Fuzzy risk assessment of a deeply buried tunnel under incomplete information

Yuanpu Xia1,†, Ziming Xiong1,2, Hao Lu1, Zhu Wen2 and Chao Ma1

1State Key Laboratory of Disaster Prevention and Mitigation of Explosion and Impact, The Army Engineering University of PLA, Nanjing 210007, People’s Republic of China
2School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, People’s Republic of China

Risk assessment has always been an important part of safety risk research in tunnel and underground engineering. Owing to the characteristics of tunnel construction, to achieve an expected risk control effect, it is necessary to carry out accurate risk assessment research according to the risk assessment concept based on the entire tunnel construction process. At present, because of the frequent occurrences of safety accidents, a variety of risk assessment models have been proposed for different tunnel projects such as subways and railway tunnels, which can be roughly classified into two types: probability-based and fuzzy set theories. However, the existing models may be more suitable for the construction stage, and the design stage lacks a reliable and practical fuzzy risk assessment method. Therefore, based on fuzzy set theory and similarity measure theory, a risk assessment model is proposed to adapt to the characteristics that the risk information is difficult to quantify the fuzziness in the design phase. Firstly, new ideas of fuzzy risk analysis are proposed to overcome deficiencies in existing methods; secondly, a new similarity measure is constructed; then fusing multi-source fuzzy information based on evidence theory, the relationship between similarity measure and mass function is established. Finally, the new method is applied to the Yuelongmen tunnel. Results show that the concept of risk control and the risk assessment model are feasible.

1. Introduction

Owing to the characteristics and the potential application value of tunnel projects, tunnels are often required in the construction of infrastructure. For example, to build highways and high-speed railways, many deeply buried, long tunnels in southwestern China have been built or are under construction. A growing
number of Chinese cities are building metro systems to solve their traffic congestion problems [1–3]. Tunnel construction involves the geotechnical engineering of large-scale complex systems, and various types of risk involved throughout the construction period [4,5]. This affects all parties involved as well as those not directly involved in the project [6]. Especially in geologically complex areas (karst developed areas, broken fault zones, areas of high crustal stress, etc.), various accidents such as collapses [7], water in-rushes [8] and rock bursts [9] can easily be triggered, which may cause casualties, economic loss, environmental harm and delay which may eventually lead to cost overruns and contract extensions [10]. The frequent occurrence of disasters in tunnel construction is increasingly arousing the attention of all parties including the public [11]. Therefore, how to reduce the occurrence of accidents during tunnel construction, decrease the number of casualties and mitigate future economic losses has always been a focus for engineers, especially in developing countries such as China.

Risk assessment is an important part of risk management, which is a tool designed to support all aspects of decision-making from the beginning to the end of a process [12]. It has been increasingly valued and regarded as an important guarantee for the safe running of large engineering projects [13]. Various risk assessment tools have been applied in tunnel engineering, including security vulnerability analysis (SVA) [14], bow-tie diagrams [15], event tree analysis (ETA) [16], fault tree analysis (FTA) [17], probabilistic risk analysis [18,19], the analytical hierarchy process (AHP) [20], grey vulnerability analysis (SVA) [14], bow-tie diagrams [15], event tree analysis (FTA) [17], probabilistic risk analysis [18,19], the analytical hierarchy process (AHP) [20], grey systems [21], Bayesian networks [22] and fuzzy sets [23]. In addition, artificially intelligent methods such as geographic information systems (GIS) [24], risk analysis software [25,26] and Monte Carlo simulations [27] have also been widely used in recent years. Probabilistic risk assessment (PRA) is the method most widely used in geotechnical engineering [13]. For example, Jurado et al. [28] developed a general PRA framework to estimate the probability of occurrence of groundwater-related hazards in tunnel construction based on FTA. Sousa & Einstein [29] introduced a dynamic Bayesian network (DBN) model for the risk of construction failure. Sejnoha et al. [30] estimated the probability of the occurrence of cave-in collapses and their consequences using FTA and ETA tools. Špačková et al. [31,32] further put forward a probabilistic model by considering the inhomogeneity of the geological conditions, the uncertainty of the consequences of damage, and human and other external factors.

With the increasing application of risk assessment models in tunnel engineering, some scholars gradually realized that insufficient information is a major problem facing risk analysts working in geotechnical engineering contexts, and the usual methods cannot support precise probabilities [33]: for example, Marques et al. [34] proposed that the conventional probabilistic approach to uncertainty can be extended to include imprecise information in the form of intervals. Beer et al. [35] have realized the role of epistemic uncertainty in civil engineering reliability evaluation and discuss the relationship between subjective probability and imprecise probability that represents epistemic uncertainty. To deal with precise information (insufficient information, vague information, descriptive language information, incomplete data, etc.), some fuzzy risk assessment models based on fuzzy set theory (first proposed by Zadeh [36]) have been proposed. The fuzzy risk assessment models can be roughly divided into the following categories:

(1) Fuzzy risk assessment function analysis model. The mechanism underpinning the development and occurrence of a geological hazard is analysed using a physico-mechanical model, then the risk assessment function is constructed, and the value of the function parameters is fuzzed. For example, firstly, the mechanical model of a breakwater failure problem is established, then extended fuzzy set theory (possibility theory and evidence theory) is used to evaluate the risk to the breakwater [37]. Hao et al. [38] first analysed the mechanism of water in-rush in karst tunnels, and then used margin theory to determine the risk of water in-rush under epistemic uncertainty.

(2) Mixed risk assessment model. The existing risk assessment methods such as Bayesian network, or FTA, are fuzzed using fuzzy set theory, and a fuzzy risk assessment model is obtained. For example, Zhang et al. [39] proposed a fuzzy probability risk decision approach for metro construction based on fuzzy set theory and FTA. Nadjafi et al. [40] proposed a dynamic FTA based on a fuzzy lower and upper (L-U) bounded failure distribution. To deal with the uncertainty during the assessment of water in-rush risk in underground engineering works, Yang et al. [41] constructed a model based on GIS and the fuzzy set theory. In addition, a Bayesian network is seen as a common tool used to deal with risk analysis, and different fuzzy Bayesian network assessment and decision models are often suggested as being applicable [42–45].

(3) Risk assessment model based on attribute measure. Wang et al. [46] and Chu et al. [47] constructed a risk assessment model of water in-rush in a karst tunnel and a coal mine based on the secondary fuzzy comprehensive evaluation, respectively, where every index was constructed by membership
degree. Li et al. [48] proposed a risk assessment model of water in-rush in a karst tunnel based on attribute measure theory.

(4) Fuzzy risk assessment model based on a similarity measure. This was first introduced by Wang in 1983 [49] and has been applied in many fields. Similar measures indicate the degree of similarity between two fuzzy numbers. Yong et al. [50] proposed a similarity measure-based method to estimate human-health risk from groundwater contamination and evaluated possible regulatory actions. Parandin et al. [51] proposed a fuzzy risk analysis algorithm to cope with fuzzy analysis problems based on a new similarity measure.

In addition, the situation where random and fuzzy variables exist in the same risk assessment model may occur causing hybrid uncertain information. For example, Xie et al. [52] developed a new hybrid reliability analysis method so that the probability analysis (PA) and interval analysis (IA) loops are decomposed into two separate loops. The existence of hybrid uncertain information increases the complexity of the risk assessment model, and so Peng et al. [53] proposed a hybrid first order reliability analysis method to improve reliability analysis of the results.

From this summary, we can see that risk assessment has aroused widespread interest among geotechnical engineers. Different risk assessment models are regularly proposed, and fuzzy risk analysis is attracting more attention, but we find that the current research into risk in tunnel engineering still has the following deficiencies.

1.1. The study of risk assessment lacks communication with decision-makers

Risk management is essentially a series of risk decision-making processes, and risk assessment should also consider how to be more conducive to decision-making [12]. Risk analysts and decision-makers generally belong to different teams, which complicates the work related to risk control [54]. For complex systems engineering such as the construction of mountain tunnels, the security issues are generally more complex, in addition to considering the severity of the risk, the decision process often needs to consider the impact of other factors. For example, the risk attitude, work experience and educational background of decision-makers may affect the level of decision-making. In addition, decision-makers always want to achieve the best risk control effect under the constraints of limited cost, time and human resources. Therefore, to better facilitate subsequent decision-making, it is necessary for risk analysts to consider the influence of decision factors when conducting risk assessments.

1.2. The feasibility of the risk assessment model is not fully considered

During construction, risk assessment work is implemented dynamically as the construction processes: in terms of time, economy, human resources and so on, will be subject to certain restrictions. No project will invest unlimited resources in risk assessment work. Furthermore, a mountain tunnel will be affected by the construction conditions, natural environmental changes such as earthquakes, rainstorms, etc. The literature [42] emphasizes the role of an expert group in the use of fuzzy Bayesian networks; due to the advantages of using experts, they are helpful when improving the reliability of risk assessment. However, it is unrealistic to get expert help and guidance at every turn: it is not always necessary to invite expert opinion.

1.3. The study of risk assessment lacks continuity

Current models are mainly aimed at the safety and risk during construction, but the construction of a tunnel involves investigation, design and construction stages. Although the risk characteristics of different stages are different, they remain closely linked. Risk assessment during the construction period is not independent of that at other stages and should be carried out on the basis of risk assessment in the design stage. Therefore, we should consider how to carry out risk assessment over the whole construction cycle.

Based on the above analysis, we believe that it is necessary to carry out the risk assessment over the entire process of tunnel construction. It can adapt to the characteristics of tunnel construction and also fully use the impact of information supplement on risk assessment to improve the science and rationality of risk assessment results. Through the risk analysis of the whole construction process, we can get to know the tunnel risks and focus on the key goals. This process can reduce the error of decision-making, improve the utilization of resources and increase the efficiency of decision-making.
We constructed a risk analysis model aiming at the whole construction process of tunnel engineering, as shown in figure 1.

From figure 1, the exploration phase is the initial evaluation, and it mainly entails qualitative analysis. It is important to initially grasp the engineering geological and hydrogeological conditions in the area where the tunnel may pass, the route selection and its feasibility, avoidance of geologically complex areas, such as faults, karst, underground rivers, etc. In the design stage, based on the site investigation, the geological conditions are analysed further, the risk of the major geologic hazards and their distribution along the route are judged, and corresponding risk control plans are formulated. Owing to the characteristics of deep tunnels, the geological information in this stage is not sufficient. Fuzzy semi-quantitative analysis may be better, although semi-quantitative evaluation cannot give accurate results. The risk assessment results can be divided into several levels (typically four or five), the entire line of tunnel can be divided into different risk grade sections according to the above risk grades, and then risk control plans are formulated which are feasible in cost and time. There are more risk factors and more sources of uncertainty in risk assessment during construction. If the quality of risk assessment result is not high enough, accidents may happen. However, the construction stage is constrained by various factors, and the resources available are also limited, which means that we have to put limited resources into the most in-need sections, namely high-risk areas. Then, the precise evaluation models described above can be used to evaluate these high-risk sections. Through continuous and dynamic risk assessment research, we can realize the target risk levels at limited economic, time and human resource cost. Risk can be reduced to an acceptable level. Of course, the realization of this goal is inseparable from the reliable and practical risk assessment results of the design stage.

At present, various risk analysis methods based on fuzzy set theory have been proposed. Different methods have their characteristics and limitations, e.g. the fuzzy risk assessment function analysis model must construct a specific function; mixed risk assessment involves the integration of different expert opinions; the risk assessment levels have been fuzzed in attribute measure-based risk analysis models, but the risk index is regarded as a precise value; the study of similarity measure-based fuzzy risk analysis has aroused wide interest in different fields and has evolved continuously [55]. By combining with the characteristics of risk assessment at the design stage, evaluation results should be reliable and easy to understand for decision-makers, and the calculations in the model should not be too complicated. Considering the above factors, we believe that the fuzzy risk assessment model, based on a similarity measure, is suitable for risk analysis at the tunnel engineering design stage.

This paper is organized as follows: an improved research idea of risk assessment has been proposed based on the research status of similar measure-based fuzzy risk analysis in §2. In §3, based on the deficiency of existing calculation methods of similarity measure, an improved formula is proposed. In §4, firstly, the similarity measure is transformed into a form of evidence (i.e. a mass function), and then based on the theory of evidence synthesis, different pieces of evidence are synthesized to obtain the assessment result. In §5, the proposed risk assessment model is applied to Yuelongmen Tunnel. The results show that the proposed method is effective and can be used as an important part of the tunnel safety risk assessment system. The main conclusions are drawn in §6.
2. Risk assessment model

2.1. The traditional fuzzy risk assessment method based on similarity measure

Although there are differences between existing fuzzy risk analysis models, the basic processes therein are similar, and the typical steps are shown below:

**Step 1.** Construction of a set of linguistic terms and its corresponding generalized trapezoidal fuzzy numbers. For example, the risk is divided into nine levels (table 1).

**Step 2.** Description of the likelihood and severity of each evaluation index with linguistic terms.

**Step 3.** Computation of the total risk $R$ of event A with the help of arithmetic operators used on generalized trapezoidal fuzzy numbers.

$$R = \frac{\sum_{i=1}^{n} \bar{w}_i \otimes \bar{R}_i}{\sum_{i=1}^{n} \bar{w}_i},$$

where $\bar{R}$ is described by a generalized trapezoidal fuzzy number.

**Step 4.** Measuring the similarities of total risk $\bar{R}$ with all given linguistic terms in Step 1, and considering the largest similarity as a risk value of the system in linguistic terms.

From the above steps, we can see that the model involves operations between different fuzzy numbers (see appendix A for details) and the acquisition of a similarity measure. However, the results of the multiplication and division with the existing formulae are an approximation of a triangular fuzzy number for simplified computation consideration, which is likely to result in partial information loss of the obtained risk value [57]. In the case of more risk indicators, it may have a greater impact on the results of risk assessment. Additionally, although different formulae are used to obtain the similarity measure, the calculation between nonlinear fuzzy numbers is still not ideal. For the fuzzy risk analysis of a tunnel, the fuzzy numbers of nonlinear parameters may be more reasonable. Therefore, to improve the reliability of tunnel engineering risk assessment, it is necessary to improve the existing risk analysis model.

### Table 1. Fuzzy representations of linguistic terms [56].

| linguistic valued trapezoidal fuzzy number | linguistic valued trapezoidal fuzzy number |
|------------------------------------------|------------------------------------------|
| absolutely low                           | fairly high                              |
| (0, 0, 0, 0, 1.0)                        | (0.58, 0.63, 0.80, 0.86, 1.0)            |
| very low                                 | high                                     |
| (0, 0, 0.02, 0.07, 1.0)                  | (0.72, 0.78, 0.92, 0.97, 1.0)            |
| low                                      | very high                                |
| (0.04, 0.1, 0.18, 0.23, 1.0)             | (0.93, 0.98, 1.0, 1.0, 1.0)              |
| fairly low                               | absolutely high                          |
| (0.17, 0.22, 0.36, 0.42, 1.0)            | (1.0, 1.0, 1.0, 1.0, 1.0)                |
| medium                                   |                                          |
| (0.32, 0.41, 0.58, 0.65, 1.0)            |                                          |
deduce the mass function of evidence theory by a similarity measure. Therefore, we can first calculate the similarity measure between the risk assessment index and different risk levels; then the assessment result can be determined by evidence theory to fuse multi-source information. The specific steps required are as follows.

Step 1. Construction of a set of linguistic terms and its corresponding generalized trapezoidal fuzzy numbers. For risk analysis in tunnel engineering, this is generally divided into four risk levels, which is different from the nine levels described above. Referring to the classification of the existing tunnel water in-rush risk assessment level [59], the risk levels are as listed in table 2.

| risk level | linguistic valued trapezoidal fuzzy number |
|------------|------------------------------------------|
| A          | (0, 0, 15, 21; 1; 1, 0.5)                |
| B          | (15, 21, 37, 45; 1; 0.5, 2)              |
| C          | (37, 45, 58, 71; 1; 0.5, 1)              |
| D          | (58, 71, 100, 100; 1; 1, 1)              |

Table 2 shows a percentage system, which is mainly used for convenience. The risk level can be represented by membership functions. The equation describing level A is written as follows:

\[
    u_A(x) = \begin{cases} 
        1, & 0 \leq x \leq 15 \\
        \left( \frac{21 - x}{6} \right)^m, & 15 \leq x \leq 21 \\
        0, & \text{others} 
    \end{cases}
\]

Step 2. Determination of the fuzzy number of each index.

For the fuzzy risk analysis based on generalized trapezoidal fuzzy numbers, the fuzzy numbers are mainly chosen from table 1 according to different linguistic terms; however, here we do not require fixed fuzzy numbers; we construct different fuzzy numbers according to different research objects and risk information. The risk assessment of water in-rush in tunnel engineering is taken as an example to discuss how to deal with fuzzy information about risk factors. The specific construction process for incomplete data is presented as follows:

(1) Assuming that \( X = \{x_1, x_2, \ldots, x_n\} \) is a set of data describing the uncertainty of a risk factor, the average value is:

\[
    m = \frac{1}{n} \sum_{i=1}^{n} x_i. 
\]

(2) The data are re-divided by the mean, and two datasets are obtained:

\[
    X_1 = \{x_i | x_i < m\}, \quad X_2 = \{x_i | x_i > m\}. 
\]

The average values of the new two datasets are then presented as:

\[
    m_1 = \frac{1}{N(X_1)} \sum_{i=1}^{n} x_i, \quad m_2 = \frac{1}{N(X_2)} \sum_{i=1}^{n} x_i. 
\]

(3) The membership function of the fuzzy number is:

\[
    u(x_i) = \begin{cases} 
        \frac{x_i - (2m_1 - m + 1)}{2m_1 - m - 1}, & 2m_1 - m + 1 \leq x_i < m - 1 \\
        1, & m - 1 \leq x_i < m + 1 \\
        \left( \frac{2m_2 - m - 1 - x_i}{2m_2 - m - 1} \right)^2, & m + 1 \leq x_i < 2m_2 - m - 1 \\
        0, & \text{others} 
    \end{cases}
\]

The values of \( m \) and \( n \) can be determined according to the actual situation of the risk indicator information. Thus, the constructed possibility distribution can objectively reflect uncertain information.
Step 3. The determination of similarity measure between risk indicator information and different risk levels. The specific formula used in this calculation is discussed in §3.
Step 4. Deriving the initial mass function from the similarity measure.
Step 5. Adjustment of initial mass function.
Step 6. Multi-source uncertainty information fusion. The evidence fusion formula is:

\[ m(A) = \begin{cases} 
0, & A = \emptyset \\
\frac{1}{1 - k} \sum_{A \cap B = A} m_1(A_i)m_2(B_j), & A \neq \emptyset, 
\end{cases} \tag{2.7} \]

where \( A, A_i, B_j \in 2^d \) and \( k = \sum_{A \cap B = \emptyset} m_1(A_i)m_2(B_j) \): Steps 4 and 5 will be elaborated in §4.

3. Calculation of similarity measure

3.1. Existing similarity measures and their limitations

The similarity measure between fuzzy numbers is an important research topic in the fuzzy risk analysis, and it has aroused the interest of many scholars. Different similarity measures have been proposed by Chen & Chen [60,61], Wei & Chen [62], Xu et al. [63] and Hejazi et al. [64] (see appendix B for details). Then, in 2015, Patra & Mondal [56] proposed a new similarity measure associating the geometric distance, area and height of generalized trapezoidal fuzzy numbers (equation (3.1)), which overcomes the shortcomings of previous similarity measures:

\[ S(\tilde{A}, \tilde{B}) = \left( 1 - \frac{\sum_{i=1}^{4} |a_i - b_i|}{4} \right) \times \left( 1 - \frac{1}{2} |A(\tilde{A}) - A(\tilde{B})| + |w_\tilde{A} - w_\tilde{B}| \right). \tag{3.1} \]

Although the above equation is improved on previous research, this method still has two significant defects; therefore, Khorshidi & Nikfalazar [65] proposed the following improved calculation method:

\[ S(\tilde{A}, \tilde{B}) = \left( 1 - \frac{\sum_{i=1}^{4} |a_i - b_i|}{4} \times d'(\tilde{A}, \tilde{B}) \right) \times \left( 1 - \frac{|A(\tilde{A}) - A(\tilde{B})| + |w_\tilde{A} - w_\tilde{B}| + (|P(\tilde{A}) - P(\tilde{B})|)/\max(P(\tilde{A}), P(\tilde{B}))}{3} \right). \tag{3.2} \]

If \( \max(P(\tilde{A}), P(\tilde{B})) = 0 \), then we have \( (P(\tilde{A}) - P(\tilde{B}))/\max(P(\tilde{A}), P(\tilde{B})) = 0 \).

In addition, Li & Zeng [66] proposed a new similarity measure between generalized trapezoidal fuzzy numbers, which could reduce the complexity of computation and improve their ability to distinguish between different values, thus broadening its scope of application:

\[ S(\tilde{A}, \tilde{B}) = se \times sp, \tag{3.3} \]

where

\[ se = \begin{cases} 
e^{-|a_1 - b_1|}, & a_1 = a_4, b_1 = b_4 \\
e^{-k+h+z+i}, & \text{otherwise} \end{cases} \quad \text{and} \quad sp = \frac{e + \min(P(\tilde{A}), P(\tilde{B}))}{e + \max(P(\tilde{A}), P(\tilde{B}))}, \]

\( k \) is the span difference, \( h \) is the centre width difference, \( s \) represents the minimum value of the difference between the left- and right-hand points of the peak value and \( z \) is the central difference for the fuzzy numbers \( \tilde{A} \) and \( \tilde{B} \), respectively. The specific formulae are as follows:

\[ k = |(a_4 - a_1) - (b_4 - b_1)|, \quad h = |(a_3 - a_2) - (b_3 - b_2)|, \]

\[ s = \min(|a_2 - b_2|, |a_3 - b_3|), \quad z = \frac{|a_4 + a_3| - |b_4 + b_1|}{2}. \]

Through the use of the above analysis, the related calculation method has been improved; however, the above methods are mainly used for generalized trapezoidal fuzzy numbers and cannot be used directly for nonlinear fuzzy numbers. For example, assuming the fuzzy numbers \( \tilde{A} \) and \( \tilde{B} \) as shown in figure 2.
Owing to the nonlinear effect, although \(a_1 = b_1, a_2 = b_2, a_3 = b_3\) and \(a_4 = b_4\), fuzzy numbers \(\bar{A}\) and \(\bar{B}\) are different. In addition, the areas \(A(\bar{A}) = A(\bar{B})\), perimeters \(P(\bar{A}) = L_1 + L_5 + L_3 + L_6, P(\bar{B}) = L_2 + L_3 + L_4 + L_6\) and heights \(w(\bar{A}) = w(\bar{B})\). If the latest research results are taken to calculate the similar measures, namely equation (3.2) or (3.3), we find that the similarity measure is 1, which is inconsistent with reality. Therefore, it is necessary to propose another nonlinear similarity measure.

### 3.2. Proposed similarity measure for nonlinear fuzzy numbers

The calculation of the similarity measure generally considers the geometric distance, distance from the centre of gravity, perimeter, area, height and so on. Therefore, based on previous studies, especially on the latest research results, we propose the following similarity measure calculation method: suppose that there are two fuzzy numbers \(\bar{A}\) and \(\bar{B}\), and the corresponding membership functions are \(u(\bar{A})\) and \(u(\bar{B})\), respectively, then:

\[
S(\bar{A}, \bar{B}) = \left(1 - \sum_{i=1}^{4} \frac{|a_i - b_i|}{4}\right) \times e^{-d'(\bar{A}, \bar{B})} \times \frac{\min(P(\bar{A}), P(\bar{B}))) + \min(A(\bar{A}), A(\bar{B})) + \min(w_{\bar{A}}, w_{\bar{B}))}{\max(P(\bar{A}), P(\bar{B}))) + \max(A(\bar{A}), A(\bar{B})) + \max(w_{\bar{A}}, w_{\bar{B}))}. \tag{3.4}
\]

where \(d'(\bar{A}, \bar{B})\) denotes the distance from the centre of gravity, \(d'(\bar{A}, \bar{B}) = \sqrt{(x_{\bar{A}} - x_{\bar{B}})^2 + (y_{\bar{A}} - y_{\bar{B}})^2}, x_{\bar{A}}\) is given by

\[
x_{\bar{A}} = \frac{\sum_{i=1}^{4} a_i w_{\bar{A}}}{\int_{a_i}^{a_{i+1}} u_{\bar{A}}(x) dx} \tag{3.5}
\]

Similarly, we can also calculate other coordinates \(y_{\bar{A}}, x_{\bar{B}}, y_{\bar{B}}\) of the centre of gravity. The perimeters and areas can be calculated thus:

\[
P(\bar{A}) = (a_4 - a_1) + \int_{a_1}^{a_4} (1 + u_{\bar{A}}(x))^{1/2} dx \tag{3.6}
\]

\[
A(\bar{A}) = \int_{a_1}^{a_4} u_{\bar{A}}(x) dx \tag{3.7}
\]

A larger \(S(\bar{A}, \bar{B})\) corresponds to a closer similarity between fuzzy numbers \(\bar{A}\) and \(\bar{B}\). It will be proved that the proposed similarity measure could satisfy some of the properties.

From equation (3.4), \(S(\bar{A}, \bar{B}) \in [0, 1]\) and \(S(\bar{A}, \bar{B}) = S(\bar{B}, \bar{A})\) are satisfied. Another property of the above formula is discussed. \(\bar{A} = \bar{B}\) if and only if \(S(\bar{A}, \bar{B}) = 1\).

**Proof.** When \(\bar{A} = \bar{B}\), it implies that \(d'(\bar{A}, \bar{B}) = 0, P(\bar{A}) = P(\bar{B}), A(\bar{A}) = A(\bar{B}), w(\bar{A}) = w(\bar{B})\), then we have \(S(\bar{A}, \bar{B}) = 1\).

If \(S(\bar{A}, \bar{B}) = 1\), it implies that \(\sum_{i=1}^{4} |a_i - b_i| = 0\), which means that \(a_1 = b_1, a_2 = b_2, a_3 = b_3, a_4 = b_4\). Also, \(d'(\bar{A}, \bar{B}) = 0\), which means that \(x_{\bar{A}} = x_{\bar{B}}\) and \(y_{\bar{A}} = y_{\bar{B}}\), \(\min(P(\bar{A}), P(\bar{B})) + \min(A(\bar{A}), A(\bar{B}))) + \min(w_{\bar{A}}, w_{\bar{B}})) = 1\), which suggests that \(P(\bar{A}) = P(\bar{B}), A(\bar{A}) = A(\bar{B})\) and \(w(\bar{A}) = w(\bar{B})\). Therefore, the fuzzy number \(\bar{A}\) would be equal to the fuzzy number \(\bar{B}\).
4. Information fusion and risk evaluation

4.1. Construction of initial mass function

The mass function obtained by the similarity measure is the key step in multi-source information fusion. The main steps for deriving the mass function by similarity measure are shown below.

1. Determination of mean $E_i$ and the variance $D_i$ for each type of uncertainty information.

$E_i = M(A) = \int_0^1 [\alpha_1(\lambda) + \alpha_2(\lambda)] \lambda \, d\lambda,$ 

where $[\alpha_1(\lambda), \alpha_2(\lambda)]$ is the cut set of $\lambda$ for the distribution of the fuzzy number.

$D_i = \delta(A) = \int_0^1 \frac{1}{6} [\alpha_2(\lambda) - \alpha_1(\lambda)]^2 \lambda \, d\lambda.$ 

Then the induced function $Q_i$ is obtained

$Q_i = \frac{E_i}{D_i}.$ 

2. The ordered function $\text{Order}_i$ can be obtained by use of

$\text{Order}_i = \frac{Q_i}{\sum_{i=1}^N Q_i}.$ 

3. The degree of support for uncertainty information at each risk level can be determined based on

$\text{Sup}(j) = \text{Order}_i \times S_{ij}, j = 1, 2, \ldots, M,$ 

where $S_{ij}$ represents the similarity measure between the $i$th risk factor information and the $j$th risk level.

4. The credibility of uncertainty information to risk level can be obtained from the use of

$\text{Crd}(j) = \frac{\text{Sup}(j)}{\sum_{j=1}^M \text{Sup}(j)},$ 

where $\sum_{j=1}^M \text{Crd}(j) = 1,$ and $\text{Crd}(j)$ can be seen as the basic trust assignment function $m(j)$ (mass function) of uncertainty information assigned to a given risk level.

4.2. Adjustment of initial mass function

Different types of uncertain information can affect assessment results from different perspectives, but the degree of impact is different. In other words, the importance is different. The above-mentioned differences cannot be reflected in the initial mass functions. Therefore, it is necessary to modify the initial mass function. This section focuses on determining the weight of each indicator and how to use weights to influence mass functions.

4.2.1. Determination of weights

To improve the reliability of each assigned weight, the influences of subjective and objective weights are considered simultaneously.

4.2.1.1. Subjective weight

AHP is the main method for subjective weighting [67]; however, this method does not consider the correlation effect among different indices. For example, surface water catchment conditions can affect the water pressure at a tunnel face, thus, the trapezoidal fuzzy number is introduced to characterize the expert information. Suppose that the influence matrix of the risk factors is as shown in table 3.

$m_{ij}(i, j \in \{1, 2, 3, 4\})$ represents the degree of influence of index $u_i$ to index $u_j$ (the scale method and its definition of $m_{ij}$ are summarized in appendix C): generally, $m_{ij} \neq m_{ji}.$ Suppose that the trapezoidal fuzzy...
numbers given by the experts are \(m_{ij} = (l_{ij}, e_{ij}, f_{ij}, s_{ij})\), where \(l_{ij} \leq e_{ij} \leq f_{ij} \leq s_{ij}\). The expected value of the trapezoidal fuzzy number is:

\[
E(a_{ij}) = bE^L(m_{ij}) + (1 - b)E^U(m_{ij}).
\]  
(4.7)

\(E^L(m_{ij}) = (l_{ij} + e_{ij})/2\) and \(E^U(m_{ij}) = (f_{ij} + s_{ij})/2\) are the left and right expected values of the trapezoidal fuzzy numbers, respectively; \(b\) represents the attitude of the experts on the results (here, \(b = 0.5\)). The fuzzy influence value of risk index \(u_i\) can be obtained using

\[
u_i = \frac{E(u_i)}{\sum_{i=1}^{n} E(u_i)}.
\]  
(4.9)

### 4.2.1.2. Objective weight

The calculation process of the objective weight is presented below [68].

(1) The correlation coefficient matrix can be established as follows:

\[
R = \begin{bmatrix}
    r_{11} & r_{12} & \cdots & r_{1n} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{n1} & r_{n2} & \cdots & r_{nn}
\end{bmatrix},
\]

where \(r_{ij}(i, j = 1, 2 \ldots n)\) is the correlation coefficient of \(u_i; u_j\) and \(r_{ij}\) can be represented by the following formula:

\[
|r_{ij}| = \frac{\text{cov}(u_i(x), u_j(x))}{\sqrt{D(u_i(x)) \sqrt{D(u_j(x))}}}
\]  
(4.10)

(2) The mean value of the correlation degree of \(u_i\) is as follows:

\[
\bar{m}_i = \frac{\sum_{j=1; j \neq i}^{n} |r_{ij}|}{(n - 1)}.
\]  
(4.11)

(3) The greater the correlation of an index, the smaller its importance. Then the objective weight can be determined as follows:

\[
w^{*o}_{ki} = \frac{\psi_k}{\sum_{k=1}^{n} \psi_k}, \quad \psi_k = \frac{1}{m_k}.
\]  
(4.12)

Then the comprehensive weight is

\[
w_i = \alpha w^{*o}_i + (1 - \alpha)w^{*s}_i
\]  
(4.13)

where \(\alpha\) is the scale factor of the subjective and objective weights (here, \(\alpha = 0.5\)).

---

### Table 3. Influence matrix of the risk assessment indices.

| \(u_1\) | \(u_2\) | \(u_3\) | \(u_4\) | \(u_5\) |
|---|---|---|---|---|
| \(m_{11}\) | \(m_{12}\) | \(m_{13}\) | \(m_{14}\) | \(m_{15}\) |
| \(\vdots\) | \(\vdots\) | \(\vdots\) | \(\vdots\) | \(\vdots\) |
| \(m_{51}\) | \(m_{52}\) | \(m_{53}\) | \(m_{54}\) | \(m_{55}\) |

---
4.2.2. The improved mass function

When the weights are determined, the factor with the maximum weight is regarded as the key factor, while the others are regarded as non-key factors. The relative weights $\beta_i$ of each factor relative to the key factor are determined. The initial mass function is modified by using the relative weights. The specific process is detailed below:

(1) Suppose that $\alpha(x_i) = \max \omega_i$, $x_i$ is the key factor and the relative weights $\beta_i$ can be obtained by using

$$\beta_i = \frac{\alpha(m_i)}{\alpha(m_f)}.$$  (4.14)

(2) The modified equation of mass function is:

$$\begin{cases} m'_i(A) = \beta_i m_i(A) \\ m'_i(\emptyset) = \beta_i m_i(\emptyset) + (1 - \beta_i). \end{cases}$$  (4.15)

Finally, the modified mass function is fused according to the synthetic rules of evidence theory.

5. Engineering application and discussion

5.1. Engineering background

In the western mountain area of China, the geological structure is complex and karst is widely distributed. A tunnel in this area is characterized by its large depth, long hole lines, high stress and high water pressure [69], which could easily cause water in-rush, or other geologically driven accidents, during construction.

The Yuelongmen tunnel of the Chengdu–Lanzhou railway (figure 3) is a control project for the railway, which has a length of approximately 20 km. It crosses multiple fault zones and the Peijiang River system, the region is seismically active and the geological environment is complex. Geological disasters such as water bursting, landslides, large deformation, rock bursts and other common geological disasters can occur. The method of simultaneous construction in multiple directions by cross-holes and inclined wells could accelerate the construction of the tunnel.

The no. 3 inclined shaft of the Yuelongmen Tunnel is located on the left-hand side of the line. The starting point is XJ3K0, which is connected to the normal line at D2K97 + 700, and the terminus of the station is at XJ3K2 + 30. The length and maximum depth of the no. 3 inclined shaft are 2030 m and 872 m, respectively. The entrance to the inclined shaft is located on the left bank of Gaochuan River, and the entrance elevation is 1280 m (figure 4).

In addition, inclined shaft no. 3 is found to have the following characteristics:

(1) The construction site is narrow and has inconvenient traffic volumes thereat.
(2) Karst development includes broken rock and a high water content.
(3) The shaft has an inclined well with an anti-slope construction that has unfavourable drainage tendencies.
(4) The construction of the inclined shaft caused the nearby river to pour into the mountain.
(5) There has been a change in the construction site of the no. 3 inclined shaft.
The above characteristics make the uncertain information more complex and the likelihood of geological disasters is higher, therefore, it is necessary to improve the quality of risk assessment and perfect risk control measures in the design stage. Taking the risk of water in-rush as an example, the no. 3 inclined shaft is evaluated and different measures are taken for different sections according to the assessment results.

5.2. Fuzzy risk analysis of tunnel water in-rush

The risk assessment of water in-rush of the no. 3 inclined well by a new fuzzy risk assessment method is presented here, and the water in-rush risk throughout the region can be better understood as a result. This can provide the basis for precise risk assessment during construction. The typical section from chainage XJ3K0 + 330 to XJ3K0 + 560 was selected to explain the proposed risk assessment model. The specific assessment process is described below, and all of the initial data used here are mainly based on the engineering geology and hydrogeological exploration report and a supplementary investigation report. Different interpretation of test results from different experts, and the results of the laboratory test are also used. On this basis, the data are extended to facilitate the construction of fuzzy numbers.

5.2.1. The construction of fuzzy numbers for uncertain information

The karst development is the internal factor for controlling the formation of water-induced disasters. The concentration and distribution of karst water are basically consistent with the trend in karst development. The water-rich nature of the karst and the karst development morphology and scale depend on the lithology and hydrodynamic conditions of karst water occurrence and migration. The following essential factors (soluble rock $x_1$, water pressure $x_2$, groundwater circulation conditions $x_3$ and surface water catchment conditions $x_4$) are selected as risk assessment indices for water in-rush in such karst tunnels based on the aforementioned analysis and with reference to the relevant literature [3,70].

5.2.1.1. Soluble rock

Soluble rock is the material foundation of karst development. Given the degree of difficulty of obtaining this type of information in the tunnel site, we determined the content of soluble rock mainly by considering the content of CaCO$_3$ to evaluate rock solubility. The effect of rock structure on rock solubility was not considered in this work. The soluble rock samples were extracted in the field and
related work was mainly analysed in the laboratory. A total of 30 samples were analysed, and the relevant results are shown in table 4.

The data are normalized, then the average values ($m = 84.12, m_1 = 78.83$ and $m_2 = 88.72$) are found. The function of fuzzy number $x_1$ is:

$$u(x_i) = \begin{cases} \frac{(x_i - 74.62)^2}{8.50}, & 74.62 \leq x_i < 83.12 \\ 1, & 83.12 \leq x_i < 85.12 \\ \frac{(92.31 - x_i)^2}{7.19}, & 85.12 \leq x_i < 92.31 \\ 0, & \text{others} \end{cases}$$ (5.1)

The distribution of the other indicial ($x_2$: water pressure at the tunnel face, $x_3$: groundwater circulation conditions and $x_4$: surface water catchment conditions) information can be obtained in the same way (see appendix D for details).

5.2.2. Determination of the initial mass function

5.2.2.1. Similarity measure

The similarity measure between the risk factors and the risk levels of the water in-rush can be obtained by using equation (3.4). Taking the rock solubility as an example, at first, the centre of gravity, perimeter and area of each risk assessment grade and risk evaluation index can be calculated as follows: Level A: $x(A) = 9.58$, $P(A) = 46.21$ and $A(A) = 19$; Level B: $x(B) = 28.33$, $P(B) = 68.66$ and $A(B) = 22.67$; Level C: $x(C) = 48$, $P(C) = 77.05$ and $A(C) = 24.83$; and Level D: $x(D) = 82.05$, $P(D) = 97.93$ and $A(D) = 35.5$; $x_1$: $x(x_1) = 83.82$, $l(x_1) = 37.34$ and $A(x_1) = 7.23$.

Then, the similarity measures are: $S_{11} = S(x_1, A) = 0.3276$, $S_{12} = S(x_1, B) = 0.2834$, $S_{13} = S(x_1, C) = 0.3096$ and $S_{14} = S(x_1, D) = 0.3331$.

Similarly, the others can also be obtained and the similarity measure matrix is as follows:

$$\begin{bmatrix}
0.3276 & 0.2834 & 0.3096 & 0.3331 \\
0.3202 & 0.5084 & 0.6671 & 0.7407 \\
0.3348 & 0.5339 & 0.6269 & 0.6006 \\
0.3178 & 0.2748 & 0.3003 & 0.3230
\end{bmatrix}$$ (5.2)

5.2.2.2. Mean, variance and induced function

For $x_1$: $E(x_1) = 83.99$, $D(x_1) = 2.75$, $Q(x_1) = 37.33$;
for $x_2$: $E(x_2) = 71.01$, $D(x_2) = 11.02$, $Q(x_2) = 6.44$;
for $x_3$: $E(x_3) = 76.8$, $D(x_3) = 8.19$, $Q(x_3) = 9.38$;
for $x_4$: $E(x_4) = 73.9$, $D(x_4) = 2.25$, $Q(x_4) = 32.84$.

5.2.2.3. Initial mass function

Firstly, the ordered functions are as follows:

Order($x_1$) = 0.434, Order($x_2$) = 0.075, Order($x_3$) = 0.109, Order($x_4$) = 0.382.
The initial mass functions $m_{ij}$ are presented as follows:

$$
\begin{align*}
 m_{11} &= m(x_1, A) = 0.2270, & m_{12} &= m(x_1, B) = 0.2532, \\
 m_{13} &= m(x_1, C) = 0.2829, & m_{14} &= m(x_1, D) = 0.2369, \\
 m_{21} &= m(x_2, A) = 0.2496, & m_{22} &= m(x_2, B) = 0.2505, \\
 m_{23} &= m(x_2, C) = 0.2503, & m_{24} &= m(x_2, D) = 0.2496, \\
 m_{31} &= m(x_3, A) = 0.2415, & m_{32} &= m(x_3, B) = 0.2530, \\
 m_{33} &= m(x_3, C) = 0.2542, & m_{34} &= m(x_3, D) = 0.2513, \\
 m_{41} &= m(x_4, A) = 0.2467, & m_{42} &= m(x_4, B) = 0.2503, \\
 m_{43} &= m(x_4, C) = 0.2571, & m_{44} &= m(x_4, D) = 0.2458.
\end{align*}
$$

5.2.3. Multi-source information fusion

Firstly, we determine the subjective and objective weights of each risk index using the aforementioned weight determination method.

5.2.3.1. Subjective weights

According to the subjective weight method described herein, we invited several experts to provide an influence matrix with different indices, and the following matrix is obtained:

$$
U = \begin{bmatrix}
(0.1,0.1,0.1,0.1) & (0.6,0.7,0.8,0.9) & (0.5,0.6,0.7,0.8) & (0.3,0.4,0.5,0.6) \\
(0.1,0.2,0.3,0.35) & (0.1,0.1,0.1,0.1) & (0.15,0.2,0.25,0.3) & (0.12,0.3,0.37,0.45) \\
(0.2,0.28,0.35,0.4) & (0.65,0.72,0.8,0.9) & (0.1,0.1,0.1,0.1) & (0.05,0.1,0.16,0.22) \\
(0.18,0.24,0.3,0.37) & (0.57,0.65,0.72,0.84) & (0.5,0.56,0.62,0.7) & (0.1,0.1,0.1,0.1)
\end{bmatrix}.
$$

According to equation (4.8), the fuzzy effect value of each index is:

$$
\begin{align*}
 u_1 &= (0.2158,0.2871,0.3925,0.5092), & u_2 &= (0.0676,0.1276,0.1906,0.1726), \\
 u_3 &= (0.1439,0.1914,0.2635,0.375), & u_4 &= (0.1942,0.2472,0.3252,0.4467).
\end{align*}
$$

The fuzzy expected value of each index can be determined by using equation (4.7):

$$
E(u_1) = 0.35115, \ E(u_2) = 0.1396, \ E(u_3) = 0.24345 \text{ and } E(u_4) = 0.3033.
$$

Then the subjective weights are: $w_{u_1}^* = 0.338, w_{u_2}^* = 0.135, w_{u_3}^* = 0.235$ and $w_{u_4}^* = 0.292$.

5.2.3.2. Objective weights

The correlation between different indices must first be determined.

$$
R = \begin{bmatrix}
1 & 0.0498 & 0.9665 & -0.2054 \\
0.0498 & 1 & 0.3169 & -0.4908 \\
0.9665 & 0.3169 & 1 & -0.2911 \\
-0.2054 & -0.4908 & -0.2911 & 1
\end{bmatrix}.
$$

The objective weights can then be determined using equations (4.10) to (4.12):

$$
\begin{align*}
 w_{u_1}^{**} &= 0.2253, & w_{u_2}^{**} &= 0.3210, & w_{u_3}^{**} &= 0.1748 \text{ and } w_{u_4}^{**} &= 0.2854.
\end{align*}
$$

Finally the comprehensive weights are: $w_{u_1} = 0.2817, w_{u_2} = 0.2280, w_{u_3} = 0.2049$ and $w_{u_4} = 0.2854$.

5.2.3.3. Information fusion

The maximum weight of $x_4$ is the key factor, and the relative weights are determined to be $\beta_{u_1} = 0.9870, \beta_{u_2} = 0.7989, \beta_{u_3} = 0.7179$ and $\beta_{u_4} = 1$. The initial mass function can be modified using equation (4.14). The fusion results are summarized in table 5.

The fusion result in table 5 indicates that the risk level of water in-rush in the section of the tunnel running from XJ3K0 + 330 to XJ3K0 + 560 is C (high risk). The water in-rush risk in this area must be heeded during construction. The remaining sections of the no. 3 inclined shaft can be evaluated by the same analytical method mentioned (table 6).
5.3. Discussion and verification of evaluation results

In table 6, there are three sections with the highest risk level (C). The length thereof is 1270 m, accounting for 62.7% of the total length, indicating that the overall water in-rush risk of the no. 3 inclined shaft is relatively high. According to the evaluation result, the decision-maker must make a corresponding plan against water in-rush risk and focus on high-risk sections of the tunnel. Owing to the constant updating of geological information during construction, the water in-rush risk may change, so special attention should be paid to changes in geological conditions during construction. The following will combine the actual situation of water burst risk in construction to illustrate the rationality of the proposed method.

On 3 March 2015, when the tunnel was excavated to XJ3K0 + 300, a small amount of water influx was noted. On 8 March 2015, the amount of water suddenly increased after a blasting operation. The construction team immediately stopped the operation, evacuated all non-essential personnel and used pumping equipment to reduce the tunnel water level. Experts were asked to assess the risk of water in-rush before further construction to improve the specified risk control scheme due to the high-risk evaluation at the design stage. One of the priority measures was to undertake advanced horizontal drilling in the tunnel face, which can not only add to the engineering geology and hydrogeological information, but also reduce the water pressure on the tunnel.

On 10 March 2015, an advanced horizontal drill (total length 38 m) was applied to the face. When drilled to XJ3K0 + 382, there was a large amount of water found in the hole, and the water pressure increased significantly. When drilling to XJ3K0 + 356, the water pressure was too great to continue drilling. After withdrawing the drill rod, a water outburst occurred over about 15 m, and the water

---

### Table 5. Multi-source information fusion: water in-rush risk assessment.

| weight coefficient $\beta$ | $m_1$ (A) | $m_2$ (B) | $m_3$ (C) | $m_4$ (D) | $m$ ($\Theta$) |
|---------------------------|----------|----------|----------|----------|----------------|
| 0.9870                    | 0.2240   | 0.2499   | 0.2792   | 0.2338   | 0.0131         |
| 0.7989                    | 0.1994   | 0.2001   | 0.2000   | 0.1994   | 0.2011         |
| 0.7179                    | 0.1734   | 0.1816   | 0.1825   | 0.1804   | 0.2821         |
| 1                         | 0.2467   | 0.2503   | 0.2571   | 0.2458   | 0               |
| fusion results            | 0.2208   | 0.2521   | 0.2896   | 0.2375   | 0               |
| evaluation result         |          |          |          |          | C               |

---

### Table 6. Evaluation of water in-rush risk in different sections of the no. 3 inclined shaft.

| sections | geological description | risk assessment results of water in-rush |
|----------|------------------------|----------------------------------------|
| XJ3K0 + 000 ~ XJ3K0 + 330 | fracture development and groundwater circulation conditions are good | C                           |
| XJ3K0 + 330 ~ XJ3K0 + 560 | a karst fissure is developed, which is beneficial to the storage of groundwater | C                           |
| XJ3K0 + 560 ~ XJ3K0 + 635 | karst section: groundwater circulation conditions are good | B                           |
| XJ3K0 + 635 ~ XJ3K0 + 815 | the joint development is dense and the groundwater is dominated by fissure water in bedrock | B                           |
| XJ3K0 + 815 ~ XJ3K1 + 090 | the overall geological conditions are relatively stable | A                           |
| XJ3K1 + 090 ~ XJ3K1 + 700 | the karst is extremely developed and the groundwater storage conditions are good | C                           |
| XJ3K1 + 700 ~ XJ3K2 + 030 | karst is slightly developed, and the water influx is relatively small during monitoring | A                           |
was clear. On 14 March 2015, another three holes were drilled, and the flow continued to increase, which resulted in the inundation of 70 m of the cave at a flow rate of about 1000 m$^3$ h$^{-1}$. At about 15.00 on 14 March 2015, there was no further increase in the amount of water in the cave. There were six sewage pumps active on the face, and the water flow was stable at about 800–900 m$^3$ h$^{-1}$ (figures 5 and 6).

At the design stage, the water in-rush risk from XJ3K0 + 330 to XJ3K0 + 560 is C (high risk), which is in line with the actual situation encountered during construction. It not only verifies the reliability of the proposed method, but also explains the rationality of the research thinking behind the safety risk assessment as applied to tunnel engineering, that is to emphasize the importance of risk assessment during construction, and pay attention to the role of the site investigation and design stages in the implementation of a safe risk control system.

6. Conclusion

With the implementation of ‘The Belt and Road’ strategy, more deeply buried, long tunnels will be constructed in complex geological conditions such as karst areas. Accurate risk analysis is an important guarantee for the safe construction of such tunnels. Based on the characteristics of tunnel engineering, it is necessary to implement an entire tunnel construction process risk analysis. Considering that current risk research is primarily focused on construction periods,
there is no reliable risk assessment model for the design phase. To further improve tunnel engineering risk research and improve the risk control effect, this paper aims to construct a risk assessment model which fully reflects the risk characteristics of the design stage. The main conclusions are as follows:

(1) Based on the characteristics of information deficiency in the design stage and its role in the risk analysis system by providing basic information guidance for risk analysis during construction, it is reasonable and necessary to construct a semi-quantitative risk assessment model. (2) The fuzzy risk analysis model based on the similarity measure, which has been a commonly used semi-quantitative evaluation method, has some deficiencies including information loss and inability to adapt to nonlinear characteristics. To improve the deficiencies of the existing models, an improved fuzzy risk assessment model based on similarity measures and evidence theory has been proposed. (3) The proposed model was used to assess water in-rush risk in the design stage of the no. 3 inclined shaft of the Yuelongmen tunnel. The tunnel is divided into different sections according to the water in-rush risk levels, and the high-risk area of water in-rush can be focused on during construction. Effective risk control measures or schemes are amended or proposed based on the accurate in-rush risk levels, and the high-risk area of water in-rush can be focused on during construction. Through semi-quantitative risk analysis in the design stage, the efficiency and effectiveness of risk control during construction can be improved.

(4) The constructed fuzzy risk assessment model is an important part of a risk analysis system of the whole construction process, which could deal with different types of risk information in the design stage of tunnel engineering, such as qualitative language, fuzzy information and interval information. The model is advantageous compared to traditional techniques using statistical data and the approach can be seen as a useful complement to existing risk analysis methods.

Data accessibility. This theoretical paper does not rely on empirical data. All the plots can be generated following the equations and instructions provided in the paper.

Authors’ contributions. Y.X. conceived, designed and performed the study. Z.X. collected and analysed the example in the paper. Z.W. and H.L. wrote the paper together. C.M. revised the paper. All authors have read and approved the final published manuscript.

Competing interests. The authors declare that there is no conflict of interest regarding the publication of this article.

Funding. This work was supported by the National Key Basic Research Program (grant no. 2013CB036005) and the National Natural Science Fund Youth Project (51608529).

Acknowledgements. We are grateful to several anonymous reviewers who provided useful advice to improve the manuscript.

Appendix A

The arithmetic operations between the generalized trapezoidal fuzzy numbers $A$ and $B$ are as follows

$$
\tilde{A} \otimes \tilde{B} = (a_1, b_1, c_1, d_1; w_1) \otimes (a_2, b_2, c_2, d_2; w_2) \\
\approx (a_1 a_2, b_1 b_2, c_1 c_2, d_1 d_2; \min(w_1, w_2)), \quad (A1)
$$

$$
\tilde{A} \oplus \tilde{B} = (a_1, b_1, c_1, d_1; w_1) \oplus (a_2, b_2, c_2, d_2; w_2) \\
= (a_1 + a_2 - a_1 a_2, b_1 + b_2 - b_1 b_2, c_1 + c_2 - c_1 c_2, d_1 + d_2 - d_1 d_2; \min(w_1, w_2)) \quad (A2)
$$

and

$$
\tilde{A} \odot \tilde{B} = (a_1, b_1, c_1, d_1; w_1) \odot (a_2, b_2, c_2, d_2; w_2) \\
\approx \left( \begin{array}{c} a_1 a_2 \\ b_1 b_2 \\ c_1 c_2 \\ d_1 d_2 \\ \end{array} \right), \min(w_1, w_2) \quad (A3)
$$

where

\[
\frac{a}{b} = \begin{cases} \frac{a}{b}, & \text{if } a < b \\ 1, & \text{otherwise} \end{cases}
\]

$$
\tilde{A} \ominus \tilde{B} = (a_1, b_1, c_1, d_1; w_1) \ominus (a_2, b_2, c_2, d_2; w_2) \\
= (a_1 - d_2, b_1 - c_2, c_1 - b_2, d_1 - a_2; \min(w_1, w_2)). \quad (A4)
$$
Appendix B

Based on a previous study [60], Chen & Chen [61] proposed a new method for calculating the similarity measure between fuzzy numbers using the concept of centre of gravity (COG).

\[
S(\bar{A}, \bar{B}) = \left( 1 - \frac{\sum_{i=1}^{4} |a_i - b_i|}{4} \right) \times \left( 1 - |x_{\bar{A}}^{*} - x_{\bar{B}}^{*}| \right)^{f(I_{\bar{A}}, I_{\bar{B}})} \times \frac{\min(y_{\bar{A}}^{*}, y_{\bar{B}}^{*})}{\max(y_{\bar{A}}^{*}, y_{\bar{B}}^{*})},
\]

where

\[
f(I_{\bar{A}}, I_{\bar{B}}) = \begin{cases} 
1 & \text{if } 0 < \frac{I_{\bar{A}} + I_{\bar{B}}}{2} \leq 1, \\
0 & \text{if } \frac{I_{\bar{A}} + I_{\bar{B}}}{2} = 0
\end{cases}, \
I_{\bar{A}} = a_4 - a_1, \quad I_{\bar{B}} = b_4 - b_1.
\]

In 2009, based on the characteristics of geometric distance, perimeter and height of fuzzy numbers, Wei & Chen [62] proposed a new method to calculate the similarity measure.

\[
S(\bar{A}, \bar{B}) = \left( 1 - \frac{\sum_{i=1}^{4} |a_i - b_i|}{4} \right) \times \frac{\min(P(\bar{A}), P(\bar{B})) + \min(w_1, w_2)}{\max(P(\bar{A}), P(\bar{B})) + \max(w_1, w_2)},
\]

where \(P(\bar{A})\) and \(P(\bar{B})\) are the perimeters of the fuzzy numbers \(\bar{A}\) and \(\bar{B}\), and the formulae are written as follows:

\[
P(\bar{A}) = \sqrt{(a_1 - a_2)^2 + w_1^2} + \sqrt{(a_3 - a_4)^2 + w_1^2} + (a_3 - a_2) + (a_4 - a_1)
\]

and

\[
P(\bar{B}) = \sqrt{(b_1 - b_2)^2 + w_1^2} + \sqrt{(b_3 - b_4)^2 + w_1^2} + (b_3 - b_2) + (b_4 - b_1).
\]

In 2010, Xu et al. [63] presented a new similarity measure for fuzzy numbers based on distance and COG.

\[
S(\bar{A}, \bar{B}) = 1 - \frac{1}{2} \times \left| \frac{\sum_{i=1}^{4} |a_i - b_i|}{4} + d'(\bar{A}, \bar{B}) \right|,
\]

where \(d'(\bar{A}, \bar{B})\) is the distance of the COG points of fuzzy numbers, as follows:

\[
d'(\bar{A}, \bar{B}) = \sqrt{(x_{\bar{A}}^{*} - x_{\bar{B}}^{*})^2 + (y_{\bar{A}}^{*} - y_{\bar{B}}^{*})^2}/\sqrt{1.25}, \quad (x_{\bar{A}}^{*}, y_{\bar{A}}^{*}) \text{ are the coordinates of COG, and the expressions are expressed as follows.}
\]

\[
x_{\bar{A}}^{*} = \begin{cases} 
\frac{y_{\bar{A}}^{*} \times (a_3 + a_2) + (a_4 + a_1) \times (w_{\bar{A}}^{*} - y_{\bar{B}}^{*})}{2w_{\bar{A}}^{*}}, & \text{if } w_{\bar{A}}^{*} \neq 0 \\
\frac{a_4 + a_1}{2}, & \text{if } w_{\bar{A}}^{*} = 0
\end{cases}
\]

and

\[
y_{\bar{A}}^{*} = \begin{cases} 
\frac{w_{\bar{A}}^{*} \times ((a_3 - a_2)/a_4 - a_1) + 2}{6}, & \text{if } a_4 \neq a_1 \\
\frac{w_{\bar{A}}^{*}}{2}, & \text{if } a_4 = a
\end{cases}
\]

Hejazi et al. [64] proposed a novel method of fuzzy risk analysis based on a new similarity measure of generalized fuzzy numbers, which considers many of their features, such as area, perimeter, height and geometric distance and could potentially overcome the limitations of the other methods. The formula of similarity measure is as follows.

\[
S(\bar{A}, \bar{B}) = \left( 1 - \frac{\sum_{i=1}^{4} |a_i - b_i|}{4} \right) \times \frac{\min(P(\bar{A}), P(\bar{B}))}{\max(P(\bar{A}), P(\bar{B}))} \times \frac{\min(A(\bar{A}), A(\bar{B}))}{\max(A(\bar{A}), A(\bar{B}))} + \frac{\min(w_{\bar{A}}, w_{\bar{B}})}{\max(w_{\bar{A}}, w_{\bar{B}})},
\]

where \(A(\bar{A})\) and \(A(\bar{B})\) are the areas of \(\bar{A}\) and \(\bar{B}\), respectively, and the formulae are as follows:

\[
A(\bar{A}) = \frac{w_{\bar{A}} \times (a_3 - a_2 + a_4 - a_1)}{2}, \quad A(\bar{B}) = \frac{w_{\bar{B}} \times (b_3 - b_2 + b_4 - b_1)}{2}.
\]
Appendix C

See table 7.

Table 7. Scale method and its definition.

| impact scale | \( m_i \) |
|--------------|----------|
| 0.1          | index \( u_i \) does not influence \( u_j \) |
| 0.3          | index \( u_i \) produces a slight effect on \( u_j \) |
| 0.5          | index \( u_i \) produces a general effect on \( u_j \) |
| 0.7          | index \( u_i \) produces a great effect on \( u_j \) |
| 0.9          | index \( u_i \) produces a significant effect on \( u_j \) |
| 0.2, 0.4     | index \( u_i \) influences \( u_j \) |
| 0.6, 0.8     | the degree is in the middle of the results for 0.5 – 0.9 |

Table 8. Monitoring data of the water pressure (MPa).

|          | 58 | 70.75 | 83.5 | 64 | 91.25 | 55.75 | 84.25 | 75.5 | 68.75 | 56 |
|----------|----|-------|------|----|-------|-------|-------|------|-------|----|
|          | 70.25 | 60.75 | 73.75 | 94 | 95.5 | 66.25 | 55.5 | 68.25 | 66.75 |    |
|          | 59.5 | 65 | 67 | 53.25 | 48.75 | 71.75 | 82 | 86.5 | 67.25 | 46.5 |

Appendix D

The distributions of the other indexes (\( x_2 \): water pressure at the tunnel face, \( x_3 \): groundwater circulation conditions, \( x_4 \): surface water catchment conditions) information are as follows.

(1) Water pressure at the tunnel face (\( x_2 \))

Water pressure is an important factor for water in-rush. Table 8 presents the data obtained on site according to the regular measurement of the drilling water pressure.

The function of fuzzy number of \( x_2 \) is:

\[
\pi(x_i) = \begin{cases} 
\frac{(x_i - 52.98)^2}{14.92}, & 52.98 \leq x_i < 67.9 \\
1, & 67.9 \leq x_i < 69.9 \\
\frac{(93.26 - x_i)^2}{23.36}, & 69.9 \leq x_i < 93.26 \\
0, & \text{others} 
\end{cases} 
\]  

(D 1)

(2) Groundwater circulation conditions (\( x_3 \)).

The erodibility of water flow is one of the basic conditions for karst development. The dissolution occurs first in the karst and water interface. The soluble rock dissolution reaches a saturation state if the dissolved ions cannot move quickly, thereby preventing further dissolution. Table 9 shows the monitoring data of the groundwater circulation condition.

The function of the fuzzy number of \( x_3 \) is:

\[
\pi(x_i) = \begin{cases} 
\frac{(x_i - 57.84)^2}{18.44}, & 57.84 \leq x_i < 76.28 \\
1, & 76.28 \leq x_i < 78.28 \\
\frac{(91.92 - x_i)^2}{13.64}, & 78.28 \leq x_i \leq 91.92 \\
0, & \text{others} 
\end{cases} 
\]  

(D 2)
Surface water catchment conditions ($x_4$)

The karst landform conditions and the formation and development of underground karst spaces promote and interact with each other. The percentage of the negative terrain area ($x_{41}$) and the slope gradient ($x_{42}$) were selected to evaluate the quality of the surface water catchment conditions. Table 10 presents the monitoring data of the surface water catchment condition.

The functions of the fuzzy number of $x_{41}$ and $x_{42}$ are:

$$\pi(x_i) = \begin{cases} (x_i - 62.01)^2, & 62.01 \leq x_i < 69.17 \\ 1, & 69.17 \leq x_i < 71.17 \\ (78.33 - x_i)^2, & 71.17 \leq x_i < 78.33 \\ 0, & \text{others} \end{cases}$$  \hspace{1cm} (D3)

and

$$\pi(x_i) = \begin{cases} (x_i - 71.93)^2, & 71.93 \leq x_i < 78.63 \\ 1, & 78.63 \leq x_i < 80.63 \\ (86.11 - x_i)^2, & 80.63 \leq x_i < 86.11 \\ 0, & \text{others} \end{cases}$$  \hspace{1cm} (D4)

The distribution of the surface water catchment conditions $x_4$ can be synthesized with $x_{41}$ and $x_{42}$ by linear weights.

References

1. Qian QH, Lin P. 2016 Safety risk management of underground engineering in China: progress, challenges and strategies. J. Rock Mech. Geotech. Eng. **8**, 423 – 442. (doi:10.1016/j.jrmge.2016.04.001)
2. Shi C, Zhong M, Nong X, He L, Shi J, Feng G. 2012 Modeling and safety strategy of passenger evacuation in a metro station in China. Saf. Sci. **50**, 1319 – 1332. (doi:10.1016/j.ssci.2010.07.017)
3. Li SC, Zhou ZQ, Ye ZH, Li LP, Zhang QQ, Xu ZH. 2015 Comprehensive geophysical prediction and treatment measures of karst caves in deep buried tunnel. J. Appl. Geophys. **116**, 247 – 257. (doi:10.1016/j.jappgeo.2015.03.019)
4. Kim YG. 2014 Application of risk analysis and assessment in tunnel design. Int. J. JRMR **5**, 11 – 18.
5. Moradi MR, Farangi MAE. 2014 Application of the risk matrix method for geotechnical risk analysis and prediction of the advance rate in rock TBM tunneling. Rock Mech. Rock Eng. **47**, 1951 – 1960. (doi:10.1007/s00603-013-0464-x)
6. Eskesen SD, Tengborg P, Kampmann J, Holst Veichters T. 2004 Guidelines for tunnelling risk management: International Tunnelling Association, Working Group no. 2. Tunn. Underground Space Technol. **19**, 217 – 237. (doi:10.1016/j.tust.2004.01.001)
7. Shin HS, Kwon YC, Jung YS, Bae G-J, Kim Y-G. 2009 Methodology for quantitative hazard assessment for tunnel collapses based on case histories in Korea. Int. J. Rock Mech. Min. Sci. **46**, 1072 – 1087. (doi:10.1016/j.jrmms.2009.02.009)
8. Li SC, Wu J, Xu ZH, Li LP. 2016 Unascertained measure model of water and mud inrush risk evaluation in karst tunnels and its engineering application. KSCE J. Civil Eng. **21**, 1170 – 1182.
53. Peng X, Wu T, Li J, Jiang S, Qiu C, Yi B. 2017 Hybrid reliability analysis with uncertain statistical variables, sparse variables and interval variables. Eng. Optimiz. 50, 1347 – 1363. (doi:10.1080/0305215X.2017.1400025)

54. Aven T, Zio E. 2011 Some considerations on the treatment of uncertainties in risk assessment for practical decision making. Reliab. Eng. Syst. Saf. 96, 64 – 74. (doi:10.1016/j.ress.2010.06.001)

55. Lin J, Chen Q. 2013 Fault diagnosis of rolling bearings based on multifractal detrended fluctuation analysis and Mahalanobis distance criterion. Mech. Syst. Signal Process. 38, 515 – 533. (doi:10.1016/j.ymssp.2012.12.014)

56. Patra K, Mondal SK. 2015 Fuzzy risk analysis using area and height based similarity measure on generalized trapezoidal fuzzy numbers and its application. Appl. Soft Comput. 28, 276 – 284. (doi:10.1016/j.asoc.2014.11.042)

57. Gao S, Zhang Z, Cao C. 2009 Multiplication operation on fuzzy numbers. J. Softw. 4, 331 – 338.

58. Lin G, Liang J, Qian Y. 2015 An information fusion approach by combining multigranulation rough sets and evidence theory. Inf. Sci. 314, 184 – 199. (doi:10.1016/j.ins.2015.03.051)

59. Wang Y, Yin X, Geng F, Jing H, Su H, Liu R. 2017 Risk assessment of water inrush in karst tunnels based on the efficacy coefficient method. Pol. J. Environ. Stud. 26, 1765 – 1775. (doi:10.15244/pjes/65839)

60. Chen SM. 1996 New methods for subjective mental workload assessment and fuzzy risk analysis. Cybern. Syst. 27, 449 – 472. (doi:10.1080/019697296126417)

61. Chen SJ, Chen SM. 2001 A new method to measure the similarity between fuzzy numbers. In Proc. 10th IEEE Int. Conf. on Fuzzy Systems, Melbourne, Australia, 25 December, pp. 1123 – 1126. IEEE.

62. Wei SH, Chen SM. 2009 A new approach for fuzzy risk analysis based on similarity measures of generalized fuzzy numbers. Expert Syst. Appl. 36, 589 – 598. (doi:10.1016/j.eswa.2007.09.033)

63. Xu Z, Shang S, Qian W, Shu W. 2010 A method for fuzzy risk analysis based on the new similarity of trapezoidal fuzzy numbers. Expert Syst. Appl. 37, 1920 – 1927. (doi:10.1016/j.eswa.2009.07.015)

64. Hejazi SR, Doostparast A, Hosseini SM. 2011 An improved fuzzy risk analysis based on a new similarity measures of generalized fuzzy numbers. Expert Syst. Appl. 38, 9179 – 9185. (doi:10.1016/j.eswa.2011.01.101)

65. Khoshdel HA, Nikfalazar S. 2016 An improved similarity measure for generalized fuzzy numbers and its application to fuzzy risk analysis. Appl. Soft Comput. 52, 476 – 486. (doi:10.1016/j.asoc.2016.10.020)

66. Li J, Zeng W. 2017 Fuzzy risk analysis based on the similarity measure of generalized trapezoidal fuzzy numbers. J. Intell. Fuzzy Syst. 32, 1673 – 1683. (doi:10.3233/JIFS-141907)

67. Zeng J, Huang G. 2017 Set pair analysis for karst waterlogging risk assessment based on AHP and entropy weight. Hydrof. Res. 49, 1143 – 1155. (doi:10.2166/hh.2017.265)

68. Cheng X, Long R, Chen H, Li W. 2018 Green competitiveness evaluation of provinces in China based on correlation analysis and fuzzy rough set. Ecol. Indic. 85, 841 – 852. (doi:10.1016/j.ecolind.2017.11.045)

69. Wang Y, Yin X, Jing H, Liu R, Su H. 2016 A novel cloud model for risk analysis of water inrush in karst tunnels. Environ. Earth Sci. 75, 1459. (doi:10.1007/s12665-016-4666-7)

70. Li L, Lei T, Li S, Zhang Q, Xu Z, Shi S, Zhou Z. 2015 Risk assessment of water inrush in karst tunnels and software development. Arab. J. Geosci. 8, 1843 – 1854. (doi:10.1007/s12517-014-1365-3)