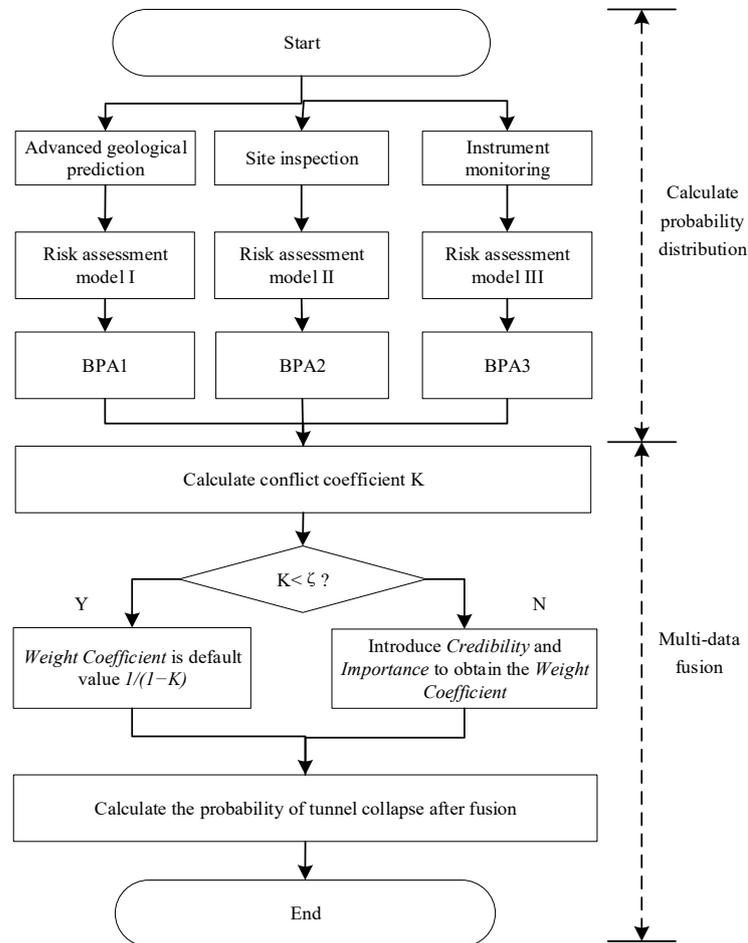


Figure 1. Flowchart of the proposed hybrid method for multi-source data fusion decision.

For advanced geological forecast data, because it constructs a data set through geophysical prospecting results and expert scoring results, CM can map qualitative concepts to quantitative data, so it can be used to process geophysical survey data. For site inspection data, the Gradient Boosting Decision Tree is used to investigate causal relationships between tunnel collapse and its influential variables based upon the risk mechanism analysis and expert scores. For instrument monitoring data, since it has been classified, SVM with the advantage of small sample classification is used for risk assessment. By using improved D-S multi-source information fusion method and adjusting high conflict evidence, useful information of different evidence sources can be extracted, so as to improve the accuracy of collapse risk assessment. A typical hazard concerning the tunnel collapse in the construction of the Fujian Wenbishan Tunnel in China is presented as a case study. The results demonstrate the feasibility of the proposed approach and its application potential.

2.1. Basic Probability Assignment Calculation of Different Evidence Sources

2.1.1. Cloud Model

Let U be a numerical representation of the quantitative domain, and C be a qualitative concept on the domain U . If there is a quantitative value $x \in U$, x is a random realization on C , and the membership $u(x) \in [0, 1]$ of C is a random number with a stable tendency, then $u(x)$ is distributed on the universe U for short as cloud, and each $(x, u(x))$ is referred to as a

cloud droplet [23]. Then x satisfies: (1) $x \in U$, (2) x is a random instantiation of concept C , (3) x satisfies the Formula (1), and the certainty of x belonging to concept C can be obtained by Formula (2).

$$\begin{cases} x \sim N(Ex, En'^2) \\ En' \sim N(En, He^2) \end{cases} \tag{1}$$

$$u(x) = \exp\left(-\frac{(x - Ex)^2}{2(En')^2}\right) \tag{2}$$

In the formula, Ex refers to the expected value of the spatial distribution of cloud droplets in the domain of discourse, which is the point that can best represent the qualitative concept. En represents a measure of the uncertainty of qualitative concepts, which can be used to describe the span of clouds and reflect the dispersion of cloud droplets. Both Ex and En can be fitted from the training set data. He is a measure of entropy En uncertainty, which represents the degree of entropy dispersion. According to Reference [24], He is 0.004 in this paper.

When modeling multiple categories on the same attribute, the generated multiple Cloud Models may appear to be overlapped together, and the degree of overlap reflects the accuracy of the model. Let $C_1(Ex_1, En_1, He_1)$ and $C_2(Ex_2, En_2, He_2)$ be two orthogonal Cloud Models. Then the intersection degree of both is defined as in Equation (3):

$$S(C_1, C_2) = \begin{cases} \frac{3(En_1+En_2)-|Ex_1-Ex_2|}{3(En_1+En_2)+|Ex_1-Ex_2|}, & 3(En_1 + En_2) - |Ex_1 - Ex_2| > 0 \\ 0, & 3(En_1 + En_2) - |Ex_1 - Ex_2| \leq 0 \end{cases} \tag{3}$$

In the formula, En_1 and En_2 are the entropy of the two models, and Ex_1 and Ex_2 are the expectation of the two models. It can be seen that when $3(En_1 + En_2) - |Ex_1 - Ex_2| \leq 0$, the orthogonal Cloud C_1 and the orthogonal Cloud C_2 have no overlapping part, and the degree of overlapping of the two is 0. There is a crossover part between Orthogonal Cloud C_1 and Orthogonal Cloud C_2 when $3(En_1 + En_2) - |Ex_1 - Ex_2| > 0$. The smaller the value of $|Ex_1 - Ex_2|$ is, the larger the overlapping part is. When $Ex_1 = Ex_2$, the expectations of C_1 and C_2 are the same, and the degree of overlap is considered to be 1.

2.1.2. Gradient Boosting Decision Tree

The forward distribution algorithm is adapted from Gradient Boosting Decision Tree [25]. The initial value of the model is set to $F_0(x)$, which is usually a constant. The model of step m is $F_m(x)$, and the newly added classification regression tree can be obtained by minimizing the loss function $L(x)$, which can be calculated using Equation (4).

$$\begin{cases} F_m(x) = F_{m-1}(x) + \alpha_m h_m(x) \\ h_m = \underset{h}{\operatorname{argmin}} \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + h(x_i)) \end{cases} \tag{4}$$

where L is the loss function, $\alpha_m h_m(x)$ is a regular term to prevent overfitting and the value range of α is $(0, 1]$.

The gradient descent method is used for training the optimal model. The negative gradient value of the loss function in the current model $F_{m-1}(x)$ is taken as the direction of gradient descent, as shown in Equation (5).

$$\begin{cases} F_m(x) = F_{m-1}(x) - \alpha_m \sum_{i=1}^N \nabla_F L(y_i, F_{m-1}(x_i)) \\ \alpha_m = \underset{\alpha}{\operatorname{argmin}} \sum_{i=1}^N L(y_i, F_{m-1}(x_i) - \alpha \frac{\partial L(y_i, F_{m-1}(x_i))}{\partial F_{m-1}(x_i)}) \\ F_m(x) = F_{m-1}(x) + v \alpha_m h_m(x) \end{cases} \tag{5}$$

In the formula, v represents learning rate. The smaller the learning rate, the more CART is needed, and the final error will be smaller. However, it will also increase training time. Therefore, it is necessary to control the learning rate and the number of CART at the same time to determine a model with fast speed and high precision.

2.1.3. Support Vector Machines

Support Vector Machine is a kind of generalized linear classifier which classifies data by supervised learning, whose output is only 0 or 1. In order to extract the correlation probability from the output of support vector machine, some scholars have proposed a variety of methods. Platt’s method [26,27] is used in the study, which uses the Sigmoid function to map the output to the interval [0, 1], as shown in Equation (6).

$$\begin{cases} P_{ab}(f(x)) = \frac{1}{1+e^{af(x)+b}} \\ f(x) = \text{sign}[(\sum_{i=1}^m \alpha_i y_i K(x_i, x)) + b] \end{cases} \quad (6)$$

In the formula, m is the size of the training data set, α_i is the Lagrange multiplier, $K(x_i, x)$ is the kernel function and b is the threshold parameter based on the training set. The parameters a and b can be obtained by minimizing the negative log-likelihood function of the training instance,

$$\min_{z=(a,b)} F(z) = -\sum_{i=1}^l (t_i \log(p_i) + (1 - t_i) \log(1 - p_i)) \quad (7)$$

$$\begin{cases} t_+ = \frac{N_++1}{N_++2} \\ t_- = \frac{1}{N_-+2} \end{cases} \quad (8)$$

where t_i is new label of class, +1 becomes t_+ , -1 becomes t_- and N_+ and N_- are the points of class 1 and class 2, as showed in Equations (7) and (8).

2.2. Improved D-S Evidence Fusion Collapse Risk Assessment

In the identification framework of Θ , an evidence can be expressed as Equation (9) [28]. In the formula, θ is any real proposition in $P(\Theta)$. The meaning of $(\theta, P_{\theta,j})$ is that the probability mass of the evidence e_j pointing to proposition θ is $P_{\theta,j}$.

$$e_j = \{ \theta, p_{\theta,j} \} | \forall \theta \subseteq \Theta, \sum_{\theta \subseteq \Theta} p_{\theta,j} = 1 \} \quad (9)$$

In the formula, e_j is evidence, and “ $j = 3$ ” represents three sources of evidence, which are advanced geological forecast data, instrument monitoring data and field inspection data. The identification framework $\Theta = \{I, II, III, IV\}$, which represents the risk levels are I, II, III and IV, respectively.

Improved D-S theory in this study combines credibility r_j ($0 \leq r_j \leq 1$) and importance t_j ($0 \leq t_j \leq 1$). Generally speaking, credibility is an objective existence, indicating the ability of e_j to evaluate correctly compared with other evidence, and importance can be subjectively determined according to the information source of the generated evidence which reflects the relative importance of e_j compared with other evidence in the combination of pieces of evidence.

Firstly, the value of credibility needs to be obtained. At present, the measurement methods of uncertainty include conflict measurement and confusion measurement [29,30]. Aggregated uncertainty is used to calculate the confidence level of each evidence learned from the probabilistic classification model, as shown in Equation (10). In the formula, $P_{\theta,i}$ is the trust of focus element θ_i , and n is the total number of focus elements, which is 4 in this study.

$$r_j = 1 + \frac{\sum_{\theta \in \Theta} p_{\theta,j} \log_2 p_{\theta,j}}{\log_2 n} \quad (10)$$

In the formula, r_j is the degree of trust. n is the total number of focal elements, which is 4 in this paper, indicating the number of risk levels in the identification framework. $P_{\odot,j}$ denotes the basic probability distribution value (BPA) corresponding to the risk level.

Secondly, the value of importance needs to be calculated. In the environment of model learning and verification, the importance rating is directly determined by its contribution to the accuracy of the model. Minimum class-wise F1-score can limit the contribution of low-quality evidence. So, the importance of each evidence is set to $\min\{\text{class-wise F1-scores}\}$ in the probability classifier, which combines precision and recall, as shown in Equation (11).

$$\begin{cases} t_j = \min\left\{\frac{P_j \cdot S_j}{P_j + S_j}\right\} \\ P_j = \frac{TP_i}{TP_i + FP_i} \\ S_j = \frac{TP_i}{TP_i + FN_i} \end{cases} \quad (11)$$

In the formula, t_j denotes the importance. TP_i means that both the true value and the predicted value are Positive. FP_i indicates that the true value is negative but the predicted value is positive. FN_i indicates that the true value is positive but the predicted value is negative.

Finally, in order to obtain non-high-conflict evidence, the original BPA needs to be adjusted by credibility and importance to obtain BPA used for evidence fusion. The adjusted evidence is fused in pairs, and the overall evidence fusion result is obtained after normalization, as shown in Equation (12).

$$\begin{cases} p_{i,e(3)} = \frac{\tilde{m}_{i,e(3)}}{\sum_{i \in \{I,II,III,IV\}} \tilde{m}_{i,e(3)}} \\ \tilde{m}_{i,e(3)} = (1 - r_3)\tilde{m}_{i,e(2)} + (1 - r_2)(1 - r_1)\tilde{m}_{i,3} + \tilde{m}_{i,e(2)}\tilde{m}_{i,3} \\ \tilde{m}_{i,e(2)} = (1 - r_2)\tilde{m}_{i,1} + (1 - r_1)\tilde{m}_{i,2} + \tilde{m}_{i,1}\tilde{m}_{i,2} \\ \tilde{m}_{i,j} = \frac{t_j p_{i,j}}{1 + t_j - r_j} \end{cases} \quad (12)$$

In the formula, $\tilde{m}_{i,j}$ represents the BPA value after the combination of trust and importance, that is, the BPA corresponding to the i focal element of evidence j . t_j and r_j are the importance and credibility corresponding to the evidence of article j , respectively. $\tilde{m}_{i,e(2)}$ represents the BPA value after the fusion of two evidences. $\tilde{m}_{i,1}$ denotes the BPA distribution of the first evidence. $\tilde{m}_{i,2}$ denotes the BPA distribution of the second evidence. $\tilde{m}_{i,3}$ denotes the BPA distribution of the third evidence. $\tilde{m}_{i,e(3)}$ represents the BPA value after the fusion of three evidences. $p_{i,e(3)}$ represents the final result of the three-evidence fusion after normalization.

3. Case Study

3.1. Case Background

Wenbishan Tunnel is a two-lane separated extra-long tunnel, located in Sanming, Fujian, with a left tunnel length of 4786 m and a right tunnel length of 4760 m. In this paper, the right-hand tunnel (YK223 + 375~YK228 + 135) is taken as the object of study, with a 360-m-long section of V-grade surrounding rock and a 1567-m-long section of IV-grade surrounding rock. The geological longitudinal section of the tunnel in the right panel is shown in Figure 2. The geological conditions in the tunnel are complex, and most of the surrounding rocks are residual clay and granite with different degrees of weathering. In the section from YK223 + 473 to YK223 + 728, the joints and fissures are developed, the rock is more broken and the rock is a mosaic fracture structure. During the construction process, it is very easy to cause tunnel collapse. Therefore, there is an urgent need to conduct a collapse risk assessment of this tunnel section to reduce the damage caused by the collapse.

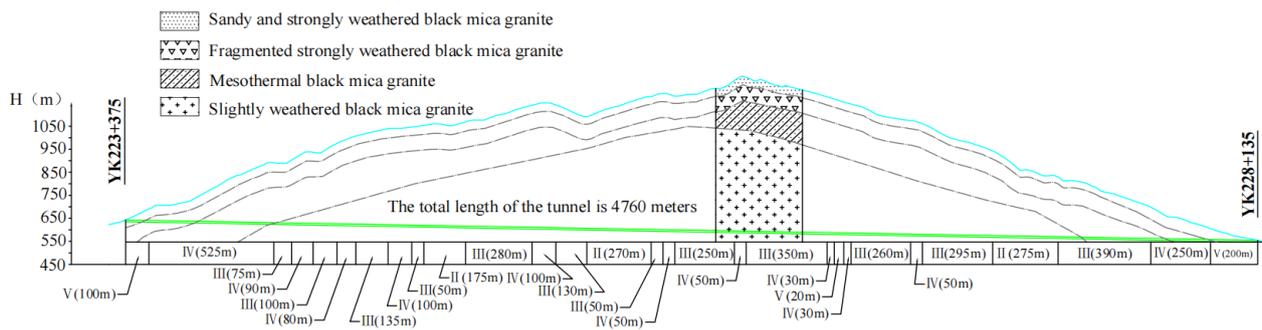


Figure 2. Longitudinal section of the right line of Wenbishan Tunnel. The black line represents the ground line, the blue line represents the surface line, and the green line represents the tunnel.

3.2. Collapse Risk Assessment Based on Advance Geological Forecast Data

According to the Guide to Construction Safety Risk Assessment for Highway Bridge and Tunnel Projects, the risk level of tunnel collapse is divided into four levels: low (level I), moderate (level II), high (level III) and very high (level IV). Combined with the effective data available for advanced geological prediction, the tunnel collapse risk evaluation index system is established, as shown in Table 1. The uniaxial compressive strength of rock is a direct reaction of rock hardness. The smaller the value, the softer the rock and the greater the risk of collapse. The integrity coefficient of rock mass can be obtained by converting the longitudinal wave velocity of rock mass. The smaller the value, the more incomplete the rock and the greater the risk of collapse. The angle between the main structural plane and the tunnel axis describes the combination relationship between the tunnel axis and the main structural plane of the chamber. The larger the angle, the more unstable the chamber and the greater the risk of collapse. The discontinuous structural plane state reflects the state of the chamber control structural plane. The more serious the corrosion, the worse the nature of the filling, the worse the stability of the chamber and the greater the risk of collapse.

Table 1. Indices and criterion for tunnel collapse risk assessment.

Evaluation Indicators	Collapse Level			
	I	II	III	IV
Uniaxial compressive strength of rock/MPa	80~120	30~80	10~30	5~10
Surrounding rock integrity factor	0.75~1	0.45~0.75	0.2~0.45	0~0.2
Angle between the main structural surface and the cave axis/°	80~90	50~80	10~50	0~10
Discontinuous structural surface state	0~0.2	0.2~0.5	0.5~0.8	0.8~1

There are four types of risk levels in the advanced geological forecast datasets, namely Level I, Level II, Level III and Level IV. Each risk level has four attributes, which are uniaxial compressive strength of rock, surrounding rock integrity factor, angle between the main structural surface and the cave axis and discontinuous structural surface state. In this study, a dataset of 100 tunnel collapse cases is collected to form a training dataset for tunnel advance geological forecasting. The four attribute values are used as inputs and the model parameters are obtained after the inverse cloud generator, as shown in Table 2. With the obtained model parameters, a forward cloud generator is created. The test samples are fed into the forward cloud generator to obtain the tunnel collapse risk assessment.

Table 2. Cloud Model parameter values of four detection indicators.

Indicators	I			II			III			IV		
	Ex	En	He	Ex	En	He	Ex	En	He	Ex	En	He
Uniaxial compressive strength of rock/MPa	80	6.67	0.004	60	8.33	0.004	20	3.33	0.004	7.5	0.83	0.004
Surrounding rock integrity factor	0.875	0.042	0.004	0.55	0.05	0.004	0.25	0.042	0.004	0.1	0.033	0.004
Angle between the main structural surface and the cave axis/°	85	1.67	0.004	65	5	0.004	30	6.67	0.004	5	1.67	0.004
Discontinuous structural surface state	0.1	0.033	0.004	0.25	0.05	0.004	0.6	0.05	0.004	0.9	0.033	0.004

In the section YK223 + 473~YK224 + 073, samples are selected every 10 m, and a total of 60 cross-sections are selected as test samples. The Cloud Model for the four risk classes under the four attributes is shown in Figure 3. There are overlaps in level I and level II of the property “uniaxial compressive strength of rock”, level III and level IV of the property “integrity coefficient of surrounding rock”, and level I and level II of the property “state of discontinuous structural surface”. According to Equation (3), the overlapping degree of uniaxial compressive strength of rock is $S_A[(I, II), (II, III), (III, IV)] = [0.38, 0, 0]$, the overlapping degree of surrounding rock integrity factor is $S_B[(I, II), (II, III), (III, IV)] = [0, 0, 0.2]$, the overlapping degree of angle between the main structural surface and the cave axis is $S_C[(I, II), (II, III), (III, IV)] = [0, 0, 0]$ and the overlapping degree of discontinuous structural surface state is $S_D[(I, II), (II, III), (III, IV)] = [0.25, 0, 0]$. It can be seen that the overlap degree is greater than 0, namely $S_A(I, II) = 0.38 > 0$, $S_B(III, IV) = 0.3 > 0$, $S_C(I, II) = 0.25 > 0$. Therefore, there will be some error in predicting the overall collapse level by using the advanced geological forecast data.

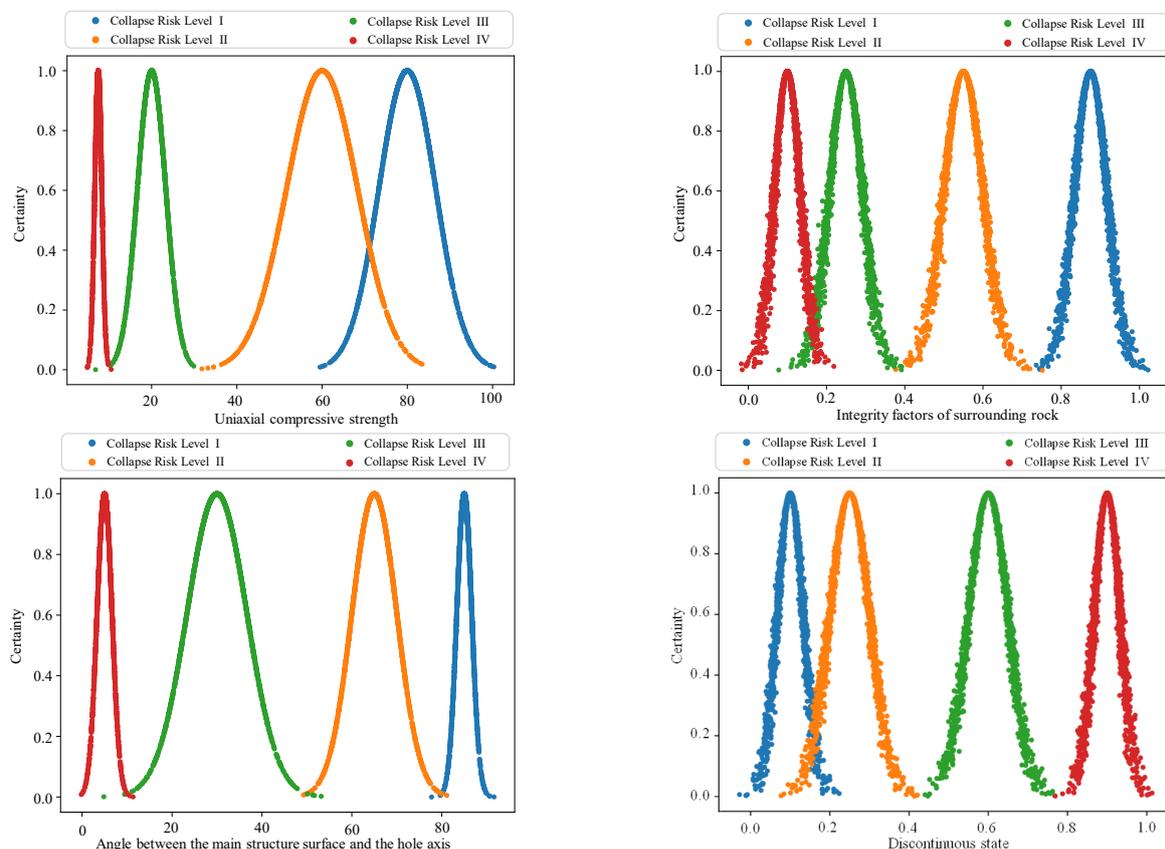


Figure 3. Cloud Model with four types of risk levels under four attributes.

The collapse risk assessment of the test samples is predicted using the forward cloud generator obtained above, whose results are shown in Table 3. The classification accuracy is 62%, which still has a big error. This may be due to the fact that expert field research methods are used to determine the actual collapse risk level in the training set, which may be subject to human judgment errors and lead to slight discrepancies between the assessment results and the actual situation.

Table 3. Risk assessment results based on advanced geological forecast data.

Index	Probability over Level I	Probability over Level II	Probability over Level III	Probability over Level IV	Predicted Label	True Label	Index	Probability over Level I	Probability over Level II	Probability over Level III	Probability over Level IV	Predicted Label	True Label
No. 1	0.35	0.45	0.1	0.1	II	I	No. 31	0	0.01	0.54	0.45	III	IV
No. 2	0.15	0.7	0	0.15	II	II	No. 32	0.4	0.55	0	0.05	II	I
No. 3	0.1	0.2	0.1	0.6	IV	III	No. 33	0.22	0.78	0	0	II	II
No. 4	0.06	0.6	0.2	0.14	II	II	No. 34	0.35	0.5	0.15	0	II	I
No. 5	0.1	0.1	0.65	0.15	III	III	No. 35	0	0.05	0.85	0.1	III	III
No. 6	0	0.18	0.2	0.62	IV	IV	No. 36	0.45	0.01	0.54	0	III	I
No. 7	0.55	0.15	0.2	0.1	I	I	No. 37	0.36	0.1	0.46	0.08	III	I
No. 8	0.2	0	0.7	0.1	III	III	No. 38	0	0.45	0.55	0	III	II
No. 9	0.1	0.9	0	0	II	II	No. 39	0	0.94	0.06	0	II	II
No. 10	0	0.88	0.12	0	II	II	No. 40	0	1	0	0	II	II
No. 11	0.74	0.06	0.1	0.1	I	I	No. 41	0.1	0.9	0	0	II	II
No. 12	0.35	0.03	0.52	0.1	III	I	No. 42	0.86	0.04	0.1	0	I	I
No. 13	0.15	0	0.75	0.1	III	III	No. 43	0	0	0.1	0.9	IV	IV
No. 14	0	0	0.95	0.05	III	III	No. 44	0	0.14	0.86	0	III	III
No. 15	0.4	0.48	0.1	0.02	II	I	No. 45	0.8	0.2	0	0	I	I
No. 16	0.85	0	0.05	0.1	I	I	No. 46	0	0.6	0.4	0	II	III
No. 17	0.7	0	0.3	0	I	III	No. 47	0	0.92	0.08	0	II	II
No. 18	0	0.76	0.04	0.2	II	II	No. 48	0.1	0.2	0.65	0.05	III	III
No. 19	0	0.4	0.6	0	III	II	No. 49	0.48	0.42	0.1	0	I	II
No. 20	0.1	0.1	0.72	0.08	III	III	No. 50	0.1	0.42	0	0.48	IV	II
No. 21	0	0.1	0.9	0	III	III	No. 51	0.6	0.1	0	0.3	I	I
No. 22	0.54	0.46	0	0	I	II	No. 52	0	0.4	0.06	0.54	IV	II
No. 23	0.58	0	0.4	0.02	I	III	No. 53	0	0.95	0	0.05	II	II
No. 24	0.05	0.75	0	0.2	II	II	No. 54	0.05	0.42	0.48	0.05	III	II
No. 25	0	0.53	0.44	0.03	II	III	No. 55	0.03	0.87	0.1	0	II	II
No. 26	0.7	0	0.15	0.15	I	I	No. 56	0.1	0.82	0	0.08	II	II
No. 27	0.75	0.1	0.15	0	I	I	No. 57	0	0	0.14	0.86	IV	IV
No. 28	0.12	0.88	0	0	II	II	No. 58	0.42	0.02	0	0.56	IV	I
No. 29	0	0	0.4	0.6	IV	III	No. 59	0	0.9	0.1	0	II	II
No. 30	0	0.85	0.15	0	II	II	No. 60	0.55	0.45	0	0	I	II

3.3. Collapse Risk Assessment Based on Site Inspection Data

Site inspection evaluation indicators can be divided into four categories: design factors, geological factors, construction factors and management factors. Design factors include excavation span and depth-to-height ratio. Geological factors include surrounding rock grade and bias. Construction factors include the main stiffness of initial support, stratum reinforcement measures, excavation methods and waterproof and drainage measures. Management factors include monitoring and measurement, construction quality qualification, accuracy of geological survey and timeliness of main support. According to the Guide to Highway Bridge and Tunnel Construction Safety Risk Assessment, the above 12 judgment indicators are divided into four levels, as shown in Table 4.

There are four collapse risk levels in the site inspection data set, which are Level I, Level II, Level III and Level IV. The number of features for each risk level is 12. The input of the test set is $\{x_1, \dots, x_{12}\}$. The data label is tunnel collapse risk grade, which can be determined according to expert site investigation. The model output is the failure probability value of tunnel collapse. The predicted result is the collapse risk level corresponding to maximum value of collapse failure, which is compared with the data label values to obtain the model prediction accuracy.

The number of features in the site inspection test set is 12, and the number of samples is 60. Due to the small number of samples, the neural network model is not applicable, so the Gradient Boosting Decision Tree model in traditional machine learning is chosen. The collapse risk assessment model is constructed by calling the GradientBoostingClassifier module in the scikit-learn library. The key parameters include `n_estimators` and `learning_rate`. The grid search method is used to find the optimal hyperparameters. The parameter search range of `n_estimators` is [60, 80, 100, 120, 14]. The parameter search range of `learning_rate` is [0.001, 0.01, 0.1, 1, 1.5, 2].

Table 4. Evaluation index and collapse risk factor grade.

Evaluating Indicator		Factor Level			
		I	II	III	IV
Design factors	Excavation span (m) (x_1)	<7	7~10	10~14	>14
	Depth-to-height ratio (x_2)	>20	15~20	10~15	<10
Geological factors	Surrounding rock grade (x_3)	81~100	61~81	41~60	<40
	Bias (x_4)	<10°	10~25°	25~40°	>40°
Construction factors	Excavation methods (x_5)	CRD	CD	Bench cut method	Full face method
	Stratum reinforcement measures (x_6)	Full section curtain grouting and advance support	Pipe shed support and advance support	Curtain grouting and advance support	Anchor bolts and advance support
	Waterproof and drainage measures (x_7)	76~100	51~75	26~50	0~25
	main stiffness of initial support (x_8)	>2	1~2	0.5~1	<0.5
Management factors	monitoring and measurement (x_9)	>4 times/day	3 times/day	2 times/day	1 times/day
	Construction quality qualification (x_{10})	90~100%	80~90%	70~80%	60~70%
	Accuracy of geological survey (x_{11})	>90%	75~90%	60~75%	<60%
	Timeliness of main support (x_{12})	<30 min	30~60 min	60~120 min	>120 min

An optimization search experiment is conducted for $n_estimators$ and learning rate. The Receiver Operating Characteristic (ROC) Curve is used for model evaluation. The closer the ROC curve is to the vertical axis, the better the model performance is, conversely, the worse the model performance is. In addition, the model can be evaluated by AUC (area under the ROC curve). Macro-averaging and micro-averaging are different methods to achieve classification accuracy calculations, and both can be used to evaluate the model. The larger the value, the larger the area, the better the model performance. The result of model parameters training is shown in Figure 4.

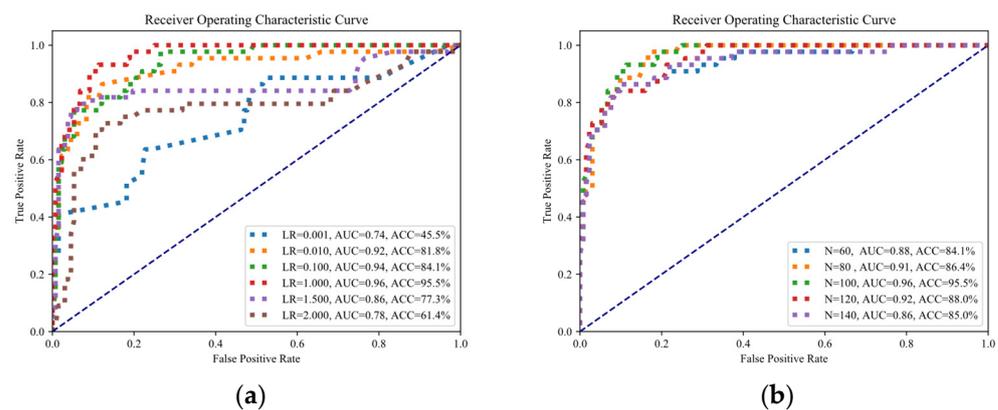


Figure 4. Receiver Operating Characteristic Curve for different $n_estimators$ and learning rates. (a) The receive operating characteristic curve of model when learning_rate is the variable. (b) The receive operating characteristic curve of model when $n_estimators$ is the variable.

Figure 4a shows the results of model training when $n_estimators$ is fixed value and learning_rate is the variable. Figure 4b shows the results of model training when learning_rate is fixed value and $n_estimators$ is the variable. As learning_rate increases, AUC and ACC (Accuracy) first increases and then decreases. As $n_estimators$ increases, AUC and ACC (Accuracy) first increases and then decreases. As can be seen from the figure, The

optimal model is obtained when learning_rate is equal to 1 and n_estimators is equal to 100. AUC and ACC are both at maximum, which are 0.96, and 95.5%. The reason is that the model will overfit if n_estimators and learning_rate are too large, while the model will underfit if n_estimators and learning_rate are too small.

Through the trained model, the above 60 tunnel sections are used as samples for risk assessment, and the results are shown in Table 5. The classification accuracy reaches 56%, which is too low to be used for construction site guidance. The results of on-site inspection come from expert experience, which are greatly influenced by subjectivity. In addition, due to the small data set and uneven distribution of the number during model training, the prediction accuracy of some risk levels is too low, and the overall accuracy is reduced.

Table 5. Risk assessment results based on site inspection data.

Index	Probability over Level I	Probability over Level II	Probability over Level III	Probability over Level IV	Predicted Label	True Label	Index	Probability over Level I	Probability over Level II	Probability over Level III	Probability over Level IV	Predicted Label	True Label
No. 1	0.4	0.2	0.1	0.3	I	I	No. 31	0	0.15	0	0.85	IV	IV
No. 2	0.25	0.55	0.1	0.1	II	II	No. 32	0	0	1	0	III	I
No. 3	0	0.2	0.8	0	III	III	No. 33	0.12	0.88	0	0	II	II
No. 4	0.08	0.52	0.3	0.1	II	II	No. 34	0.05	0.95	0	0	II	I
No. 5	0.3	0.4	0	0.3	II	III	No.35	0	0.1	0.8	0.1	III	III
No. 6	0.1	0.3	0.5	0.1	III	IV	No. 36	0.44	0.56	0	0	II	I
No. 7	0.9	0.1	0	0	I	I	No. 37	0.45	0	0.55	0	III	I
No. 8	0.1	0.1	0.6	0.2	III	III	No. 38	0	0	0.85	0.15	III	II
No. 9	0	0	0.96	0.04	III	II	No. 39	0	0.84	0.16	0	II	II
No. 10	0.18	0.82	0	0	II	II	No. 40	0.12	0.78	0.1	0	II	II
No. 11	0.82	0	0.03	0.15	I	I	No. 41	0	0.74	0.16	0.1	II	II
No. 12	0.4	0.6	0	0	II	I	No. 42	0.84	0.14	0.02	0	I	I
No. 13	0	0.1	0.9	0	III	III	No. 43	0.1	0.8	0.1	0	II	IV
No. 14	0	0	1	0	III	III	No. 44	0.15	0	0.85	0	III	III
No. 15	0.05	0.2	0.1	0.65	IV	I	No. 45	0.86	0	0.04	0.1	I	I
No. 16	0.7	0.3	0	0	I	I	No. 46	0	0	0.35	0.65	IV	III
No. 17	0	0.55	0.45	0	II	III	No. 47	0.92	0	0.08	0	I	II
No. 18	0	0.85	0.15	0	II	II	No. 48	0	0	0.85	0.15	III	III
No. 19	0.2	0	0	0.8	IV	II	No. 49	0.58	0	0.42	0	I	II
No. 20	0.7	0.1	0.1	0.1	I	III	No. 50	0	0.4	0.6	0	III	II
No. 21	0.1	0	0.9	0	III	III	No. 51	0.8	0	0.2	0	I	I
No. 22	0.56	0.44	0	0	I	II	No. 52	0	0.48	0.52	0	III	II
No. 23	0	0.6	0.4	0	II	III	No. 53	0	0.9	0.1	0	II	II
No. 24	0	0.86	0	0.14	II	II	No. 54	0	0	0.25	0.75	IV	II
No. 25	0	0.08	0.4	0.52	IV	III	No. 55	0.1	0.8	0	0.1	II	II
No. 26	0.85	0.05	0	0.1	I	I	No. 56	0	1	0	0	II	II
No. 27	0.75	0.15	0.1	0	I	I	No. 57	0	0	0.25	0.75	IV	IV
No. 28	0	0.75	0	0.25	II	II	No. 58	0.45	0.01	0	0.54	IV	I
No. 29	0.75	0.1	0.15	0	I	III	No. 59	0.2	0.7	0.1	0	II	II
No. 30	0.05	0.95	0	0	II	II	No. 60	0.6	0.4	0	0	I	II

3.4. Collapse Risk Assessment Based on Instrument Monitoring Data

The instrument monitoring data include surface settlement, vault settlement displacement and horizontal convergence displacement in the shallow buried section, which reflects the stability of the tunnel support after the initial lining. The effect of surface settlement for deep tunnels is often not considered, and only the vault settlement displacement and horizontal convergence displacement are selected for collapse risk analysis. According to the national standard for Technical Specification for Highway Tunnel Construction (JTG/T 3660-2020), and Technical Specification for Monitoring and Measurement of Highway Tunnels (DB 35/T 1067-2010), the daily change rate and cumulative deformation are taken as indicators. The farther the monitoring point is from the palm surface, the larger the accumulated displacement limit value is, so the cumulative displacement value needs to be multiplied by a factor ζ . These two judgment indicators are divided into four levels, as shown in Table 6. The accumulated displacement needs to be multiplied by a factor ζ according to the distance between the measurement point and the palm surface. According to the Standard for Technical Specification for Monitoring and Measurement of Highway Tunnels (DB 35/T 1067-2010), the relationship between ζ and D is shown in Table 7, where B is the span of the excavated tunnel.

The instrument monitoring datasets have four types of risk levels, which are Level I, Level II, Level III and Level IV. The number of features for each risk level is 2, which are vault settlement displacement and horizontal convergence displacement. The inputs to the test set are the monitored values of vault settlement and horizontal convergence. The tunnel collapse level is the output of model. Then, according to the most unfavorable principle,

the larger collapse risk level is used as the label. The output is the tunnel collapse failure probability value. The collapse level corresponding to the maximum value of collapse failure probability is the collapse level predicted by the model. The model prediction accuracy is obtained by comparing the model prediction values with the labeled values.

Table 6. Classification of monitoring measurement data.

Tunnel Collapse Level	Daily Deformation (mm/d)	Cumulative Deformation (mm)
I (Safe)	$0 \leq x < 2$	$0 \leq y < 50$
II (Deformation is formed)	$2 \leq x < 5$	$50 \leq y < 100$
III (Small-scale collapse)	$5 \leq x < 10$	$100 \leq y < 200$
IV (Large-scale collapse)	$10 \leq x < 20$	$200 \leq y < 300$

Table 7. Coefficient (ζ) of cumulative deformation (y).

Distance from Monitoring Point to Palm Face (D)	1B	2B	3B	4-6B
ζ	0.5	0.75	0.85	1

A collapse risk assessment model is built based on the scikit-learn machine learning library. The number of features in the instrument monitoring datasets is 2. The test set sample is 60 tunnel cross-sections in Section 3.2. The supervised learning model is chosen because the samples need to be trained with labels. In this paper, a Support Vector Machine with the key parameters of kernel function, kernel function parameters and penalty parameters is selected. Based on the fact that both the number of features and the number of samples are small, and the number of features is much smaller than the number of samples, Radial Basis Function (RBF) is selected as the kernel function. The key parameters are the penalty parameter C and the kernel function parameter gamma. A grid search method to find the optimal hyperparameters (C, gamma). The parameter optimization range of C is 0.05, 0.1, 0.15 and 0.2. The parameter optimization range of gamma is 0.001, 0.002, 0.004 and 0.006. The classification probability distribution chart of tunnel collapse risk level is shown in Figure 5.

Figure 5a shows the classification probability distribution chart when gamma is 0.004, and C are 0.05, 0.1, 0.15 and 0.2, respectively. The result shows the relationship between classification probability and color. The darker the color is, the lower the classification probability value is. It can be seen that there is a point where it has clearly deviated from the other points when C is 0.05, and the level is II. However, all points are gathered together when C is 0.1, 0.15, 0.2, and the level is II. When C is 0.05, 0.1, 0.15 and 0.2, the model classification accuracies are 63.6%, 68.3%, 68.3% and 68.3%, respectively. This is due to the fact that the larger the parameter C is, the easier the model is to be overfitted. From the above analysis, the optimal value of C is 0.1. Figure 5b shows the classification probability distribution chart when C is 0.1, and gamma are 0.001, 0.002, 0.004 and 0.006, respectively. It can be seen that there are two points where it has clearly deviated from the other points when gamma is 0.001, and the level is II. There is one point where it has clearly deviated from the other points when gamma is 0.002, and the level is II. However, all points are gathered together when gamma is 0.004, 0.006, and the level is II. When gamma is 0.001, 0.002, 0.004 and 0.006, the model classification accuracies are 63.6%, 68.3%, 68.3% and 68.3%, respectively. This is due to the fact that the larger the parameter gamma is, the easier the model is to be overfitted. From the above analysis, the optimal value of gamma is 0.004. In summary, the optimal parameter C is 0.1, and the optimal parameter gamma is 0.004.

The collapse risk assessment of the test samples based on instrument monitoring data is predicted, and the results are shown in Table 8, which can obtain risk assessment accuracy. Obviously, a large part of the prediction results deviates from the true value. This is due to the error of monitoring data caused by the operation error of construction personnel, or the judgment error caused by the fact that the vault settlement displacement and horizontal convergence displacement cannot fully reflect the surrounding rock support state.

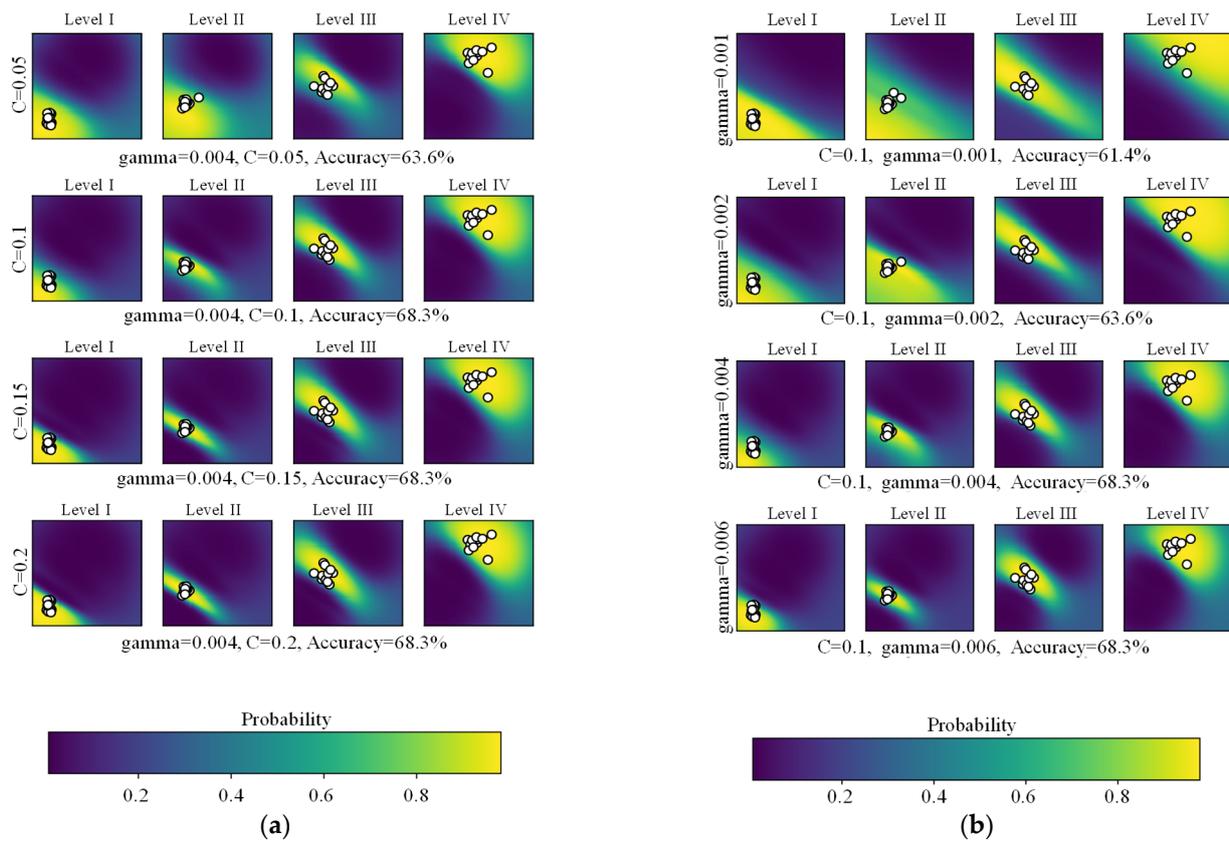


Figure 5. Classification probability distribution for different parameters. (a) C are 0.05, 0.1, 0.15 and 0.2. (b) gamma are 0.001, 0.002, 0.004 and 0.006.

Table 8. Risk assessment results based on instrument monitoring data.

Index	Probability over Level I	Probability over Level II	Probability over Level III	Probability over Level IV	Predicted Label	True Label	Index	Probability over Level I	Probability over Level II	Probability over Level III	Probability over Level IV	Predicted Label	True Label
No. 1	0.5	0.3	0	0.2	I	I	No. 31	0	0	0	1	IV	IV
No. 2	0	0.3	0.6	0.1	III	II	No. 32	0.35	0.65	0	0	II	I
No. 3	0.2	0.6	0.1	0.1	II	III	No. 33	0	0.75	0.25	0	II	II
No. 4	0.04	0.1	0.5	0.36	III	II	No. 34	0.4	0.55	0	0.05	II	I
No. 5	0	0.1	0.4	0.5	IV	III	No. 35	0	0	0.9	0.1	III	III
No. 6	0.02	0.1	0.28	0.6	IV	IV	No. 36	0.44	0	0.56	0	III	I
No. 7	0.7	0.2	0.1	0	I	I	No. 37	0.46	0	0.04	0.5	IV	I
No. 8	0	0	0.86	0.14	III	III	No. 38	0	1	0	0	II	II
No. 9	0.1	0.9	0	0	II	II	No. 39	0.25	0.58	0.17	0	II	II
No. 10	0	0.86	0.14	0	II	II	No. 40	0	0.78	0	0.22	II	II
No. 11	1	0	0	0	I	I	No. 41	0.05	0.95	0	0	II	II
No. 12	0.4	0	0.6	0	III	I	No. 42	0.4	0.6	0	0	II	I
No. 13	0	0.32	0.68	0	III	III	No. 43	0	0	0.2	0.8	IV	IV
No. 14	0.1	0.1	0.7	0.1	III	III	No. 44	0	0.2	0.6	0.2	III	III
No. 15	0.4	0.06	0.54	0	III	I	No. 45	0.9	0.1	0	0	I	I
No. 16	0.65	0.05	0.2	0.1	I	I	No. 46	0	0.6	0.4	0	II	III
No. 17	0	0.1	0.9	0	III	III	No. 47	0	0.85	0	0.15	II	II
No. 18	0.2	0.75	0.05	0	II	II	No. 48	0	0.25	0.75	0	III	III
No. 19	0	1	0	0	II	II	No. 49	0.55	0.45	0	0	I	II
No. 20	0	0	0.95	0.05	III	III	No. 50	0	0.4	0	0.6	IV	II
No. 21	0	0.2	0.8	0	III	III	No. 51	0.75	0.1	0.1	0.05	I	I
No. 22	0.6	0.35	0	0.05	I	II	No. 52	0.56	0.44	0	0	I	II
No. 23	0	0	1	0	III	III	No. 53	0.2	0.7	0	0.1	II	II
No. 24	0	0.68	0	0.32	II	II	No. 54	0	0.85	0.15	0	II	II
No. 25	0	0.58	0.42	0	II	III	No. 55	0	0.9	0.1	0	II	II
No. 26	0.8	0	0	0.2	I	I	No. 56	0.1	0.6	0.1	0.2	II	II
No. 27	0.6	0.3	0.1	0	I	I	No. 57	0	0	0.3	0.7	IV	IV
No. 28	0	0.85	0.1	0.05	II	II	No. 58	0.42	0.1	0	0.48	IV	I
No. 29	0	0	1	0	III	III	No. 59	0.1	0.8	0.1	0	II	II
No. 30	0.05	0.65	0	0.3	II	II	No. 60	0.7	0.25	0.05	0	I	II

3.5. Collapse Risk Assessment Based on Multi-Source Data Fusion Method

In order to solve the problem of unreliable evaluation results of a single information source, the improved D-S theory is used to fuse multi-source data. This method combines the different results of the above three single evaluation models. The importance value is

chosen based on the classification measures of each model by the rule (min{F1-scores}). The classification performance for the above three models is summarized by Equation (11). The result is shown in Table 9. According to the rule (min{classwise F1-scores}), the importance values of each model in the fusion process are [0.5; 0.33; 0.57]. To illustrate the fusion process, six test samples were selected for analysis. The three-source information fusion training results are compared with three single-source training results, as shown in Table 10. In the table, bold values represent the probability of the maximum risk level obtained by the model. The following conclusions can be obtained:

- (1) The multi-source information fusion model has good fault tolerance. The fusion model can correct the wrong classifier result by the important rating and reliability. Taking the test sample section No. 5 as an example, the collapse failure probability value obtained by the three single-information source model are $p_{\theta,1} = [0.1, 0.1, 0.65, 0.15]$, $p_{\theta,2} = [0.3, 0.4, 0, 0.3]$ and $p_{\theta,3} = [0, 0.1, 0.4, 0.5]$. According to Equation (10), the credibility can be obtained as $[r_1, r_2, r_3] = [0.733, 0.685, 0.707]$. It can be seen from Table 9 that the importance value of the fusion model is $[t_1, t_2, t_3] = [0.5, 0.33, 0.57]$. According to Equation (12), the value of BPA, which is combined credibility and importance rating, can be obtained as

$$\begin{bmatrix} \tilde{m}_{I,1} & \tilde{m}_{II,1} & \tilde{m}_{III,1} & \tilde{m}_{IV,1} \\ \tilde{m}_{I,2} & \tilde{m}_{II,2} & \tilde{m}_{III,2} & \tilde{m}_{IV,2} \\ \tilde{m}_{I,3} & \tilde{m}_{II,3} & \tilde{m}_{III,3} & \tilde{m}_{IV,3} \end{bmatrix} = \begin{bmatrix} 0.068 & 0.068 & 0.438 & 0.109 \\ 0.000 & 0.297 & 0.243 & 0.135 \\ 0.000 & 0.153 & 0.102 & 0.256 \end{bmatrix}.$$

Firstly, evidence e_1 and evidence e_2 are fused, and the probability distribution result after fusion is $[\tilde{m}_{I,e(2)}, \tilde{m}_{II,e(2)}, \tilde{m}_{III,e(2)}, \tilde{m}_{IV,e(2)}] = [0.020, 0.119, 0.299, 0.083]$. Then it is fused with evidence e_3 , and the probability distribution result after fusion is $[\tilde{m}_{I,e(3)}, \tilde{m}_{II,e(3)}, \tilde{m}_{III,e(3)}, \tilde{m}_{IV,e(3)}] = [0.006, 0.068, 0.133, 0.067]$. After normalization, the final BPA result after fusion is $[P_{I,e(3)}, P_{II,e(3)}, P_{III,e(3)}, P_{IV,e(3)}] = [0.1, 0.2, 0.4, 0.3]$. The actual collapse level is Level III. As shown in Table 10, wrong conclusions are given by traditional D-S model. The reason is that the improved D-S model takes into account the importance and credibility, which allows e_1 to correct e_2 and e_3 . Through the above analysis, the traditional D-S model is highly sensitive to high conflict evidence when there is a high conflict of evidence, so its performance is poor. In contrast, the improved D-S model, which can be corrected for incorrect classifiers based on the importance rating and confidence of the correct classifier, has better fault tolerance.

- (2) The multi-source information fusion method proposed in this paper simultaneously considers information from three sources: advance geological forecast data, site inspection data and instrument monitoring data, which provide a more comprehensive understanding of tunnel collapse risk, thus reducing data uncertainty and improving the accuracy of assessment. Compared with the single information source risk assessment method, the multi-source information fusion assessment has higher accuracy.
- (3) When the evaluation results of three single information sources has high conflict (e.g., Tunnel section No. 5), the fusion result of the improved D-S theory is better than the traditional D-S theory. The traditional D-S theory accumulates consensus support only and rejects a proposition completely if it is opposed by any evidence, no matter what support it may acquire from any other evidence. As a result, when three kinds of single information evaluation give high conflict results, the traditional D-S theory will give a fusion result contrary to common sense. The improved method has high accuracy when merging high conflict information sources because of considering the importance rating and credibility.

Table 9. Classifier training results.

Evidence	Level	Precision	Recall	Specificity	F1-Score	Support
E1	I	0.62	0.5	0.89	0.55	16
	II	0.74	0.68	0.83	0.71	25
	III	0.56	0.6	0.84	0.58	15
	IV	0.38	0.75	0.91	0.5	4
E2	I	0.6	0.56	0.86	0.58	16
	II	0.68	0.6	0.8	0.64	25
	III	0.53	0.54	0.84	0.53	15
	IV	0.25	0.5	0.89	0.33	4
E3	I	0.67	0.5	0.91	0.57	16
	II	0.75	0.72	0.83	0.73	25
	III	0.69	0.73	0.89	0.71	15
	IV	0.5	1	0.93	0.67	4

Table 10. Fusion results of the probability over classes for six samples.

Sample Index	Classifier Outputs	Probability over Level I	Probability over Level II	Probability over Level III	Probability over Level IV	Predicted Label	True Label	Description
No. 1	E1	0.35	0.45	0.1	0.1	II	I	K = 0.897, t = [0.5, 0.33, 0.6], r = [0.790, 0.707, 0.671]
	E2	0.4	0.2	0.1	0.3	I	I	
	E3	0.5	0.3	0	0.2	I	I	
	D-S	0.68	0.26	0	0.06	I	I	
	Improved D-S	0.5	0.3	0	0.2	I	I	
No. 2	E1	0.15	0.7	0	0.15	II	II	K = 0.883, t = [0.5, 0.33, 0.6], r = [0.741, 0.790, 0.791]
	E2	0.25	0.55	0.1	0.1	II	II	
	E3	0	0.3	0.6	0.1	III	II	
	D-S	0	0.9	0	0.1	II	II	
	Improved D-S	0	0.8	0.2	0	II	II	
No. 3	E1	0.1	0.2	0.1	0.6	IV	III	K = 0.968, t = [0.5, 0.33, 0.6], r = [0.763, 0.834, 0.892]
	E2	0	0.2	0.8	0	III	III	
	E3	0.2	0.6	0.1	0.1	II	III	
	D-S	0	0.75	0.25	0	II	III	
	Improved D-S	0	0.33	0.67	0	III	III	
No. 4	E1	0.06	0.6	0.2	0.14	II	II	K = 0.862, t = [0.5, 0.33, 0.6], r = [0.801, 0.735, 0.834]
	E2	0.08	0.52	0.3	0.1	II	II	
	E3	0.04	0.1	0.5	0.36	III	II	
	D-S	0	0.9	0.1	0	II	II	
	Improved D-S	0	0.8	0.2	0	II	II	
No. 5	E1	0.1	0.1	0.65	0.15	III	III	K = 0.9735, t = [0.5, 0.33, 0.6], r = [0.733, 0.685, 0.707]
	E2	0.3	0.4	0	0.3	II	III	
	E3	0	0.1	0.4	0.5	IV	III	
	D-S	0	0.15	0	0.85	IV	III	
	Improved D-S	0.1	0.2	0.4	0.3	III	III	
No. 6	E1	0	0.18	0.2	0.62	IV	IV	K = 0.874, t = [0.5, 0.33, 0.6], r = [0.786, 0.779, 0.834]
	E2	0.1	0.3	0.5	0.1	III	IV	
	E3	0.02	0.1	0.28	0.6	IV	IV	
	D-S	0	0	0.2	0.8	IV	IV	
	Improved D-S	0	0.05	0.15	0.8	IV	IV	

4. Discussion

There is no doubt that the single-source information assessment method can also assess the risk of tunnel collapse. However, a single source of information does not fully reflect the environment of tunnel construction, resulting in a certain deviation and low accuracy of the assessment results. In order to compare the single information source method and the multi-source information fusion method, the single information source model,

the traditional multi-source information fusion method and the improved multi-source information fusion method are used to evaluate the collapse risk. The risk assessment of 60 tunnel sections in Wenbishan Tunnel is conducted, and the evaluation results of different models are shown in Figure 6. In order to illustrate that the proposed method can improve the high conflict evidence, the high conflict risk of YK223 + 473 tunnel section in the fault zone is evaluated and analyzed. The two-dimensional scatter plot of traditional evidence conflict and improved model evidence conflict is shown in Figure 7. The following conclusions can be obtained:

- (1) Figure 6 is the confusion matrix diagram of different models. The ordinate represents the test results, the abscissa represents the real results and the numbers in the diagram represent the number of tunnel sections. In Figure 6a, the total number of tunnel sections is 60, which is the sum of all tunnel sections in the figure. In the figure, the number 8 indicates that the number of tunnel sections with 'risk prediction grade I and actual grade I' is 8. Similarly, the number 4 indicates that the number of tunnel sections with 'risk prediction grade II but actual grade I' is 4. Therefore, the number of tunnel sections with the same prediction level and the actual level can be obtained by adding the numbers of the diagonals in the figure, that is, the number of tunnel sections predicted by the model is correct. The ratio of the number of tunnel sections to the total number of tunnel sections can obtain the prediction accuracy of the model. The accuracy of Advanced Geological Prediction (Figure 6a) is 62% $((8 + 17 + 9 + 3)/(8 + 4 + 3 + 1 + 3 + 17 + 3 + 2 + 2 + 2 + 9 + 2 + 0 + 0 + 1 + 3))$. Similarly, the accuracy of Field Inspection (Figure 6b) is 56%, the accuracy of Instrument Monitoring (Figure 6c) is 68%, the accuracy of Traditional D-S (Figure 6d) is 73% and the accuracy of Improved D-S (Figure 6e) is 88%.
- (2) The accuracy of the single-source information evaluation method (Figure 6a–c) is less than 70%, which cannot provide accurate decision-making suggestions for construction. The single information source method does not fully consider the risk factors of collapse, making the assessment slightly biased. The accuracy of the traditional D-S evidence fusion method is 73% (Figure 6d), which is slightly higher than that of the single-source information evaluation method, but the accuracy is still low, and it cannot provide accurate guidance for the construction site. As shown in Figure 6e, the multi-source information fusion method has a high evaluation accuracy (88%). This is because the multi-source information model comprehensively considers the advanced geological forecast, on-site inspection and monitoring data, making the evaluation model closer to the actual situation. At the same time, the proposed method considers the importance and credibility, ameliorating the high conflict evidence, making full use of the available information, and improving the accuracy of the evaluation results.
- (3) The traditional D-S method and the proposed method are used to evaluate the risk of section YK223 + 473 respectively. Taking the conflict threshold as the reference value, the traditional D-S method (Figure 7a) has some evidence conflict values higher than the threshold and are more discrete. After the proposed method (Figure 7b) is identified and adjusted, the excess evidence conflict has been greatly improved, which is basically below the threshold. In addition, the adjusted evidence conflict values are relatively concentrated and less discrete. The analysis results further show that the proposed method not only has higher recognition accuracy, but also can effectively improve the high evidence conflict.

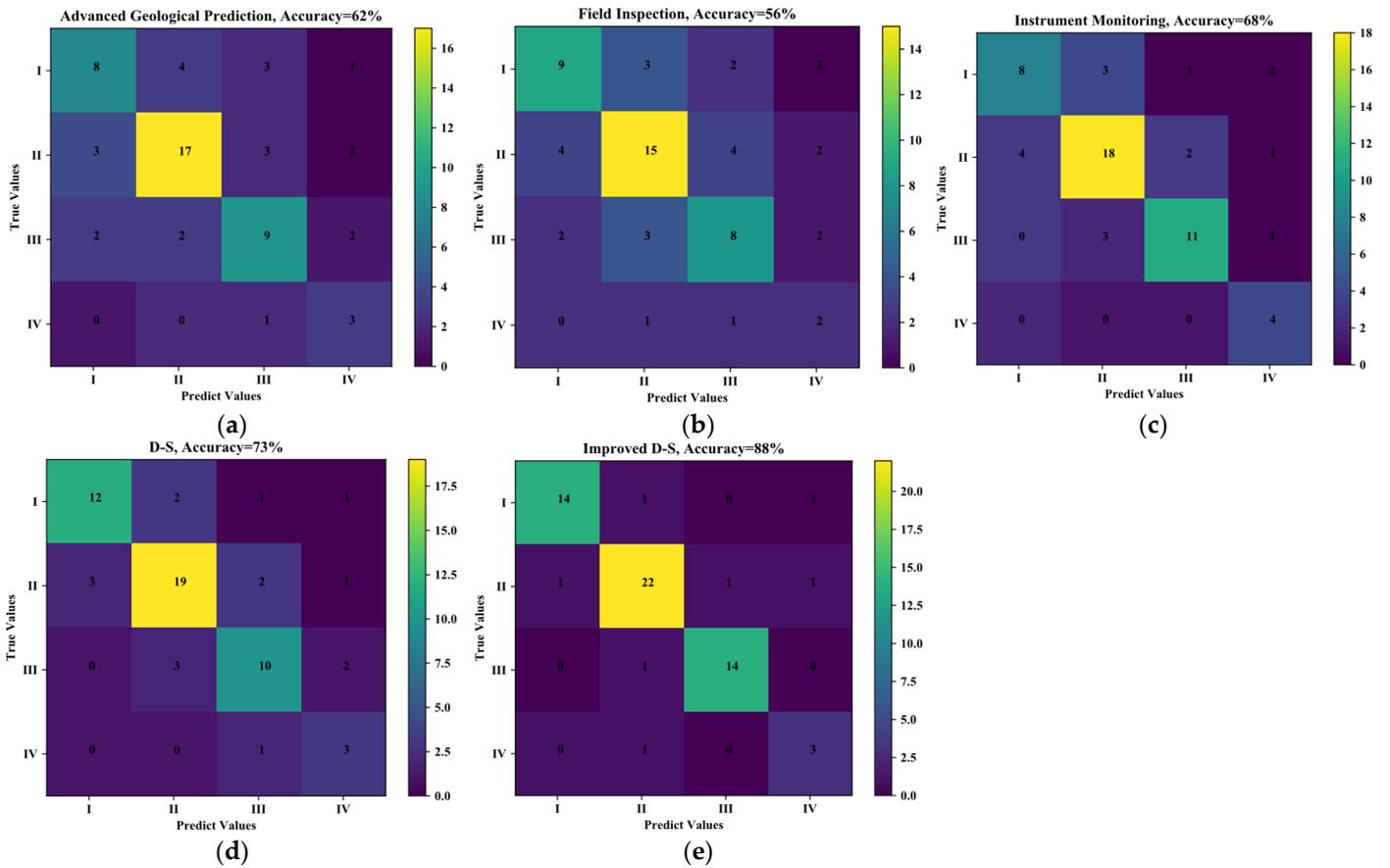


Figure 6. Confusion matrix diagram of different models. (a) Advanced Geological Prediction. (b) Field Inspection. (c) Instrument Monitoring. (d) D-S. (e) Improved D-S.

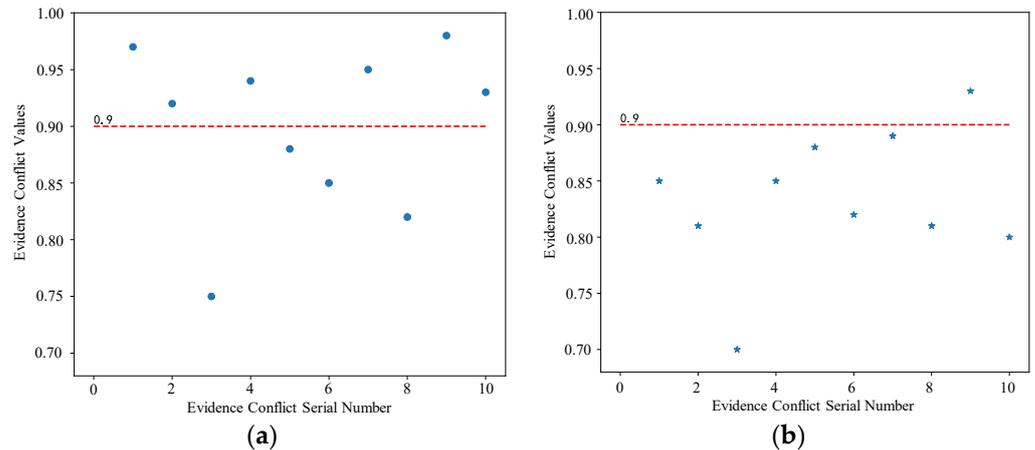


Figure 7. Scattered point distribution of three evidence conflict values. (a) Traditional D-S model. (b) Improved D-S model. The red line represents the conflict threshold. The dots and stars represent evidence conflict value.

5. Conclusions and Future Work

Based on the improved D-S model, a high-accuracy multi-source data fusion method is proposed in this research, which can achieve accurate assessment of tunnel collapse risk. In this method, the evidence conflict coefficient K is used as the identification index, and the credibility and importance are introduced. The weight coefficient is determined according to whether the conflicting evidence is divided into two situations. The advanced geological forecast data, on-site inspection data and instrument monitoring data are trained by CM,

GBDT and SVC, respectively, to obtain the initial BPA value. Finally, the identified conflict evidence is adjusted by combining the weight coefficient, and the overall collapse risk value is obtained by fusing the evidence from different sources. The methods developed in this paper have the following innovations and capabilities:

- (1) It can synthesize multi-source information to obtain a more accurate result for tunneling collapse risk assessment. Due to many influencing factors, the tunneling collapse risk assessment is a multi-attribute decision making problem. Single-source assessment methods have difficulty fully considering all risk factors, resulting in biased prediction results. The performance of the fused model is better than the single information sources model with higher precision.
- (2) Compared with the traditional D-S theory, the improved method has more advantages in dealing with high conflict information. When the risk assessment results of three single information sources are inconsistent, the improved fusion model considers the importance rating and credibility of the assessment results, which improves the accuracy of the final assessment results.

The method proposed in this paper still has limitations. Experts still need to participate in the entire assessment process, which means that a truly fully automated risk assessment has not been achieved. The tunnel collapse training data set is still very small, and a big data system needs to be developed nationwide. In addition, this method cannot predict the risk status of the next construction process, and further research is needed.

Author Contributions: Conceptualization, B.W., J.Z. and R.Z.; methodology, J.Z., R.Z. and W.Z.; software, J.Z., R.Z. and W.Z.; validation, J.Z., R.Z. and W.Z.; formal analysis, J.Z., R.Z. and W.Z.; investigation, J.Z., R.Z. and W.Z.; resources, J.Z., R.Z. and W.Z.; writing—original draft preparation, J.Z.; writing—review and editing, J.Z.; visualization, J.Z., R.Z. and W.Z.; supervision, B.W. and C.L.; project administration, B.W. and C.L.; funding acquisition, B.W. and C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of China (Grant Numbers: 52168055 and 52278397), the Natural Science Foundation of Jiangxi Province (Grant Number: 20212ACB204001), “Double Thousand Plan” Innovation Leading Talent Project of Jiangxi Province (Grant Number: jxsq2020101001) and Jiangxi Province Graduate Innovation Special Fund Project (Grant Number: YC2022-B179). Their support is gratefully acknowledged.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Chen, W.; Zhang, G.-H.; Wang, H.; Chen, L.-B. Risk assessment of mountain tunnel collapse based on rough set and conditional information entropy. *Rock Soil Mech.* **2019**, *40*, 3549–3558. [[CrossRef](#)]
2. Zhang, N.; Di, Y.-T.; Li, S.-R. Application of AHP-LEC Method to Risk Assessment of Subway Tunnel Construction. *J. Mil. Transp. Univ.* **2018**, *20*, 85–89. [[CrossRef](#)]
3. Zhai, Q.; Gu, W.-H. Risk Assessment of Tunnel Collapse by EW-AHP and Unascertained Measure Theory. *Saf. Environ. Eng.* **2020**, *27*, 92–97. [[CrossRef](#)]
4. Saaty, T.-L. Correction to: Some mathematical concepts of the analytic hierarchy process. *Behaviormetrika* **2021**, *48*, 193–194. [[CrossRef](#)]
5. Zhang, C.-X.; Wu, S.-C.; Wu, J. Study on risk assessment model of collapse during construction of mountain tunnel and its application. *J. Saf. Sci. Technol.* **2019**, *15*, 128–134. [[CrossRef](#)]
6. Xie, Z.-L.; He, X.-B. Analysis of Risks in Long-distance Floating Transportation of Immersed Tunnel Segment in Complex Inland River and Their Countermeasures. *Tunn. Constr.* **2016**, *36*, 1095–1100. [[CrossRef](#)]
7. Wu, Z.-H. Study on Identification and Assessment of Safety Risk in Shield Construction of Crossing-River Tunnel. Master’s Thesis, Guangxi University, Guangxi, China, 2019. (In Chinese).
8. Zhang, G.-H.; Chen, W.; Jiao, Y.-Y.; Hao, W.; Cheng, T.-W. A failure probability evaluation method for collapse of drill-and-blast tunnels based on multistate fuzzy Bayesian network. *Eng. Geol.* **2020**, *276*, 105752. [[CrossRef](#)]

9. Mortazavi, A. Bayesian Interactive Search Algorithm: A New Probabilistic Swarm Intelligence Tested on Mathematical and Structural Optimization Problems. *Adv. Eng. Softw.* **2021**, *155*, 102994. [[CrossRef](#)]
10. Mortazavi, A. Interactive fuzzy Bayesian search algorithm: A new reinforced swarm intelligence tested on engineering and mathematical optimization problems. *Expert Syst. Appl.* **2022**, *187*, 115954. [[CrossRef](#)]
11. Kamal, A.; Mortazavi, A.; Cakici, Z. Optimal Design of RC Bracket and Footing Systems of Precast Industrial Buildings Using Fuzzy Differential Evolution Incorporated Virtual Mutant. *Arab. J. Sci. Eng.* **2023**, *48*, 1–17. [[CrossRef](#)]
12. Khakzad, N.; Khan, F.; Amyotte, P. Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. *Reliab. Eng. Syst. Saf.* **2011**, *96*, 925–932. [[CrossRef](#)]
13. Chen, W.; Wang, W.; Zhang, G.-H.; Wang, C.-T.; Zhong, G.-Q. Evaluation of Tunnel Collapse Susceptibility Based on T-S Fuzzy Fault Tree and Bayesian Network. *J. Shang Hai Jiao Tong Univ.* **2020**, *54*, 820–830. [[CrossRef](#)]
14. Guo, D.; Li, J.-H.; Li, X.; Li, Z.-F.; Li, P.-X.; Chen, Z.-Y. Advance prediction of collapse for TBM tunneling using deep learning method. *Eng. Geol.* **2022**, *299*, 106556. [[CrossRef](#)]
15. Wei, X.-Y.; Jin, C.-L.; Gong, L.; Zhang, X.; Ma, M.-H. Risk evaluation of railway tunnel water inrush based on PCA-improved RBF neural network model. *J. Railw. Sci. Eng.* **2021**, *18*, 794–802. [[CrossRef](#)]
16. Zhang, P.-F.; Li, T.-R.; Wang, G.-Q. Multi-source information fusion based on rough set theory: A review. *Inf. Fusion* **2021**, *68*, 85–117. [[CrossRef](#)]
17. Wu, B.; Qiu, W.; Huang, W.; Meng, G.; Nong, Y.; Huang, J. A Multi-Source Information Fusion Evaluation Method for the Tunneling Collapse Disaster Based on the Artificial Intelligence Deformation Prediction. *Arab. J. Sci. Eng.* **2022**, *47*, 5053–5071. [[CrossRef](#)]
18. Zuo, Y.L.; Zhu, H.H.; Li, X.J. Estimation of rock mass discontinuity spacing distributions using maximum entropy principle. *Chin. J. Rock Mech. Eng.* **2017**, *36* (Suppl. S1), 3492–3498. [[CrossRef](#)]
19. Pan, Y.; Zhang, L.-M.; Wu, X.-G.; Skibniewski, J.-M. Multi-classifier information fusion in risk analysis. *Inf. Fusion* **2020**, *60*, 121–136. [[CrossRef](#)]
20. Zhang, L.-M.; Wu, X.-G.; Zhu, H.-P.; AbouRizk, M.-S. Perceiving safety risk of buildings adjacent to tunneling excavation: An information fusion approach. *Autom. Constr.* **2016**, *73*, 88–101. [[CrossRef](#)]
21. Li, S.-C.; Liu, C.; Zhou, Z.-Q.; Li, L.-P.; Shi, S.-S.; Yuan, Y.-C. Multi-sources information fusion analysis of water inrush disaster in tunnels based on improved theory of evidence. *Tunn. Undergr. Space Technol.* **2021**, *113*, 103948. [[CrossRef](#)]
22. Wu, B.; Qiu, W.; Huang, W. A multi-source information fusion approach in tunnel collapse risk analysis based on improved Dempster-Shafer evidence theory. *Sci. Rep.* **2022**, *12*, 3626. [[CrossRef](#)] [[PubMed](#)]
23. Guo, Q.; Wen, W.-L.; Wang, Y.-N.; Qi, L.-G. Basic Probability Assignment Generation Method and Application Based on Cloud Model. *J. Electron. Inf. Technol.* **2022**, *44*, 1–8. [[CrossRef](#)]
24. Zhang, L.-M.; Wu, X.-G.; Ding, L.-Y. A novel model for risk assessment of adjacent buildings in tunneling environments. *Build. Environ.* **2013**, *65*, 185–194. [[CrossRef](#)]
25. Meng, W.-L. Prediction of Ground Deformation and Assessment of Constructions Risks of Excavation Face during Shield Tunneling in Weathered Stratum. Master's Thesis, Shanghai Jiao Tong University, Shanghai, China, 2019. (In Chinese). [[CrossRef](#)]
26. Liu, Y.; Lian, J.; Bartolacci, M.-R.; Zeng, Q.-A. Density-Based Penalty Parameter Optimization on C-SVM. *Sci. World J.* **2014**, *2014*, 851814. [[CrossRef](#)]
27. Wang, Y.; Zhang, L.-M. Feature-based evidential reasoning for probabilistic risk analysis and prediction. *Eng. Appl. Artif. Intell.* **2021**, *102*, 104237. [[CrossRef](#)]
28. Xu, X.-B.; Zheng, J.; Yang, J.-B.; Xu, D.-L.; Chen, Y.-W. Data classification using evidence reasoning rule. *Knowl.-Based Syst.* **2017**, *116*, 144–151. [[CrossRef](#)]
29. Yang, Y.; Han, D.-Q. A new distance-based total uncertainty measure in the theory of belief functions. *Knowl.-Based Syst.* **2016**, *94*, 114–123. [[CrossRef](#)]
30. Gao, X.-X.; Chen, M.-Y.; Wang, T.-Y. Design and optimization for the separation of a ternary methyl methacrylate-methanol-water mixture to save energy. *Energy Sources Part A Recovery Util. Environ. Eff.* **2020**, *2020*, 1–10. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.