



Enhanced safety prediction of vault settlement in urban tunnels using the pair-copula and Bayesian network



Xianguo Wu^{a,1}, Zongbao Feng^{a,1}, Yang Liu^{b,c,*}, Yawei Qin^{a,d,**}, Tingyou Yang^e, Junchao Duan^{a,e}

^a School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China

^b ZhongNan Hospital of Wuhan University, Wuhan University, Wuhan 430071, PR China

^c School of Economics and Management, Wuhan University, Wuhan 430072, PR China

^d Lanzhou University of Science and Technology, Lanzhou, Hansu 730050, China

^e China Construction Third Engineering Bureau Group CO., Ltd, Wuhan 430065, China

ARTICLE INFO

Article history:

Received 21 April 2022

Received in revised form 6 September 2022

Accepted 5 October 2022

Available online 13 October 2022

Keywords:

Urban tunnel

Vault settlement

Safety prediction

Pair-copula Bayes

Key risk factors

ABSTRACT

To analyze and study the key control factors of the risk of vault deformation in tunnels excavated at small clear distances, a risk assessment method based on a pair-copula Bayesian network (PCBN) model is proposed. Based on an analysis of the factors affecting the settlement of the vault of an excavated tunnel, an index system for the risk assessment of the vault settlement is established. A PBN model based on a pair-copula function and BN can effectively deal with the complex risk system and the correlation problems within the risk system. Using the constructed PBN model, a risk analysis of the dome settlement caused by the excavation in the Donghu Tianshan tunnel in Wuhan is carried out, and the risk status of the dome settlement of the tunnel is determined to be basically safe. It can be determined from this research that a pair-copula correlation concept combined with a BN retains the strength of both concepts, and a safety risk analysis method for the settlement of tunnel vaults based on the PBN model is proposed to perform real-time and effective safety risk assessment and provide decision support for the shield construction stage. Through a correlation analysis of the risk indicators of the risk system, it is determined that the risk indicators that have the most significant impact on the deformation of the vault are the state of the groundwater, the complexity of the construction environment, the soil quality of the tunnel arch bottom, and the soil quality of the tunnel vault. These risk indicators are used as key risk indicators for decision-making. Based on this method, decision-making suggestions for reducing the risk of vault deformation are proposed. Good results are achieved in project implementation, and safety management of the vault deformation of a small clear excavated tunnel is provided.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

With rapid economic development, cities and densely populated urbanized areas face a very large demand for subway facilities [1]. For the construction of urban tunnels, a small clear-distance undercut tunnel is a commonly adopted structural form.

* Corresponding author at: ZhongNan Hospital of Wuhan University, Wuhan University, Wuhan 430071, PR China; School of Economics and Management, Wuhan University, Wuhan 430072, PR China.

** Corresponding author at: School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China; China Construction Third Engineering Bureau Group CO., Ltd, Wuhan 430065, China.

E-mail addresses: dabailiu@whu.edu.cn (Y. Liu), qinyawei@hust.edu.cn (Y. Qin).

¹ Xianguo Wu and Zongbao Feng contributed equally to the work.

This form has many advantages and can be used for tunnel excavation under relatively specific geological conditions [2]. However, during the process of tunnel excavation, the excavation machine easily causes the surface of the vault to move during engineering excavation [3], producing changes in the displacement of the vault, leading to stress concentration [4]. At the same time, as excavation continues, the surrounding shear stress will continue to increase [5]. The stress of the nearby soil layer changes with the change in the stress until it reaches a new equivalent state with the stress and stabilizes [6]. In addition, during the construction process, the buried depth of the tunnel, the size of the excavated section [7], the construction method and the lining support method can all affect the magnitude of the settlement and deformation of the tunnel vault [8]. Therefore, real-time dynamic safety assessment and decision-making for the settlement of a vault caused by subway tunnel excavation is important for engineering.

Nomenclature	
Abbreviation	
PCBN	Pair-copula BN
BNs	Bayesian networks
AI	Artificial intelligence
ML	Machine learning
AIC	Akaike information criterion
BIC	Bayesian information criterion
DAG	Directed acyclic graph
ANNS	Artificial neural networks
SVMs	Support vector machines
DDF	Density distribution function
Nomenclature	
$v = (v_1, \dots, v_j, \dots, v_d)$	A d -dimensional vector
v_j	Any one of the vectors $1 \leq j \leq d$
$C_{uv_j v_{-j}}$	Binary copula function.
$F(x_1, x_2, \dots, x_n)$	Joint distribution $F(x_1), F(x_2), \dots, F(x_n)$ of a marginal distribution
$X_{pa(t_j)}$	Set of parent nodes of random variable X in period t_j
T	The period length
V	Vertex set
E	The edge set
D^m	The moral graph of D
SD	The result of normalizing the deformation value monitored during tunnel construction
OD	The actual deformation value monitored during tunnel construction
MX	Maximum values of the actual deformation value
MN	Minimum values of the actual deformation value
N	The number of measured data features of the risk index
k	The number of optimal functions
z	The dimension of all indicator measurement values
d_k^2	The ranking difference in risk indicator measurement values.
$F^*(x)$	The optimal distribution function of the risk index
$F_n(x)$	The fitted distribution function of the measured value
$\gamma_{X,Y}$	The linear correlation coefficient
$\sigma_X^2 \sigma_Y^2$	The variance of two random variables (X, Y)
$Cov(X, Y)$	The covariance of two random variables (X, Y)
X1	Surrounding rock geology
X2	Tunnel design
X3	Construction level
X4	Organizational management
C1	Soil quality of the arch top
C2	Soil quality of the arch bottom
C3	The groundwater state
C4	Buried depth of the tunnel
C5	The influence of the slope coefficient
C6	Proper cover span ratio
C7	The construction method
C8	Support construction
C9	Tunneling speed
C10	The complexity of the construction environment
C11	The experience level of the construction team
C12	The maturity of construction technology

To ensure that subway construction will not cause damage to urban underground pipelines and surrounding buildings, thus affecting the normal life of surrounding residents [9], many researchers have performed extensive investigations on the prediction of settlement deformation caused by subway construction [10]. A survey of the related literature shows that existing research is mainly based on the use of empirical formulas, centrifugal tests, and numerical analysis to study the impact of dome settlement caused by subway tunnel excavation [11]. The empirical formula method is simple and practical and was widely used in early research [12]. For example, according to the equivalent layering method, Zhou et al. regarded an existing tunnel as an equivalent rock stratum and showed a good fit of the empirical formula to the measurement data [13]. Based on the defined longitudinal excavation coefficient y , the lateral ground settlement of the dynamic prediction model of Jiang et al. was established [14]. Tang proposed a modified Peck formula considering geological conditions to predict the surface settlement of a double tunnel and compared it with the measured data of other cases [10]. However, an empirical formula is highly subjective and is only suitable for a rough estimation of surface settlement. Many researchers have developed physical models, such as centrifugal models, to study the influence of the settlement of the vault caused by tunnel excavation. Lei et al. conducted model tests and numerical simulations to investigate the ground deformation caused by passing through overlapping tunnels [15]. Shi et al. designed and carried out a three-dimensional centrifuge test to study the three-dimensional response of the ground and pipeline during the excavation of an underlying tunnel and determined the main influencing factors of the tunnel excavation on the ground settlement [16]. Qiu et al. used a numerical simulation method to analyze the formation deformation evolution law and predicted the formation deformation of the model in combination with actual measurements [17]. However, centrifugation experiments are time-consuming and costly. With the rapid development of computer technology, various numerical simulation analysis methods, such as finite element and finite difference methods, have been widely used for tunnel construction due to their advantages in analysis and modeling [18]. Yang used a 3D numerical simulation to study the deformation characteristics of strata at different stages of tunnel excavation and explored the influencing factors to ensure the safety of ground traffic and the smooth progress of the project [19]. Numerical analysis is often carried out using simplifying assumptions for modeling analysis and does not take into account the uncertainty of multisource risk factors, which is different from the actual situation of a project, so that safety hazards during the construction process may not be identified. It is also difficult to realize real-time decision-making and security risk control [20].

2. Literature review

In recent years, innovation in the computer field has promoted the application of artificial intelligence (AI) to all industries [21], and machine learning (ML) technology has been widely used for a wide range of civil engineering problems [22]. To solve the problem of sustainable building material selection in uncertain environments, an integrated multistandard large group decision-making framework was developed [23]. Chen developed a new MCGDM bid evaluation method based on ELECTRE III, in which the generalized comparison language expression (gcle) is used to evaluate the performance of bidders [24]. Liu et al. proposed the standard and evaluation index system of the water seepage level, constructed a sample data set from the monitoring records for the construction of a random forest model, and obtained good results [25]. Some researchers have used artificial neural networks (ANNs) [11] and support vector machines (SVMs) [22] to predict the risk of shield tunnel vault settlement [26]. Santos studied the settlement of a concrete-supported tunnel of São Paulo Metro Line 2 using an artificial neural network and discussed the influence of network training parameters on the quality of the results [27]. Zhang proposed a combination of a wavelet packet transform and least squares support vector machine to improve the accuracy and reliability of the ground settlement caused by tunnels [9]. However, traditional machine learning methods, such as support vector machines and artificial neural networks [28], are inefficient and are not sufficiently accurate for analyzing the likelihood of risks in a dynamic environment. For example, the support vector product algorithm is difficult to implement for large-scale training samples, and it faces difficulties in solving multiclassification problems. Artificial neural networks cannot explain their reasoning process and the basis for their conclusions, and their theory and learning algorithm need to be further improved. Their applications usually do not consider the correlation between different indicators [29]. Correlation analysis is a method for feature selection in these machine learning techniques [30]. In addition, the variance matrix in the ML techniques changes depending on the correlation between the variables.

To comprehensively consider the uncertainties of multiple risk factors and objectively mine the data related to risk factors, more advanced methods, such as Bayesian networks (BNs), have been used [31]. BNs have diagnostic and predictive analysis capabilities, so that they are widely used in quantitative risk analysis [32]. Compared with other methods, some of the advantages of a BN are that the cumulative distribution function can perform probabilistic updates when new information is available over time [33], realize real-time reasoning between network node factors, and better solve uncertainty problems [34]. Risk assessment based on a BN model has been widely used in construction projects [35]. Wu based a BN on the risk assessment of an urban integrated pipe gallery and verified the effectiveness of the method [36]. Wang proposed a combination of numerical simulations and a BN to quantitatively analyze the factors that cause stratum settlement due to shield tunnel construction and analyzed the main parameters that affect stratum settlement [37]. Zhou conducted a risk assessment of directly buried underground sewer pipes based on BNs [38]. However, traditional BNs are more suitable for analyzing discrete node variables, and it is difficult to achieve dynamic evaluation of network nodes. Existing hybrid BNs and dynamic BNs can better solve continuous node variables and node dynamic probabilities. However, the assumption of Gaussian correlation exists in the description of node correlation, and it is difficult to construct a nonnormal correlation structure relationship between nodes.

Copula theory is a theory based on Sklar's theorem. It uses marginal functions to describe the distribution of random variables and copula functions to describe the correlation between

variables [39]. With the development of computer technology, copula theory has been effectively applied in the fields such as finance [40], seismology [41], hydrology [42], and meteorology [43]. Copula theory can more reasonably evaluate the reliability of complex coupled systems [44]. Therefore, some scholars integrate the copula concept into a BN model to realize the dynamic risk evaluation of complex systems. To reduce the uncertainty of the length of the railway interruption time [45], Hashemi used a copula BN model to perform probabilistic safety analysis on the multiprobability of process facilities [46]. By combining copula functions and BNs, Yu analyzed multiple environmental risk indicators of large reservoirs, providing valuable decision support tool for reservoir managers and scientists [10]. Pair-copula is a high-dimensional dependency modeling method developed on the basis of copula theory. A rattan structure provides a feasible and effective modeling method for high-dimensional copula theory. Kurowicka et al. applied the pair-copula concept to BNs and obtained a new pair-copula construction method, which was defined as a pair-copula BN model (PCBN) [47]. Based on this work, this paper conducts risk analysis and qualitative and quantitative correlation analysis using an established PCBN model, realizes real-time assessment of shield construction risks and identification of key safety control factors, and provides decision support for construction safety management.

The main research questions are as follows: (1) How can a PCBN model be established to carry out an effective safety risk assessment of shield construction risks and provide decision-making support for it? (2) How can a risk analysis be conducted for the construction of a tunnel excavated with a small clear distance? (3) How can the risk factors in the shield construction risk evaluation index system be analyzed? This paper proposes a risk analysis method based on a pair-copula BN. Based on the measured data of the project, the construction risk of a shield tunnel is modeled. Then, based on the constructed PCBN model, risk analysis and node correlation analysis are carried out to determine the shield construction risk status and the risk factors in the risk system that are more relevant to the settlement of the tunnel vault.

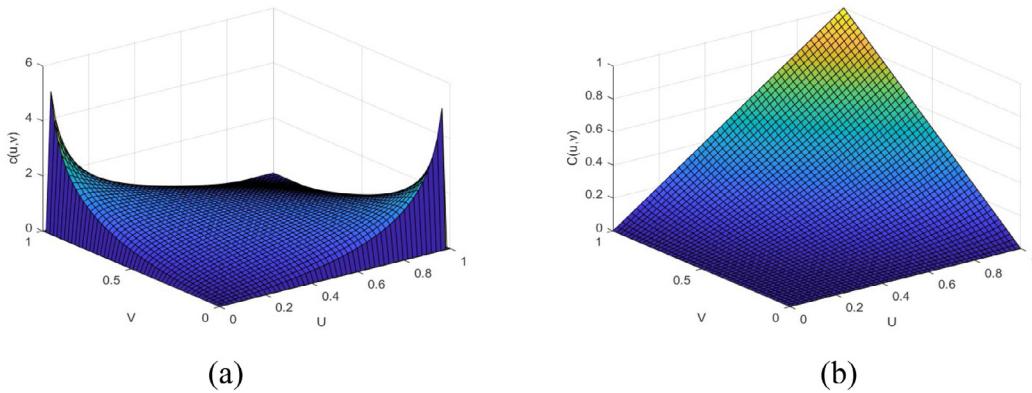
The contributions of this research are as follows: (a) Through a combination of pair-copula theory and a BN, based on the measured data, the most marginal distribution of risk indicators and the correlation coefficients between the risk indicators are determined, and a small net distance that is more suitable for an actual project is constructed. The PCBN model of the risk evaluation of the construction of an excavated tunnel can obtain more reliable evaluation results. (b) Through correlation analysis, the key factors in the risk system that are more relevant to the project construction risk are identified. Based on this, decision-making suggestions to reduce the construction risk can be proposed, providing a basis for the risk control decision.

The rest of this paper is organized as follows. Section 2 introduces the relevant basic theory of the PCBN model. Section 3 introduces the step-by-step process based on the PCBN model. Section 4 analyzes specific cases. Section 5 discusses protective measures. Section 6 summarizes the paper.

3. Methodology

3.1. Pair-copula theory

A copula function can be used to characterize the correlation between two random variables [48]. The copula method is a commonly used method based on Sklar's theorem to construct a multivariate joint distribution model. Sklar's theorem states that $F(x_1, x_2, \dots, x_n)$ is the joint distribution $F(x_1), F(x_2), \dots, F(x_n)$

**Fig. 1.** Gaussian copula function.

of a marginal distribution and that there exists for all real n-dimensional copulas, x_1, x_2, \dots, x_n [49] that can be expressed using formula (1).

$$F(x_1, x_2, \dots, x_n) = C[F_1(x_1), F_2(x_2), \dots, F_n(x_n)] \quad (1)$$

It is written in the form of a joint density function:

$$f(x_1, x_2, \dots, x_n) = f_1(x_1) \cdot f(x_2|x_1) \cdots f(x_n|x_1, \dots, x_{n-1}) \quad (2)$$

Based on this idea, by multiplying the pair-copula density function of each edge condition probability density function of a variable, the multivariate joint distribution density function of the variable can be obtained. The specific form is as follows:

$$f(u|v) = c_{uv|v_{-j}}(F(x|v_j), F(v_j|v_{-j})) \cdot f(u|v_{-j}) \quad (3)$$

In the above formula, $v = (v_1, \dots, v_j, \dots, v_d)$ represents a d-dimensional vector, v_j is any one of the vectors $1 \leq j \leq d$ and $v_{-j} = (v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_d)$. This means that if v_j is removed from v and the distribution function is restored with the density function (3), then formula (4) is the result.

$$f(u|v) = \frac{\partial C_{uv|v_{-j}}(F(x|v_j), F(v_j|v_{-j}))}{\partial F(u|v_{-j})} \quad (4)$$

where $C_{uv|v_{-j}}$ is a binary copula function.

A Gaussian copula can easily construct the dependency relationship to ensure the accuracy of the copula function to the original parameter data [50]. An example is provided in Fig. 1, which shows the two-dimensional frequency histogram of the data obtained using a Gaussian copula function model for simulation. From the probability density function graph, it can be intuitively seen that the Gaussian copula function has strong tail clustering and exhibits relatively regular tail symmetry. This shows that a Gaussian copula function can capture the correlation between random variables well.

3.2. BN

A BN reflects the connection relationship between complex variables in specific problems through directed acyclic graphs [51]. It is a network model based on probability theory with strong uncertain knowledge expression and reasoning capabilities for any BN [52]. For random variable $X = (x_{t1}, x_{t2}, \dots, x_{tn})$, the corresponding joint probability distribution can be expressed in the following form [53]:

$$f(x_{t1}, x_{t2}, \dots, x_{tn}) = \prod_{\forall t_j \in T} f(x_{t_j} | X_{pa(t_j)}) \quad (5)$$

In formula (5), $X_{pa(t_j)}$ is the set of parent nodes of random variable X in period t_j , and T represents the period length [47].

If the random variable X is a time series and the parent node of x_{ti} is $(x_{t_{j-1}}, \dots, x_{t1})$, then formula (5) can be transformed into formula (6) [54].

$$f(x_{t1}, x_{t2}, \dots, x_{tp}) = \prod_{\forall t_j \in T} f(x_{t_j} | x_{t1}, \dots, x_{t_{j-1}}) \quad (6)$$

3.3. PCBN model

A pair-copula BN relies on the conditional independence between a DAG coded variable and uses the conditional probability density parameterized by the copula function to connect the nodes of the network for continuous data to accurately control the form of the univariate marginal distribution and describe the degree of node dependence [55]. Let $D = (V, E)$ denote the DAG corresponding to a BN, V is the vertex set, E is the edge set, and D^m is the moral graph of D .

Let P denote an absolute continuous probability measure on R^d , where $d := |V|$. Let X be a random variable with probability distribution P . If the probability density function of P is f and P satisfies the d-Markov property, Sklar's theorem shows that the probability density distribution function (DDF) of P can be uniquely decomposed into a series of univariate marginal distributions F_1, F_2, \dots, F_d and a copula function C . Bauer [56] proved that copula function C can be further decomposed into a series of (conditional) pair-copula functions $C_{v,\omega|pa(v)}$, where $v \in V$, $\omega \in pa(v)$. A BN provides the copula with a new type of PCC model, in which each (conditional) pair-copula corresponds to an edge $\omega \rightarrow v$ in network structure D , connecting a node with its parent node. Therefore, the probability density function f of P is finally decomposed into Eq. (7):

$$\begin{aligned} f(x) = & \prod_{v \in V} f_v(x_v) \prod_{\omega \in pa(v)} C_{v,\omega|pa(v;\omega)} \\ & \times (F_{v|pa(v;\omega)}(x_v | x_{pa(v;\omega)}), F_{\omega|pa(v;\omega)}(x_\omega | x_{pa(v;\omega)})) \end{aligned} \quad (7)$$

where $x = (x_v)_{v \in V} \in R^d$. This model is called the PCBN model.

3.4. PCBN model of interaction risk during the construction of an excavated tunnel

This paper proposes a pair-copula Bayesian method that combines copula functions and BNs to assess the interaction risk during the construction of an excavated tunnel. Fig. 2 shows the overall workflow of the proposed risk assessment method. It includes three main stages: (1) PCBN model design; (2) risk analysis; and (3) correlation analysis.

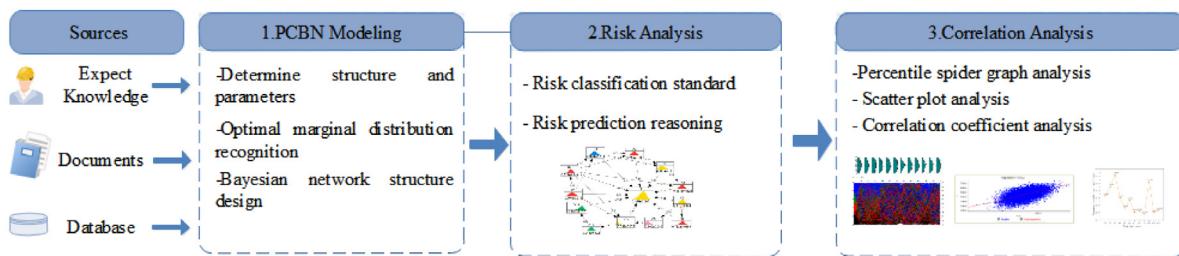


Fig. 2. The dynamic evaluation process of the construction risk of an undercut tunnel based on the PCBN model.

Table 1
Expression and statistical characteristics.

Function type	Normal	Weibull	Gamma	Exponential
Expression	$\Phi\left(\frac{x-p}{q}\right)$	$1 - \exp\left[-\left(\frac{x}{p}\right)^q\right]$	$\frac{1}{\Gamma(q)} \int_0^{px} t^{q-1} e^{-t} dt$	$1 - e^{-px}$
Variance σ^2	$\sigma^2 = q^2$	$\sigma^2 = p^2 \left[\Gamma\left(1 + \frac{2}{q}\right) - \Gamma^2\left(1 + \frac{1}{q}\right) \right]$	$\sigma^2 = \frac{q}{p^2}$	$\sigma^2 = \frac{1}{q^2}$
Mean μ	$\mu = p$	$\mu = p\Gamma\left(1 + \frac{1}{q}\right)$	$\mu = \frac{q}{p}$	$\mu = \frac{1}{p}$

3.4.1. Constructing a tunnel deformation risk assessment model

In this paper, a BN model combined with a pair-copula function is used to establish a risk assessment model for the vault deformation of a small clear-distance undercut tunnel based on the PCBN model.

1. Data preprocessing and fitting

(1) Data preprocessing

① Data collection

Due to the monitoring method and technology, the monitored data are abnormal. To solve this problem, the tunnel deformation data obtained from monitoring are normalized to reduce the inaccuracy of the analysis results caused by abnormal monitoring data, thereby increasing the reliability and accuracy of the PCBN model.

② Normalized processing of measured data

The monitored deformation data from the tunnel construction process are normalized, and a risk assessment of the tunnel vault deformation is carried out using formula (8).

$$SD = [OD - (1-d)MN] / [(1+d)MX - (1-d)MN] \quad (8)$$

where SD is the result of normalizing the deformation value monitored during tunnel construction; OD is the actual deformation value monitored during tunnel construction; MX and MN are the maximum and minimum values of the actual deformation value, respectively; and d is the confidence interval of the actual value of the deformation monitored during the construction of the tunnel.

③ Data distribution law fitting

① Alternative optimal marginal distribution function

The optimal function among the normal, Weibull, gamma, and exponential functions that conform to the distribution law of the measured data is selected. The expressions and statistical characteristics of the above four functions are shown in Table 1.

② Determine the optimal distribution function

The quality of the risk index evaluation is closely related to the evaluation criteria. Due to their simple evaluation and high evaluation effectiveness, the Akaike information criterion and Bayes information criterion are the most commonly used methods in related scientific research. To ensure the scientific rationality of the risk index distribution function and the accuracy of the PCBN model, the risk index distribution function in the PCBN model is determined by comparing the selection results of the Akaike and Bayes information criteria [57], the measured risk index data are adjusted, and the distribution function of the optimal risk index limit minimizes the baseline value of

the Akaike and Bayesian information. The calculation process is shown in formulas (9)–(11):

$$-2 \sum_{i=1}^N \ln c(u_{1i}, u_{2i}) + k \ln N \quad (9)$$

$$-2 \sum_{i=1}^N \ln c(u_{1i}, u_{2i}) + 2k \quad (10)$$

$$\begin{cases} u_{1i} = \frac{\text{rank}(x_{1i})}{N+1} & i = 1, 2, \dots, N \\ u_{2i} = \frac{\text{rank}(x_{2i})}{N+1} \end{cases} \quad (11)$$

where N represents the number of the measured data features of the risk index, and k represents the number of optimal functions. In this paper, k is set to 1.

2. Structural design of the BN

(1) Construction of the tunnel vault deformation risk assessment map: Combining the tunnel deformation research results and related monitoring data, directional connecting lines are used to connect the parent and child nodes that cause the tunnel vault deformation risk indicators to form a directed and acyclic risk assessment network;

(2) Calculation of the correlation between the indicators: According to the corresponding indicator measurement value, formula (12) is used to calculate the Spearman correlation coefficient β between the risk indicators.

$$\beta = 1 - \frac{6 \sum_{k=1}^z d_k^2}{z(z^2-1)} \quad (12)$$

where z is the dimension of all indicator measurement values, and d_k^2 is the ranking difference in risk indicator measurement values.

3. Model verification

To verify the suitability of the pair copula model, it is necessary to verify the fit of the constructed pair copula network model. The K-S, A-D and RMSE methods are used to calculate the degree of fit of the edge distribution of a node in the copula-Bayes network. Formula (13) describes the statistics of the K-S method:

$$T = \sup_x |F^*(x) - F_n(x)| \quad (13)$$

where $F^*(x)$ is the optimal distribution function of the risk index and $F_n(x)$ is the fitted distribution function of the measured value.

$F^*(x)$ is substituted into formula (14) to calculate the statistical characteristics of the distribution function and test whether the sample obeys the selected optimal distribution.

$$W_n^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \left\{ \log F^*(x_i) + \log (1 - F^*(x_{n+1-i})) \right\} \quad (14)$$

Formula (15) expresses the root mean square error value (RMSE) of the sample value x_k^{est} with total sample size N and monitoring value x_k^{obs} :

$$RMSE = \sqrt{\frac{\sum_{k=1}^N (x_k^{est} - x_k^{obs})^2}{N}} \quad (15)$$

4. Risk assessment

The standardized measurement data of the risk indicators are input into the established PCBN model, the risk distribution function that meets the qualitative rules is calculated, the digital characteristics of the function are calculated and compared with the risk level interval, and finally, the vault deformation risk level is determined.

3.4.2. Correlation analysis

Because previous methods cannot accurately analyze the relationship between risk factors, their simulation results are distorted. A pair-copula model is used in this paper that can effectively capture the correlation between the parameters. Combining BN theory, the accuracy correlation was accurately identified and reconstructed. Based on 12 risk indicators in the constructed PCBN model, vault deformation was evaluated for risk, and percentile spider graph and other analytical methods were used to analyze the common risk factors related to the magnitude of the vault deformation in the risk system, and the correlation ranking of each index obtained can be used to guide an operation team to focus on the more important risk factors.

1. Percentile spider graph

Percentile spider graph [2] reflects the relationship between various risk indicators and the deformation of the vault during tunnel construction and reflects the interrelationship between the various factors that affect the deformation of the tunnel vault.

2. Scatter chart

In this paper, joint distribution scatter diagrams of the risks C_i affecting the safety of tunnel construction and the vault deformation risk T during tunnel construction are used to determine the risks affecting the safety of tunnel construction and the C_i-T correlation of vault deformation during tunnel construction. Through the distribution law of the scatter diagram of each risk index and the fitted straight line, the largest risk factor affecting the deformation of the tunnel vault can be determined.

3. Correlation coefficient

Pearson's digital characteristics were used to evaluate the linear correlation between each risk index C_i and the vault deformation risk T and to measure whether the relationship was negative and significant [58]. A regression coefficient can be used to measure the degree of interpretation between two indicators. To address the problem of the nonlinear correlation in an unsafe system, Spearman's coefficient is selected to determine whether there exists a linear correlation between the risk index C_i affecting the safety of tunnel construction and the vault deformation risk of tunnel construction and to determine the sign and significance of the relationship [59].

(1) Pearson correlation coefficient [58,60]

The following formula $\gamma_{X,Y}$ represents the linear correlation coefficient of two random variables (X, Y) :

$$\gamma_{X,Y} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} \quad (3.24)$$

where $\sigma_X^2 \sigma_Y^2$ is the variance of two random variables (X, Y) and $Cov(X, Y)$ expresses the covariance of two random variables (X, Y) .

(2) Regression coefficient

The Pearson correlation coefficient can express the linear correlation of two random variables (X, Y) , and the regression coefficient can calculate the linear correlation rate of two random variables (X, Y) . The following formula represents the calculation expression of the regression coefficient a of variable Y to dependent variable X .

$$a = \frac{\sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)}{\sum_{i=1}^n (x_i - \mu_x)^2} \quad (3.25)$$

(3) Spearman correlation coefficient [59]

The Spearman correlation coefficient is usually used to characterize the nonlinear correlation between two random variables (X, Y) . The expression for the calculation of the Spearman correlation coefficient is shown in the following formula.

$$\rho_{X,Y} = 3 \left\{ \begin{array}{l} \Pr[(X_1 - X_2)(Y_1 - Y_3) > 0] \\ - \Pr[(X_1 - X_2)(Y_1 - Y_3) < 0] \end{array} \right\} \quad (3.26)$$

where (X_1, Y_1) , (X_2, Y_2) , and (X_3, Y_3) represent three different independent and identically distributed random variables, respectively. The positive and negative correlation coefficients $\rho_{X,Y}$ of the Spearman rank correlation coefficient are consistent with the connection direction of X and Y . In addition, when $\rho_{X,Y} = 0$, it indicates that variable y is independent of variable X .

4. Case study

4.1. Project overview

The East Lake Passage Project is a key project in Wuhan and is an important part of the regional transportation network connecting Hankou and Wuchang. The passage starts from the main line viaduct (Hongmiao Interchange) of the Shuidong Section of the Second Ring Line, ends at the intersection of Yujiahu Road and Yujiahubei Road, and traverses the Donghu Scenic Area. Its length is approximately 9.86 km. The tunnel section includes 6.88 km of the East Lake section and a 1.22-km-long Tuanshan section. The Tuanshan undercut tunnel is designed as a two-way six-lane tunnel, and the distance between the two tunnels is between 10.6 and 16.7 m, which is a small clear-distance undercut tunnel, and the maximum tunnel depth is 50 m. In addition, the geological conditions along the Tuanshan Tunnel are complex and changeable. The V-grade surrounding rock along the entire line accounts for approximately 70% of the total length of the tunnel. Most of the rocks distributed at the site are soluble carbonate rocks. The lake is rich in groundwater and has large reserves of karst water. Tunnel excavation is prone to major disasters such as water inrush and collapse. Fig. 3 shows the effect of the implementation of Tuanshan's small clear-distance wide-body tunnel undercut.

4.2. PCBN model construction

1. Establishment of a risk indicator system

Research and analysis show that aspects of the surrounding rock and soil quality, tunnel design [2], construction level and organization management are the main factors affecting the deformation of tunnel construction vaults [61].

Surrounding rock geology X1: the disturbance due to tunnel construction on the stress field of the surrounding rock geology cannot be restored to the original state in a short time [62], so that the soil stress released during excavation will inevitably change the initial geological stress state [63]. The soil quality (C1, C2) of the arch top differs strongly from that of the arch



Fig. 3. The effect of the implementation of Tuanshan's undercut wide-body tunnel with a small clear distance.

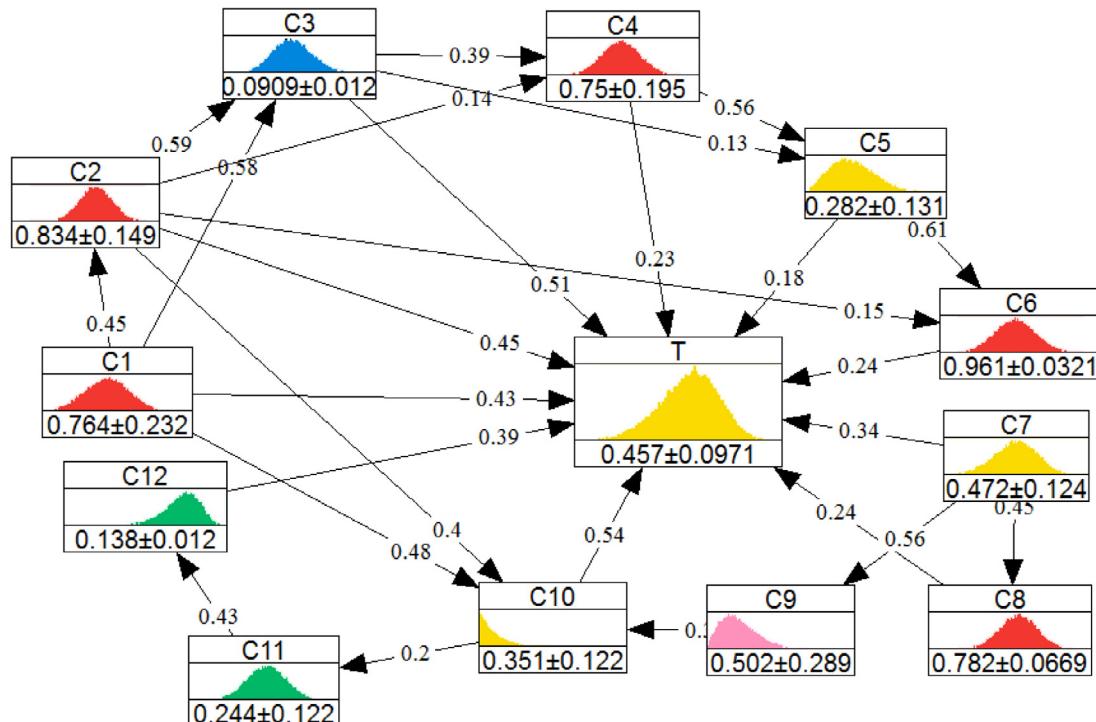


Fig. 4. Directed acyclic association model.

bottom, and all soil parameters may change in different sections. In addition, the erosion effect of the groundwater state (C3) on a tunnel structure will damage the lining concrete, particularly the weak parts such as construction joints, thus posing a serious threat to the overall quality.

Tunnel design X2: since the actual stress distribution in the underground surrounding rock state and disturbed soil cannot be accurately predicted, it is necessary to adapt the original design plan to ensure the feasibility and safety of the construction plan [64]. The stress status of the structure and the deformation of the ground surface are affected by the burial depth of the tunnel (C4), and the internal force of the structure increases with increasing burial depth, threatening the safety of the structure under force [65]. The influence of the slope coefficient (C5) on the stress field of the soil is more complicated, because it makes the soil uplift or subside to different degrees [2]. In addition, choosing a proper cover span ratio (C6) is beneficial for reducing surface deformation and stabilizing the safety state of the tunnel structure [10].

Construction level X3: The construction method (C7), support construction (C8), and tunneling speed (C9) are several common influencing factors that characterize the construction level [66]. Different construction methods have different construction characteristics. Choosing appropriate construction methods according

to the actual situation of a specific project is beneficial to reducing ground settlement. Excavation and support construction is a key part of preventing tunnel collapse and water leakage. Choosing a suitable excavation speed for the project's working conditions is conducive to the formation of the overall stress state in the early stage, as well as for improving the safety of the structure [2].

Organizational management (X4): many specialties and a large number of personnel are involved in underground engineering construction. Establishing an efficient organization and management system is an important part of safe construction [67]. The complexity of the construction environment (C10) is closely related to engineering risks and has an important effect on the status of personnel, construction materials, and construction machinery. Since an actual operation process often encounters various emergencies, the experience level (C11) of the construction team is highly important for the safety of the project. The maturity of construction technology (C12) is a summary of an organization's experience in successfully applying various construction methods to engineering cases and is an important factor ensuring the smooth progress of tunnel construction.

2. PCBN model design and verification

In this paper, a long-distance tunnel with a small clear distance is taken as an example. In Section 4.2, an evaluation index system for the settlement of the tunnel vault was constructed,

Table 2

Function types and test characteristic values of various risk indicators.

Index	A-D		K-S		μ value	Root mean square error	σ value	Function type	Judgment criteria	
	P value	Wn2	P value	T value					AIC	BIC
C1	0.982	0.190	0.639	0.069	0.230	2.139	0.2302	Weibull	62.39	65.17
C2	0.638	0.271	0.618	0.084	0.176	0.590	0.1501	Normal	40.36	45.29
C3	0.426	0.408	0.434	0.108	0.465	1.129	0.3840	Gamma	82.36	85.79
C4	0.855	0.236	0.687	0.081	0.177	0.361	0.1952	Normal	-76.22	-70.41
C5	0.826	0.244	0.678	0.083	0.175	0.364	0.2320	Weibull	-33.66	-27.13
C6	0.930	0.163	0.809	0.067	0.356	2.618	0.2330	Gamma	36.81	38.07
C7	0.631	0.207	0.619	0.060	0.215	3.208	0.1190	Weibull	109.34	116.83
C8	0.323	0.413	0.484	0.092	0.215	2.134	0.1222	Normal	-52.47	-48.93
C9	0.748	0.215	0.598	0.087	0.155	1.604	0.2131	Weibull	-73.52	-71.39
C10	0.201	0.507	0.285	0.120	0.3514	1.238	0.1212	Exponential	69.47	75.26
C11	0.681	0.266	0.682	0.084	0.203	3.744	0.1025	Normal	134.58	137.69
C12	0.659	0.237	0.574	0.125	0.234	3.567	0.1205	Weibull	82.72	84.07

and the 12 identified risk factors were used as the 12 nodes of the BN DAG graph. Combined with expert experience and relevant literature data, the DAG chart of risk evaluation is constructed, as shown in Fig. 4. Fig. 4 shows the correlation coefficients between the indicators calculated based on the measured data of the 12 indicators comprising the PCBN model of the vault deformation risk. The monitoring data are normalized according to formula (9), and the AIC and BIC values of each risk indicator for the four candidate distributions are calculated according to formulas (10)–(12).

According to the minimum AIC and BIC value principle, the optimal marginal distribution function is determined, and the specific calculation and recognition results are shown in Table 2. At the same time, the K-S, A-D and RMSE methods are used to test the goodness of fit of the fitted results of the marginal distribution of each risk index according to the formula. The significance level is set to $\alpha = 0.05$. If the statistical P value is greater than the significance level, the assumed marginal distribution is fitted. By contrast, the assumed marginal distribution is not fitted. The final results of the test statistics T and P are shown in Table 2.

3. Risk assessment of vault deformation

Using the risk value calculated based on the constructed PCBN model, the subsurface excavated tunnel vault settlement risk status can be divided into five risk levels, and the vault settlement risk level is mapped to [0, 1] by combining engineering practice and expert experience. The intervals correspond to safe, safer, basic safe, generally dangerous, and dangerous in order, as shown in Table 3. The average value of the vault deformation risk of the small clear-distance excavated tunnel is 0.457, the calculated value of the dispersion coefficient is 0.0971, the final maximum value of the vault deformation risk is 0.5541, and the minimum value is 0.3599. Combining the calculation records and Table 3, it is observed that the vault deformation risk of 0.5541 corresponds to vault safety level IV, and the vault deformation risk of 0.3599 corresponds to vault safety level III. It is observed from the safety state classification that the vault deformation safety state of the tunnel construction project is between the basically safe state and the more dangerous state. To comprehensively evaluate the deformation risk of tunnel vault construction, the high-risk factors affecting tunnel vault construction are then analyzed to form a comprehensive and systematic vault deformation risk system to achieve effective management and control of tunnel vault engineering.

4.3. Correlation analysis of indicators

1. Percent spider graph analysis

A percentile cobweb chart is an effective method for the analysis of the relationship between the vault settlement risk value T and the influencing factors X_i . In the cobweb chart, the abscissa represents each influencing factor, and the ordinate is the percentile value point of each factor. Fig. 5 is a percentile spider diagram under five risk levels, from which it is observed that the high-probability value and low-probability value of the vault deformation risk T during the construction of the small clear-distance tunnel are related to the tunnel vault and arch bottom. Soil quality (C1, C2), groundwater state (C3) and construction environment complexity (C10) are related. To further determine the correlation between the vault deformation risk T during the tunnel construction process and the various factors of the tunnel vault, a percentile spider graph of the joint distribution of the parameters under conditionalization is used for analysis, and the results are shown in Fig. 6, in which the high and low probability values of T are conditioned. The high-probability value is represented by the blue sample line, and the low-probability value is represented by the black sample line. When the risk value T is high, C1, C2, C3, and C10 values are also high, and are concentrated at the top of the spider graph; when the risk value T is low, C1, C2, C3, and C10 values are also low, and are concentrated at the bottom of the spider graph. It is determined that the risk value T has a fairly high level of correlation with C1, C2, C3, and C10. Therefore, the above indicators are used as the decisive indicators for the risk T of tunnel vault deformation.

2. Scatter chart

Using the updated parameters of the PCBN model of tunnel crown settlement constructed above, the joint distribution scatter diagram and fitting curve of each risk index (CI) affecting the safety of tunnel construction and the crown settlement risk T during tunnel construction are drawn to observe the dependency relationship of each risk index in the capture risk system, as shown in Fig. 7(a)–(m). According to the relevant theory, if the scatter points of the parameter distribution are distributed at an oblique 45° angle on the plane formed by X-Y or the slope of the straight line fitted by the scatter diagram is larger, it indicates a greater correlation between the parameters. Fig. 7(a)–(m) show that the relationships between vault settlement risk T and risk factors during the construction of small-spacing tunnels in the background project differ. Compared with other risk factors, the scattered points formed by the vault settlement risk T and the groundwater state (C3) factor during the construction of the small spacing tunnel are more concentrated, and the slope of the fitted straight line is close to 1 (the angle of the line with the x-axis is 45°), indicating that the vault settlement risk during the tunnel construction process is closely related to the

Table 3
Vault deformation risk grade standard.

Grade	I	II	III	IV	V
Safety state	Safety	Safer	Basically safe	More dangerous	Danger
Risk interval value	[0,0.1]	(0.1,0.25]	(0.25,0.50]	(0.50,0.75]	(0.75,1.0]

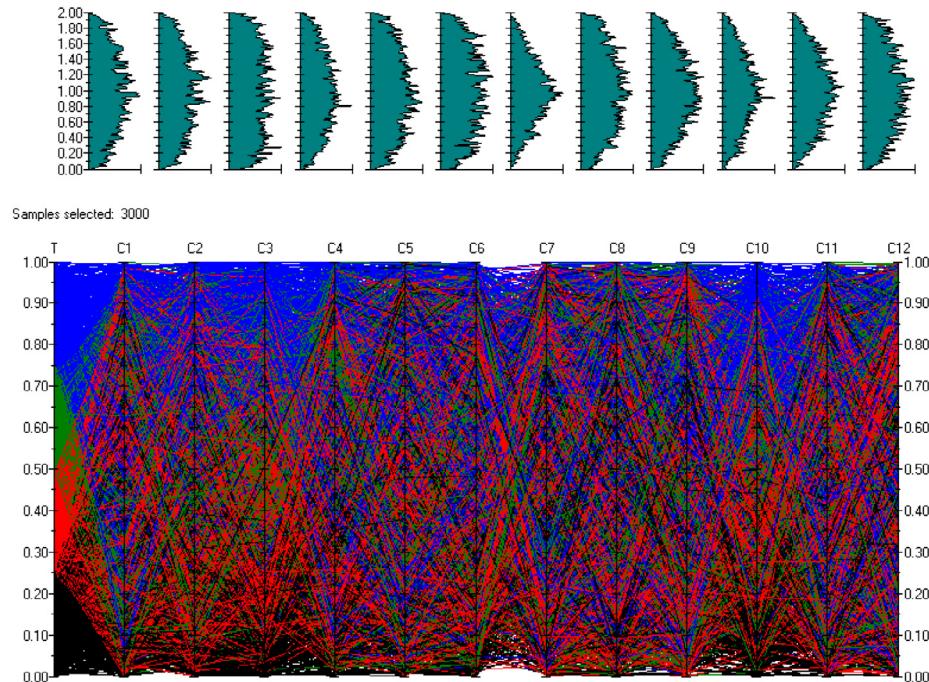


Fig. 5. Percentile spider graph.

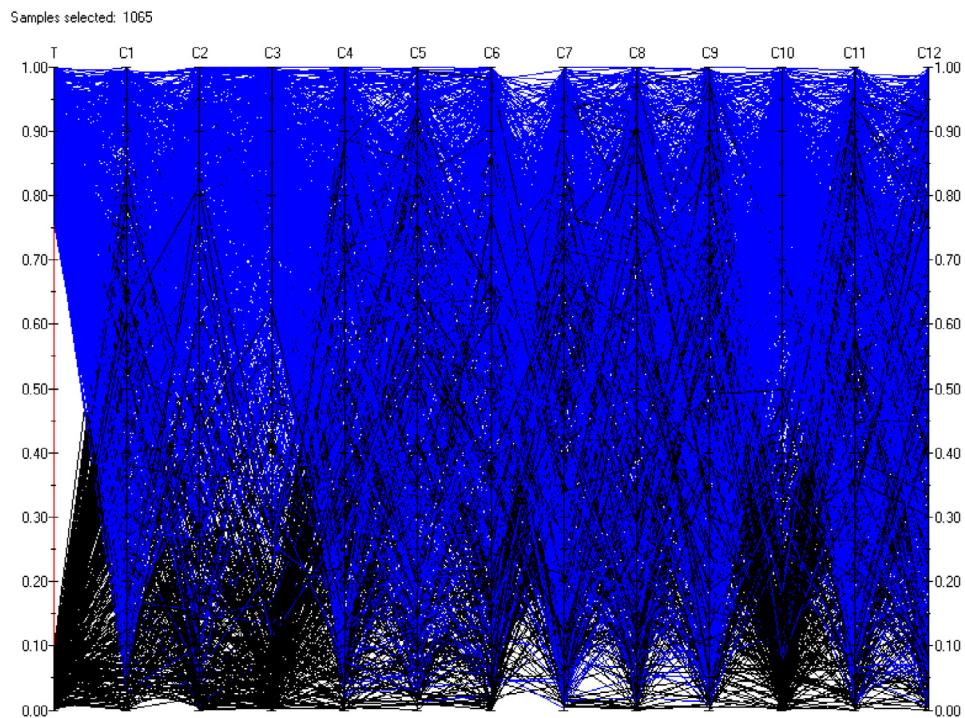


Fig. 6. Percentile spider diagram of the joint distribution of the parameters under conditionalization.

groundwater state. The scattered distribution formed by the arch crown settlement risk T during tunnel construction and the “soil at the bottom of the tunnel (C2)” presents a clear oval distribution

with an incline angle of 45° , demonstrating a strong correlation. The scattered distribution of vault settlement risk T and “soil quality of tunnel vault (C1)” and “complexity of construction

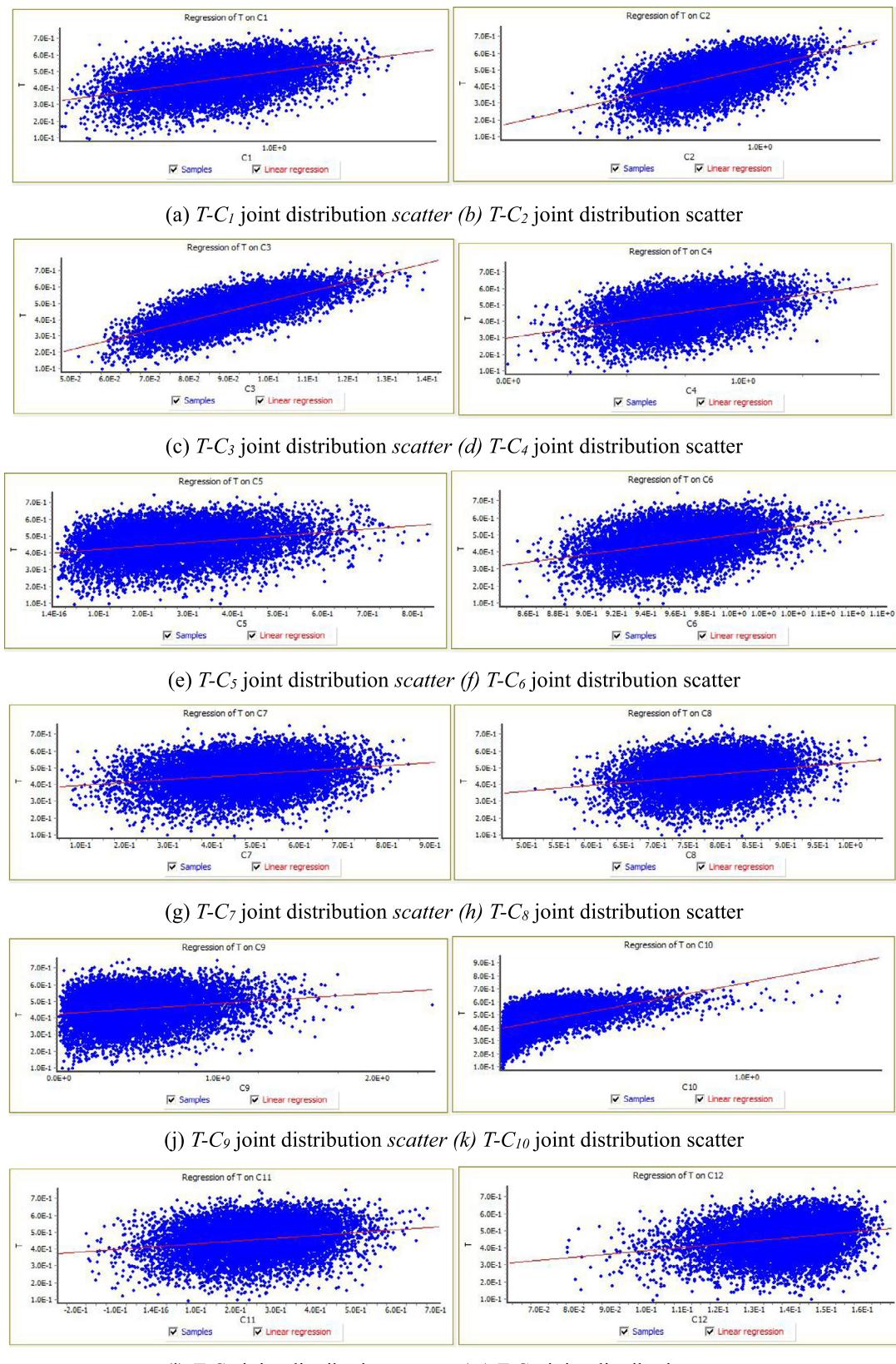


Fig. 7. Scatter plot of the joint distribution of each indicator and risk T risk.

environment (C10)" during tunnel construction is not as dense as that of vault settlement risk T and "soil quality of tunnel bottom (C2)" and "groundwater state (C3)" during tunnel construction,

but the slope of the fitting curve exceeds 45° . This shows that the vault settlement risk T in the process of tunnel construction is highly correlated with "soil quality of tunnel vault (C1)" and

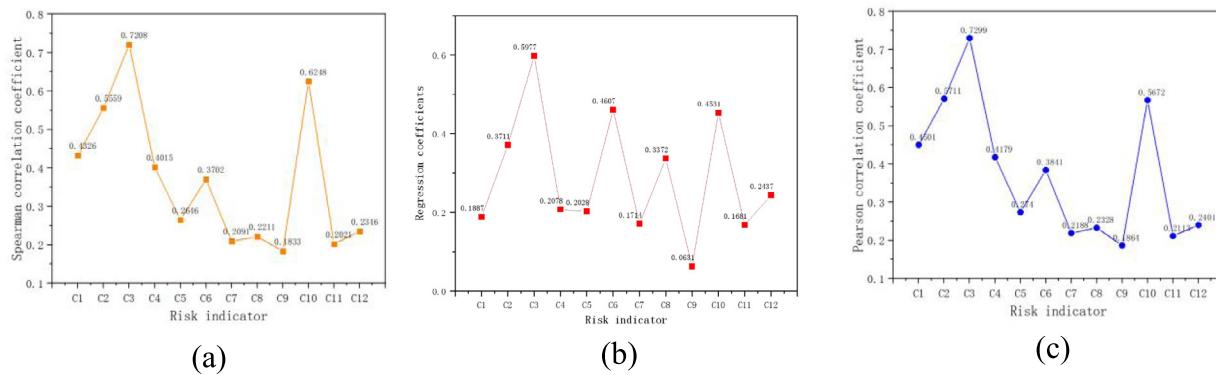


Fig. 8. Correlation coefficient of index C_i with respect to T .

“complexity of construction environment (C10)”, but not with “tunnel buried depth (C4)”, “longitudinal spacing of tunnel face (C5)”, “tunnel coverage ratio (C6)”, “construction method (C7)”, “support and reinforcement method and time (C8)”, and “advancing speed of working face (C9)”. The slopes of the fitting curves for “construction team experience level (C11)” and “construction technology maturity (C12)” are lower than 45° , and in particular, the slopes of the data for vault settlement risk T plotted versus “tunnel longitudinal spacing (C5)”, “construction method (C7)” and “working face advancing speed (C9)” approach 0, showing weak correlation. Therefore, according to the scatter diagram analysis of various indicators, in this project, the soil quality of the tunnel crown (C1), the soil quality of the tunnel bottom (C2), the state of groundwater (C3), and the complexity of the construction environment (C10) are the decision-making indicators for reducing t-risk, which is consistent with the conclusions obtained from the spider's Web diagram.

3. Correlation coefficient analysis

Based on the PCBN model constructed in this paper for the vault deformation risk of the excavated tunnel, in the calculation of the vault deformation risk system, the 12 risk indicators affecting the safety of tunnel construction are related using the Pearson correlation coefficient and the regression of the vault deformation risk T during the tunnel construction process. The calculated values of the Pearson correlation coefficient and the Spearman correlation coefficient are shown in Fig. 8. The Pearson correlation coefficient is used to measure the development trend of the degree of linear correlation between the risk index (C_i) and the vault deformation risk T by calculating a regression coefficient among the risk factors. To obtain the linear correlation slope among the factors, the Spearman correlation coefficient among the risk factors is calculated to measure the nonlinear correlation between the risk index (C_i) and the vault deformation risk T . Table 4 shows the categories of the significance of correlation based on the correlation coefficient used in this article.

According to the analysis of the linear and nonlinear correlation, C3, C10, and C2 are the most critical risk indicators of the risk system. Table 4 and Fig. 8 show that the Pearson coefficients and regression coefficients of the top three absolute values are analyzed. Among the 12 indicators that affect the risk of tunnel construction vault deformation, the Pearson coefficient values of C3, C2 and C10 are 0.7297, 0.5707, and 0.5668, respectively, and are all significantly positively correlated. In addition, the maximum value of slope a is 5.9080 for C3, indicating that C3 has the greatest correlation with the vault deformation risk T . At the same time, the relationship between the main causes of vault deformation in tunnel engineering may also be nonlinear. By observing and comparing Table 4 and Fig. 8, it is observed that the correlation coefficients of the three largest Spearman rank

Table 4
The relationship between the correlation coefficients and correlation.

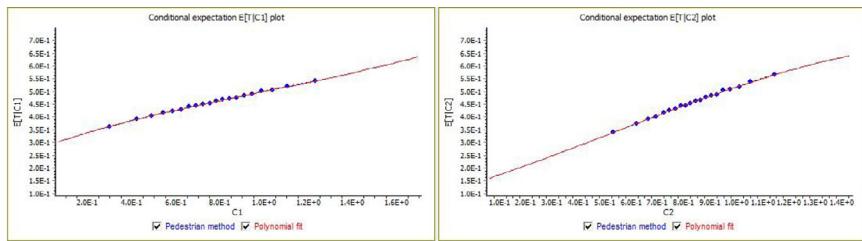
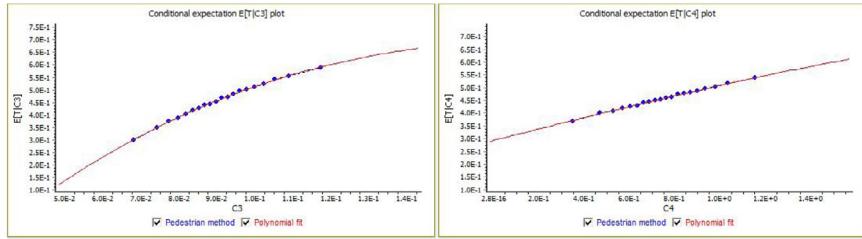
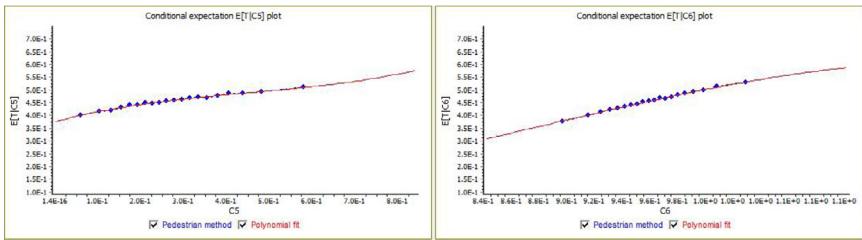
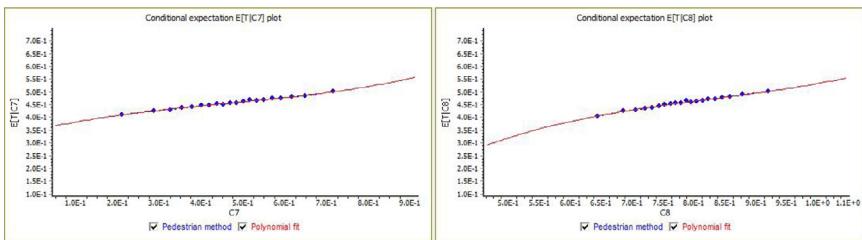
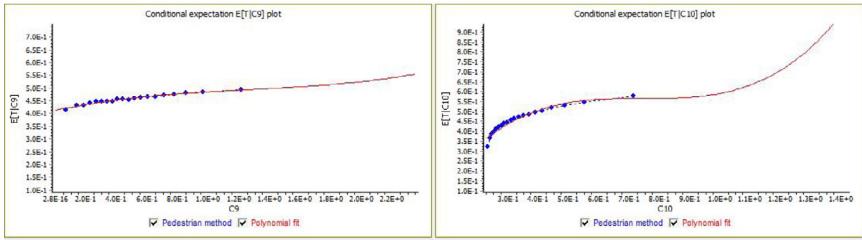
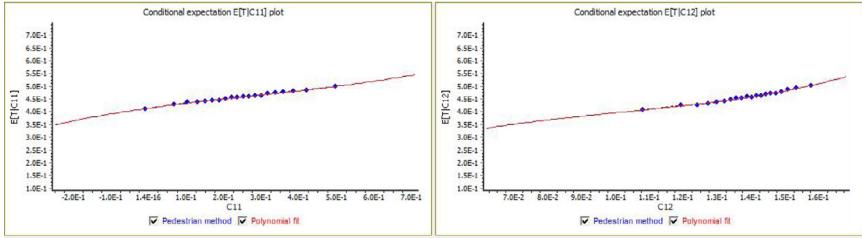
Correlation	Negative value	Positive value
Irrelevant	-0.09~0.0	0.0~0.09
Low correlation	-0.3~0.1	0.1~0.3
Moderately relevant	-0.5~0.3	0.3~0.5
Significant correlation	-1.0~0.5	0.5~1.0

correlation coefficients are C3 and C10, C2, with the values of 0.7299, 0.5711, and 0.5672, respectively, all in the range of 0.5 to 1, showing significant nonlinear correlation. In summary, C3, C10, and C2 are the key risk indicators of the safety system of tunnel construction vault deformation.

4. Tail correlation analysis

In tunnel engineering, when the risk value of vault settlement is large, the upper tail correlation of the risk index measures the impact of the changes in 12 risk indexes in the tunnel construction risk system on the risk value of the vault structure. A higher upper tail correlation corresponds to a higher risk of tunnel vault formation. In tunnel engineering, changing the risk index will cause significant changes in the risk value of the arch crown structure. Since the change risk of each risk index affecting the safety of tunnel construction at a higher value is far greater than its risk in other value ranges, the upper tail correlation index between each risk index affecting the safety of tunnel construction and the risk of vault settlement during tunnel construction has a strong correlation with the risk management of vault settlement caused by tunnel construction.

To determine the maximum upper tail phase value of each risk index, with the change in the value of the vault settlement factor (C1) affecting the tunnel construction process, the change in the upper tail phase value of each risk index (C1) affecting the safety of tunnel construction and the vault settlement risk during tunnel construction are fit. The corresponding upper tail correlation estimation diagrams of each risk index (C1) affecting the safety of tunnel construction and the vault settlement risk $t(i = 1, 2, 12)$ during tunnel construction are drawn, as shown in Fig. 9(a)–(m), respectively. According to Fig. 9, despite the differences in the upper tail correlation coefficient between each risk index (C1) affecting the safety of tunnel construction and the crown settlement risk T during tunnel construction, the same change trends are observed; that is, the value of the crown settlement factor (C1) during tunnel construction is positively correlated with the upper tail correlation coefficient, but there are differences in the growth range and growth mode. In particular, the crown settlement risk T during tunnel construction increases nearly linearly with the soil quality of the tunnel crown (C1), the soil quality of the tunnel bottom (C2), the tunnel buried depth (C4) and the tunnel coverage ratio (C6), while other indicators

(a) $T-C_1$ upper tail correlation estimation diagram. (b) $T-C_2$ upper tail correlation estimation diagram.(c) $T-C_3$ upper tail correlation estimation diagram. (d) $T-C_4$ upper tail correlation estimation diagram.(e) $T-C_5$ upper tail correlation estimation diagram. (f) $T-C_6$ upper tail correlation estimation diagram.(g) $T-C_7$ upper tail correlation estimation diagram. (h) $T-C_8$ upper tail correlation estimation diagram.(j) $T-C_9$ upper tail correlation estimation diagram. (k) $T-C_{10}$ upper tail correlation estimation diagram.(l) $T-C_{11}$ upper tail correlation estimation diagram (m) $T-C_{12}$ upper tail correlation estimation diagram**Fig. 9.** $Ci-T$ upper tail correlation estimation diagram.

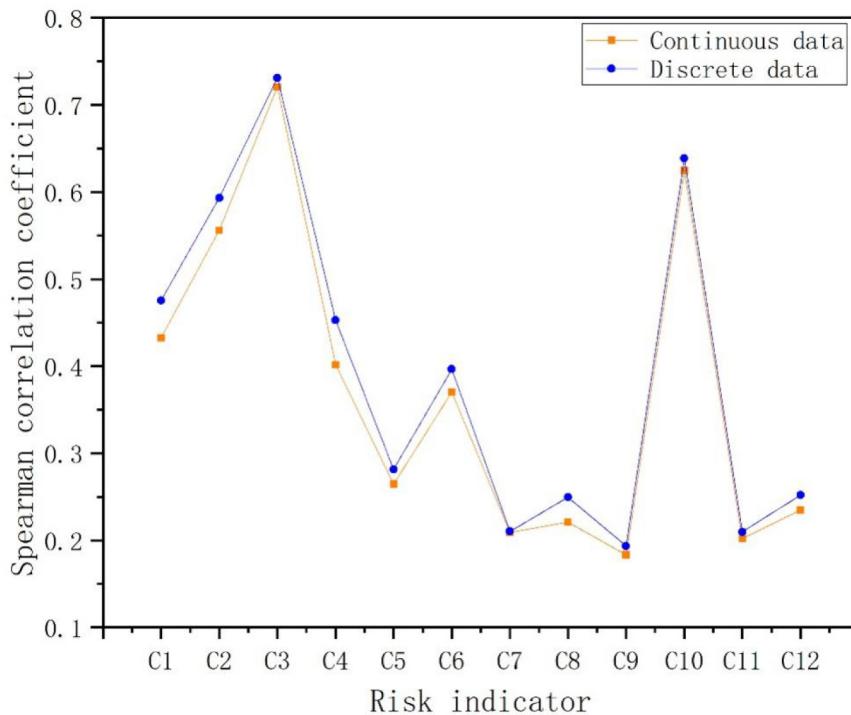


Fig. 10. C_i - T correlation under different types of data.

show a non-linear increase. Moreover, the vault settlement risk T during tunnel construction and the soil quality (C1) of the tunnel vault, the soil quality (C2) of the tunnel bottom, the groundwater state (C3) and the complexity of the construction environment (C10) show very pronounced increasing trends. These results show that in practical engineering, when the vault settlement risk t of the underground tunnel is high, the risk indicators are the soil quality (C1) of the tunnel vault, the soil quality (C2) of the tunnel bottom, and the groundwater state (C3). The increase in the complexity of the construction environment (C10) will lead to a significant increase in the risk of vault settlement during tunnel construction. Conversely, the reduction in the risk indicators soil quality of the tunnel vault (C1), soil quality of the tunnel bottom (C2), groundwater state (C3) and complexity of the construction environment (C10) will lead to a significant reduction in the risk of vault settlement during tunnel construction.

Fig. 9 shows that among the risk indicators with a larger increasing trend, namely the soil quality of the tunnel crown (C1), the soil quality of the tunnel bottom (C2), the groundwater state (C3) and the complexity of the construction environment (C10), the maximum result of the upper tail correlation coefficient is the maximum value of the upper tail correlation curve in Fig. 9(k), that is, the upper tail correlation between the crown settlement risk T during tunnel construction and the complexity of the construction environment (C10). The second largest results are presented in Fig. 9(a)–(c), the maximum values of the upper tail correlation curves, that is, the upper tail correlations between the crown settlement risk T during tunnel construction and the soil quality of the tunnel crown (C1), the soil quality of the tunnel bottom (C2), and the groundwater state C3. This shows that when the crown settlement risk t in the process of tunnel construction is high, reducing the soil quality of the tunnel crown (C1), soil quality of the tunnel bottom (C2), groundwater state (C3) and complexity of the construction environment (C10) sequentially can effectively control the risk t caused by tunnel crown settlement in the process of tunnel construction. In the process of tunnel construction, combined with the measured values of on-site tunnel deformation monitoring, when it is found that the

risk value of vault settlement is large, priority should be given to the risk of reducing the construction environment complexity (C10), which is the index with the greatest dependence on the top and tail of the vault settlement risk t in the process of tunnel construction, to effectively reduce the risk of vault settlement. If the risk of construction environment complexity (C10) of the project is low, the risk of groundwater state (C3), soil quality of the tunnel crown (C1) and soil quality of the tunnel bottom (C2) should be reduced to reduce the risk of tunnel crown settlement.

5. Discussion

5.1. C_i - t correlation under different types of data

By determining the best fitting marginal distribution, the PCBN model can effectively capture the correlation between the variables in the risk analysis of actual tunnel crown settlement. Compared with discrete data, continuous data can better reflect the characteristics and laws of data. The difference in the marginal distribution types of variables will affect the construction of the PCBN model and then will affect the results of risk analysis. This case uses the normal, Weibull, gamma and exponential distributions as alternative distribution functions to fit the characteristics of 12 risk factors. Now, 12 risk factors are considered as discrete data. By comparing the correlation between T and each risk factor C_i under two different types of data, as shown in Fig. 10, the following conclusions can be obtained.

Different marginal distribution types of the variables will affect the specific value of the correlation coefficient that determines the accuracy of the model. However, from the qualitative point of view, the dependencies between the variables have not changed significantly. The Spearman correlation coefficient in Fig. 10 is used to show the correlation strength between the variables. For both discrete data and continuous data, the overall trend of the change of the correlation coefficient among the variables is roughly the same. However, the correlation value of the discrete data variable is slightly higher than that of the



Fig. 11. Reduced risk of groundwater.

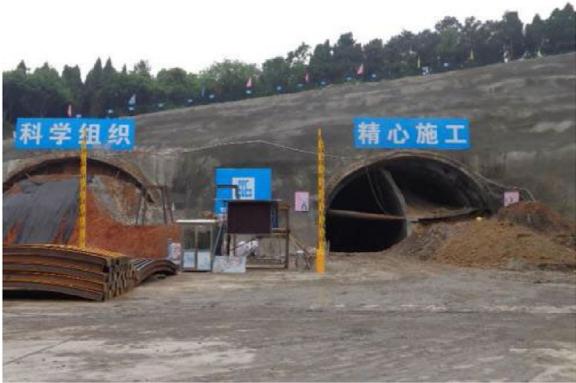


Fig. 12. Reduced complexity of tunnel construction.

continuous data variable, indicating that the model established by the discrete data distribution variable slightly overestimates the relevant risk intensity. Therefore, it is necessary to verify the marginal distribution type of each variable in the PCBN model in order to more accurately measure the correlation between the variables and more effectively provide support for the risk analysis of tunnel crown settlement.

5.2. Risk analysis results and control measures

(1) Based on the spider web diagram, scatter diagram, correlation coefficient and upper tail correlation analysis results, it is shown that the risk indicators that are highly correlated with the vault settlement risk are the key risk indicators in the vault settlement risk system of underground excavation tunnels, namely, the soil quality of the tunnel vault (C1), the soil quality of the tunnel bottom C2, the groundwater state (C3), and the complexity of the construction environment (C10). According to the calculated correlation coefficients, the order of correlation is groundwater state (C3), complexity of construction environment (C10), soil quality at the bottom of the tunnel (C2), and soil quality at the top of the tunnel (C1);

- 1) According to the upper tail correlation analysis results, the order of correlation is the complexity of the construction environment (C10), the state of groundwater (C3), the soil quality of the tunnel vault (C1), and the soil quality of the tunnel bottom (C2).

- 2) Based on the analysis and suggestions of PCBN, the goal of effectively reducing the risk of tunnel vault settlement in the process of tunnel construction is achieved by controlling the risk of vault settlement more effectively when such risk is high. Effective drainage measures (as shown in Fig. 11) can be adopted to reduce the risk caused by groundwater. The stratum with poor geological conditions should be reinforced by small conduit grouting (as shown in Fig. 12). For large-span and small-spacing tunnels, to reduce the excessive settlement of the arch crown, a temporary inverted arch and temporary steel arch of the middle diaphragm should be added during the tunnel excavation in order to reduce construction safety risks (as shown in Fig. 13). When the left and right tunnels are constructed simultaneously, to reduce the settlement risk due to the mutual influence of the tunnels in the process of simultaneous excavation, the excavations of the left and right tunnels must be at least 50 m apart (as shown in Fig. 14). After the initial lining of the tunnel is completed, the construction of the tunnel invert and secondary lining must be carried out promptly to achieve the stability of the tunnel excavation section structure and ensure that the construction risk of the tunnel vault workshop is minimized (as shown in Figs. 15 and 16). The adoption of this series of measures has effectively reduced the risk of vault settlement of small spacing tunnels and has achieved good results.

6. Conclusion

To address vault deformation that has a strong impact on the safety risk of undercut tunnel construction, this paper adopts a Bayesian model combined with a copular function to establish a risk index system, designs a Bayesian-copula model (PCBN), and performs model analysis in three steps. The highest deformation risk is evaluated. Then, the proposed PCBN model is applied to a small clear-distance undercut tunnel project in Tianshan Donghu, Wuhan, and effective assessment and decision-making of the vault deformation risk are carried out to verify the effectiveness of the method.

The conclusions of this case study are as follows: (1) The optimal marginal distribution and numerical characteristics of each risk indicator are determined using the measured data based on the copula function, the correlation coefficient characteristics of each risk indicator and the vault deformation risk are calculated, and then the BN is used according to the correlation to enable intelligent decision-making regarding the risk of tunnel construction vault deformation. (2) After the construction of the PCBN model, the vault deformation risk and risk indicators are analyzed based on a cobweb diagram, scatter diagram, and the correlation coefficients, and the key risk indicators that are more strongly related to the vault deformation risk are found to include groundwater, the soil quality of the vault and the bottom of the arch, and the complexity of the construction environment. (3) In view of the key risk factors, it is recommended to adopt effective waterproofing and drainage measures to reduce the risk of groundwater and reduce the complexity of the tunnel construction environment to reduce the risk of vault deformation. Based on the PCBN model analysis and the actual project situation, the proposed targeted suggestions achieved good results with project implementation.

It must be pointed out that the research described in this article has certain limitations. Although some multifaceted characteristic parameters have been collected, they are still not comprehensive in terms of personnel, machinery and equipment, and

materials. This affects the construction of the PCBN model, and other risk factors need to be considered in the future to establish a more complete indicator system. In addition, the case used in this paper is for Wuhan geology, whereas many different geological conditions are encountered in the subway tunnels built in China, such as the karst cave landform of Guizhou Metro Line 3. The total number of samples used in this paper is limited, and the research results will be limited by the number of samples.

Therefore, it is necessary to consider other risk factors and establish a more complete index system in the future. In the future, this method page will be applied to tunnel construction safety risk assessment for additional different geological types.

CRediT authorship contribution statement

Xianguo Wu: Methodology, Investigation, Writing – original draft. **Zongbao Feng:** Funding acquisition, Writing – review & editing. **Yang Liu:** Conceptualization, Methodology, Writing – review & editing. **Yawei Qin:** Supervision, Writing – review & editing. **Tingyou Yang:** Methodology, Writing – review & editing. **Junchao Duan:** Writing – review & editing.



Fig. 13. Preventing and controlling the arch of the front and arch tunnel.



Fig. 14. Reducing the complexity of the tunnel construction environment (The single-side cavern of the tunnel is constructed successively at an interval of more than 50 m to reduce the construction risk.)

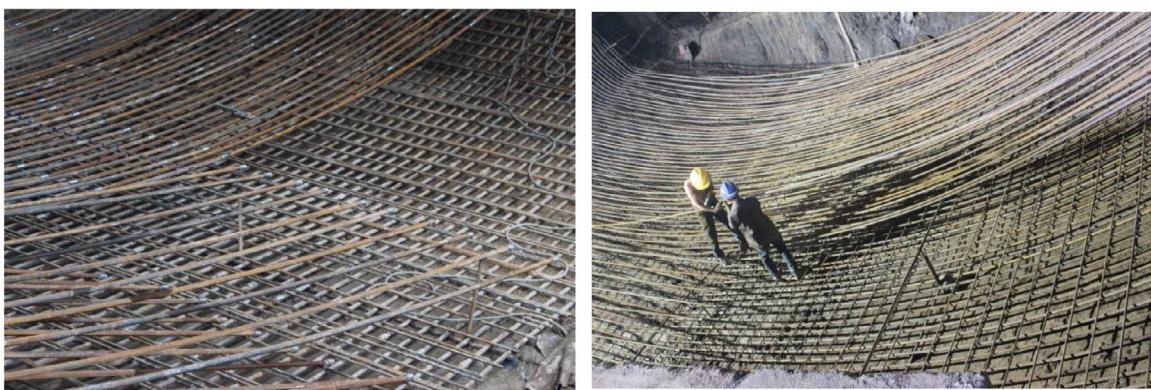


Fig. 15. Follow-up construction of the permanent reinforced concrete inverted arch.

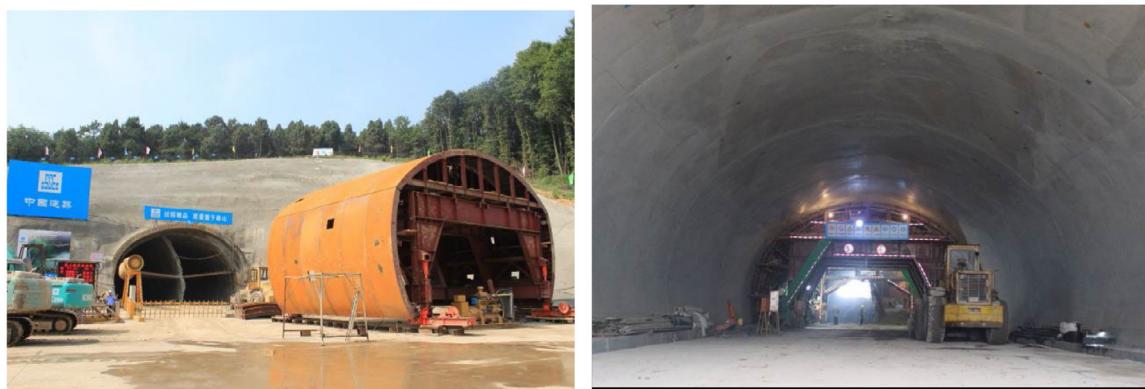


Fig. 16. Timely follow-up of the second lining concrete construction to reduce safety risks.

References

- [1] L.M. Zhang, Y. Zhang, H.X. Li, Z. Lei, Estimating long-term impacts of tunnel infrastructure development on urban sustainability using granular computing, *Appl. Soft Comput.* 113 (2021).
- [2] L.C. Yang, Y. Chen, et al., Place-varying impacts of urban rail transit on property prices in Shenzhen, China: Insights for value capture, *Sustain. Cities. Soc.* 58 (2020) 102140.
- [3] L.Y. Ding, L.M. Zhang, X.G. Wu, M.J. Skibniewski, Q.H. Yu, Safety management in tunnel construction: Case study of Wuhan metro construction in China, *Saf. Sci.* 62 (2014) 8–15.
- [4] F. Wang, H. Li, Towards reliability evaluation involving correlated multivariates under incomplete probability information: A reconstructed joint probability distribution for isoprobabilistic transformation, *Struct. Saf.* 69 (2017) 1–10.
- [5] Y.M. Zhao, J.S. Wu, R. Zhou, J.T. Cai, Y.P. Bai, L. Pang, Effects of the length and pressure relief conditions on propagation characteristics of natural gas explosion in utility tunnels, *J. Loss Prev. Process Ind.* 75 (2022).
- [6] M. Asadi, M. Eftekhari, M.H. Bagheripour, Evaluating the strength of intact rocks through genetic programming, *Appl. Soft Comput.* 11 (2011) 1932–1937.
- [7] D.X. Zhao, B.J. He, Effects of architectural shapes on surface wind pressure distribution: case studies of oval-shaped tall buildings, *J. Build. Eng.* 12 (2017) 219–228.
- [8] H.Y. Zheng, Y. Deng, Y. Hu, Fuzzy evidential influence diagram and its evaluation algorithm, *Knowl.-Based Syst.* 131 (2017) 28–45.
- [9] L.M. Zhang, L.Y. Ding, X.G. Wu, M.J. Skibniewski, An improved Dempster-Shafer approach to construction safety risk perception, *Knowl.-Based Syst.* 132 (2017) 30–46.
- [10] B.T. Yin, B.Y. Li, G. Liu, Z.Y. Wang, B.J. Sun, Quantitative risk analysis of offshore well blowout using bayesian network, *Saf. Sci.* 135 (2021).
- [11] L.M. Zhang, P.H. Lin, Multi-objective optimization for limiting tunnel-induced damages considering uncertainties, *Reliab. Eng. Syst. Safe.* 216 (2021) 107945.
- [12] Q. Fang, D.L. Zhang, Q.Q. Li, L.N.Y. Wong, Effects of twin tunnels construction beneath existing shield-driven twin tunnels, *Tunn. Undergr. Space Technol.* 45 (2015) 128–137.
- [13] Z. Zhou, Y. Chen, Z.Z. Liu, L.W. Miao, Theoretical prediction model for deformations caused by construction of new tunnels undercrossing existing tunnels based on the equivalent layered method, *Comput. Geotech.* 123 (2020).
- [14] S. Jiang, Y. Zhu, Q. Li, H. Zhou, H.L. Tu, F.J. Yang, Dynamic prediction and influence factors analysis of ground surface settlement during tunnel excavation, *Rock Soil Mech.* 43 (2022) 195–204.
- [15] H.Y. Lei, Y.N. Liu, Y.J. Zhang, Y. Hu, Ground deformation behaviour induced by overlapped shield tunnelling considering vibration loads of subway train in sand, *Acta Geotech.*.
- [16] J.W. Shi, Y. Wang, C.W.W. Ng, Three-dimensional centrifuge modeling of ground and pipeline response to tunnel excavation, *J. Geotech. Geoenvir. Eng.* 142 (2016).
- [17] M. Qiu, G. Yang, J. Duan, P. Zhang, Evolution rules and numerical simulation of stratum deformation induced by close-spaced twin tunnels shield construction, *J. Nat. Disasters* 30 (2021) 60–69.
- [18] X.X. Huang, J.Q. Chen, H.P. Zhu, Assessing small failure probabilities by AK-SS: An active learning method combining Kriging and subset simulation, *Struct. Saf.* 59 (2016) 86–95.
- [19] C.Y. Zhao, M.F. Lei, C.H. Shi, H.R. Cao, W.C. Yang, E. Deng, Function mechanism and analytical method of a double layer pre-support system for tunnel underneath passing a large-scale underground pipe gallery in water-rich sandy strata: A case study, *Tunn. Undergr. Space Technol.* 115 (2021).
- [20] R.Z. Liang, W.B. Wu, F. Yu, G.S. Jiang, J.W. Liu, Simplified method for evaluating shield tunnel deformation due to adjacent excavation, *Tunn. Undergr. Space Technol.* 71 (2018) 94–105.
- [21] F.Q. Meng, B.J. He, et al., Sensitivity analysis of wind pressure coefficients on CAARC standard tall buildings in CFD simulations, *J. Build. Eng.* 16 (2018) 146–158.
- [22] Y. Liu, Y. Cao, et al., Prediction of the durability of high-performance concrete using an integrated RF-LSSVM model, *Constr. Build. Mater.* 356 (2022) <http://dx.doi.org/10.1016/j.conbuildmat.2022.129232>.
- [23] Z.S. Chen, L.L. Yang, K.S. Chin, Y. Yang, W. Pedrycz, J.P. Chang, L. Martinez, M.J. Skibniewski, Sustainable building material selection: An integrated multi-criteria large group decision making framework, *Appl. Soft Comput.* 113 (2021).
- [24] Z.S. Chen, X. Zhang, R.M. Rodriguez, W. Pedrycz, L. Martinez, Expertise-based bid evaluation for construction-contractor selection with generalized comparative linguistic ELECTRE III, *Autom. Constr.* 125 (2021).

- [25] Y. Liu, H.Y. Chen, L.M. Zhang, X.J. Wang, Risk prediction and diagnosis of water seepage in operational shield tunnels based on random forest, *J. Civ. Eng. Manag.* 27 (2021) 539–552.
- [26] M. Saridemir, Genetic programming approach for prediction of compressive strength of concretes containing rice husk ash, *Constr. Build. Mater.* 24 (2010) 1911–1919.
- [27] O.J. Santos, T.B. Celestino, Artificial neural networks analysis of São Paulo subway tunnel settlement data, *Tunn. Undergr. Space Technol.* 23 (2008) 481–491.
- [28] K. Nadig, W. Potter, G. Hoogenboom, R. McClendon, Comparison of individual and combined ANN models for prediction of air and dew point temperature, *Appl. Intell.* 39 (2013) 354–366.
- [29] H.G. Yu, J.F. Tao, C.J. Qin, M.Y. Liu, D.Y. Xiao, H. Sun, C.L. Liu, A novel constrained dense convolutional autoencoder and DNN-based semi-supervised method for shield machine tunnel geological formation recognition, *Mech. Syst. Signal Process.* 165 (2022).
- [30] S. Kumar, I. Chong, Correlation analysis to identify the effective data in machine learning: Prediction of depressive disorder and emotion states, *Int. J. Environ. Res. Public Health* 15 (2018).
- [31] A. Qazi, M. Simsekler, Assessment of humanitarian crises and disaster risk exposure using data-driven Bayesian networks, 52, 2020.
- [32] S. Samet, A. Miri, E. Granger, Incremental learning of privacy-preserving Bayesian networks, *Appl. Soft Comput.* 13 (2013) 3657–3667.
- [33] S.H. Lee, J.E. Kang, C.S. Park, D.K. Yoon, S. Yoon, Multi-risk assessment of heat waves under intensifying climate change using Bayesian networks, 50, 2020, 101704.
- [34] C.H. Lo, Y.K. Wong, A.B. Rad, Bond graph based Bayesian network for fault diagnosis, 11, 2011, pp. 1208–1212.
- [35] D. Cinar, G.J.K.-B.S. Kayakutlu, Scenario analysis using Bayesian networks: A case study in energy sector, 23, 2010, pp. 267–276.
- [36] B. Chen, Q. Liu, H.Y. Chen, L. Wang, T.T. Deng, L.M. Zhang, X.G. Wu, Multiobjective optimization of building energy consumption based on BIM-DB and LSSVM-NSGA-II *, *J. Clean. Prod.* 294 (2021).
- [37] Y.Y. Zhang, C.F. Zhu, Q.R. Wang, LightGBM-based model for metro passenger volume forecasting, *Iet Intell. Transp. Syst.* 14 (2020) 1815–1823.
- [38] R. Zhou, W.P. Fang, J.S. Wu, A risk assessment model of a sewer pipeline in an underground utility tunnel based on a Bayesian network, *Tunn. Undergr. Space Technol.* 103 (2020).
- [39] C.Q. Guo, F. Khan, S. Imtiaz, Copula-based Bayesian network model for process system risk assessment, *Process Saf. Environ. Prot.* 123 (2019) 317–326.
- [40] K. Sukcharoen, D.J. Leatham, Hedging downside risk of oil refineries: A vine copula approach, *Energy Econ.* 66 (2017) 493–507.
- [41] K. Goda, Statistical modeling of joint probability distribution using copula: Application to peak and permanent displacement seismic demands, *Struct. Saf.* 32 (2010) 112–123.
- [42] E. Sharifi, B. Saghafian, R. Steinacker, Copula-based stochastic uncertainty analysis of satellite precipitation products, *J. Hydrol.* 570 (2019) 739–754.
- [43] A.F.S. Ribeiro, A. Russo, C.M. Gouveia, P. Pascoa, Copula-based agricultural drought risk of rainfed cropping systems, *Agricult. Water Manag.* 223 (2019).
- [44] Z.Q. Dong, G. Li, G.H. Lu, B. Song, H.N. Li, Copula-based joint probabilistic model of earthquakes and rain for the failure assessment of masonry-adobe structures, *J. Build. Eng.* 42 (2021).
- [45] X.S. Tang, D.Q. Li, C.B. Zhou, K.K. Phoon, L.M. Zhang, Impact of copulas for modeling bivariate distributions on system reliability, *Struct. Saf.* 44 (2013) 80–90.
- [46] S.J. Hashemi, F. Khan, S. Ahmed, Multivariate probabilistic safety analysis of process facilities using the copula Bayesian network model, *Comput. Chem. Eng.* 93 (2016) 128–142.
- [47] L. F. Marta, Lascio Di, D.J.K.B. Marta, System, A copula-based clustering algorithm to analyse EU country diets, 2017.
- [48] F.M.L. Di Lascio, M. Disegna, A copula-based clustering algorithm to analyse EU country diets, *Knowl.-Based Syst.* 132 (2017) 72–84.
- [49] S.D. Eskesen, P. Tengborg, J. Kampmann, T.H. Veichert, Guidelines for tunnelling risk management: International tunnelling association, Working Group (2), *Tunn. Undergr. Space Technol.* 19 (2004) 217–237.
- [50] S.R. Kasa, S. Bhattacharya, V. Rajan, Gaussian mixture copulas for high-dimensional clustering and dependency-based subtyping, *Bioinformatics* 36 (2020) 621–628.
- [51] C. Chen, L.M. Zhang, R.L.K. Tiong, A novel learning cloud Bayesian network for risk measurement, *Appl. Soft Comput.* 87 (2020).
- [52] Y. Liu, X.J. Wang, et al., Enhancing public building energy efficiency using the response surface method: An optimal design approach, *Environ. Impact Assess. Rev.* 87 (2021) 106548.
- [53] J.F. Unger, C. Konke, An inverse parameter identification procedure assessing the quality of the estimates using Bayesian neural networks, *Appl. Soft Comput.* 11 (2011) 3357–3367.
- [54] T. Parhizkar, F. Aramoun, S. Esbati, Y. Saboohi, Efficient performance monitoring of building central heating system using Bayesian network method, 26, 2019, 100835.
- [55] A. Bauer, C. Czado, Pair-copula Bayesian networks, 2012.
- [56] A. Bauer, C. Czado, T. Klein, Pair-copula constructions for non-Gaussian DAG models, *Canad. J. Statist.-Revue Canadienne de Statistique* 40 (2012) 86–109.
- [57] Y. Liu, H.Y. Chen, et al., Energy consumption prediction and diagnosis of public buildings based on support vector machine learning: A case study in China, *J. Clean. Prod.* 272 (2020) 122542.
- [58] L.C. Yang, Y. Liang, et al., COVID-19 effects on property markets: The pandemic decreases the implicit price of metro accessibility, *Tunn. Undergr. Space Technol.* 125 (2022) 104528.
- [59] J.A. Kumar, S. Abirami, Aspect-based opinion ranking framework for product reviews using a Spearman's rank correlation coefficient method, *Inform. Sci.* 460 (2018) 23–41.
- [60] Y.K. Zhang, Y.L. Li, J.Z. Song, X.L. Chen, Y. Lu, W.K. Wang, Pearson correlation coefficient of current derivatives based pilot protection scheme for long-distance LCC-HVDC transmission lines, *Int. J. Electr. Power Energy Syst.* 116 (2020).
- [61] A. Yazdani-Chamzini, Proposing a new methodology based on fuzzy logic for tunnelling risk assessment, *J. Civ. Eng. Manag.* 20 (2014) 82–94.
- [62] Z.S. Chen, X. Zhang, et al., Expertise-based bid evaluation for construction-contractor selection with generalized comparative linguistic ELECTRE III, *Automat. Constr.* 125 (2021) 103578.
- [63] G. Andreotti, C.G. Lai, Use of fragility curves to assess the seismic vulnerability in the risk analysis of mountain tunnels, *Tunn. Undergr. Space Technol.* 91 (2019).
- [64] C.R. Nott, S.M. Olcmen, D.R. Lewis, K. Williams, Supersonic, variable-throat, blow-down wind tunnel control using genetic algorithms, neural networks, and gain scheduled PID, *Appl. Intell.* 29 (2008) 79–89.
- [65] Z.S. Chen, X. Zhang, et al., Expertise-structure and risk-appetite-integrated two-tiered collective opinion generation framework for large scale group decision making, *IEEE. T. Fuzzy. Syst.* (2022) 3179594, <http://dx.doi.org/10.1109/TFUZZ.2022.3179594>.
- [66] L. Cui, Q. Sheng, Y.K. Dong, M.X. Xie, J.H. Huang, Semianalytical analysis of overexcavation and critical support pressure for support design in TBM tunneling through squeezing rock condition, *Int. J. Geomech.* 21 (2021).
- [67] M. Zhu, S. Liu, J.J.A.I. Jiang, A hybrid method for learning multi-dimensional Bayesian network classifiers based on an optimization model, 44, 2016, pp. 123–148.