Abstract
This study presents a fuzzy Bayesian network (FBN) method to analyze the influence on the safety risk of railway passenger transport applying different risk control strategies. Based on the fuzzy probability of the basic event determined by the expert group decision method, the proposed FBN method can reasonably predict the probability of railway passenger safety risk. It is also proven that control the risk in the safety management of railway passenger transport will be the most effective way to reduce the risk probability of the railway passenger transport safety.

Keywords
railway passenger transport, fuzzy Bayesian network, probabilistic forecast modeling, safety risk analysis, fuzzy probability reasoning

1 Introduction
In the last several decades, the Chinese railway passenger transportation with the rapid development of high-speed train technology has entered the golden age. Unfortunately, the incidence of railway safety accidents also significantly increased with the emergence of this phenomenon. According to the statistics published by the National Railway Administration of China, the average accident rate has reached 1.47% and caused a direct economic loss of 62.92 million yuan annually. Though many safety precautions have been attempted to prevent railway accidents, these current measures are difficult to determine which the main risk factor affecting railway passenger safety is. In this regard, it is necessary to study a new method that considers the uncertainty relationship of different risk factors to provide decision support for ensuring the safety of railway passenger transportation.

In fact, many risk-based analysis methods for avoiding casualties and property losses have been developed, including pre-hazard analysis (PHA) (Khakzad et al., 2013; Savage, 2005), fault tree analysis (FTA) (Fink et al., 2014; Johnston, 2000), safety check list (Muttram, 2002), risk matrix analysis and risk probability analysis (Khakzad et al., 2011; An et al., 2011). In particular, The Bayesian network modeling method based on probability statistics is a powerful tool for dealing with the uncertainty and causality of the multiple influencing factors (Holický et al., 2013; Trucco et al., 2008). It can also be used to effectively resolve the correlation between security risk factors at all levels of the network and to provide an early warning to prevent railway accidents (Lee et al., 2008; Helge and Verner, 2007; Paul and Maiti, 2007). In addition to the safety evaluation, BNs has now been widely used in various fields (Anacleto et al., 2013; Eleye-Datubo et al., 2008; Leu and Chang, 2013).

It is difficult to calculate the prior probability of root nodes by the traditional BN method (Liu et al., 2015; Ayha and Ismail, 2011). The group decision-making technique based upon fuzzy set theory (FST) is generally used to obtain the prior probability of root nodes more accurately (Horčík, 2008). From the perspective of government regulation, this research has newly proposed an approach to predict the risk probabilities of the railway passenger transport safety in different management scenarios.
2 Data Survey

The research field of this study is mainly focused on the safety risk in the process of railway passenger transportation in China. According to the Yearbook of China Railway published by the China Railway Corporation, the railway passenger transportation system in China has a high incidence of major railway accidents. Moreover, the statistics shown in Fig. 1 and Fig. 2 illustrate the types and causes of major railway accidents. As we can see from Fig. 1, the train collision and derailment accounted for 68.89% of the types of railway passenger accidents, while train fires and explosions each accounted for 13.33% and 17.78%. It is obviously recognized that the collision and derailment is the main type of railway passenger accidents.

Also national statistics shown in Fig. 2, attribute more than fifty percentages of the accidents at rail to human and management errors, while only 13.33% to equipment failures and 15.56% to environment factors. It is obviously recognized that the human and management elements play the major role in most railway passenger transportation accidents.

3 Methods and process

According to the analysis of existing statistical data of the railway accidents, it is found that the risk factors (such as personnel factors, management factors) which affect the safety of railway passenger transport are independent from each other, and the probability of occurrence of these risk factors is highly uncertain. FST provides an analytical technique to deal with the inaccuracy of the previous failure probability (Halliwell and Shen, 2009). Adopted the advantages of the BN and FST, a risk analysis method based on fuzzy Bayesian network is proposed in this study.

3.1 BN model construction

In the practice of railway passenger transport, the occurrence probability of the basic risk factors is difficult to be accurately calculated (i.e. the data is rare). However, Khakzad et al. (2011) summarize the simplified procedure which consists of graphical and numerical mapping FT into BN is shown in Fig. 3.

![Fig. 3 Transformation flow chart from FT to BN.](image)

In this study, the risk factors (i.e. individual factors, management factors, equipment factors and environmental factors) were defined as intermediate nodes of the BN. Subsequently, the BN model of RPTS is newly developed, as shown in Fig. 4. In addition, the descriptions of all nodes in the BN of RPTS are illustrated in Table 1.

![Fig. 4 Established Bayesian network model for RPTS.](image)

3.2 Fuzzy probability assessment

In view of the method described in previous section, it is reasonable to carry out the safety risk analysis in various railway safety control strategy applying the built BN model. Nevertheless, the prior information based on expertise knowledge (i.e., the node prior probability) needs to be provided.
into the Bayesian network model. Unfortunately, according to previous related research (Ozbay and Noyan, 2006; Liu et al., 2014; Zhang et al., 2014), it is hard to accurately evaluate the prior probability simply based on expertise experiential knowledge. In order to obtain more precise prior probabilities of the root nodes, Zhang et al. (2014) have proposed an expertise-based fuzzy probability calculation method which is used in this study to estimate the probabilities of the nodes applying triangle fuzzy numerals.

Table 1 Descriptions of nodes in the BNs of RPTS

| Nodes | Descriptions                      |
|-------|-----------------------------------|
| T     | Railway passenger transportation safety |
| A₁    | Personnel risk                    |
| A₂    | Management risk                   |
| A₃    | Equipment risk                    |
| A₄    | Environmental risk                |
| X₁    | Personnel Composition             |
| X₂    | Staff training                    |
| X₃    | Staff assessment                  |
| X₄    | Safety production objectives      |
| X₅    | Safety production system          |
| X₆    | Safety production process         |
| X₇    | Safety education and publicity     |
| X₈    | Emergency rescue system           |
| X₉    | Equipment failure rate            |
| X₁₀   | Maintenance protection            |
| X₁₁   | Natural environment               |
| X₁₂   | Social environment                |

The process of this method contains four sequential steps of probability interval division, expert credibility calculation, probability fuzzification and probability defuzzification. The specific steps of the fuzzy probability calculation method in each of these steps are presented below.

Step 1 Probability interval division
A reasonable interval division is helpful to improve the accuracy of expert investigation. Therefore, this research introduces the Wickens’ 7-level theory (Wickens and Hollands 2000) to express the probability range of natural language division. And the probability of the event is divided into very high (VH), high (H), the higher (FH), medium (M), the lower (FL), low (L), very low (VL) 7 intervals, as seen in Fig. 5., and the corresponding form of fuzzy numbers level is shown in Table 2.

Step 2 Expert credibility calculations
In this step, the expert reliability index which taking educational background, working years and title level into consideration is presented to improving the rationality of expert knowledge. The expert reliability index, denoted by ω in this study, is calculated by Eq. (1).

$$\omega = \sum_{i=1}^{3} w_i q_i$$

Where $q_i$ refers to the credibility of the $i$th expert level as shown in Table 3; $w_i$ refers to weight values of the three influencing factors. Here the weight value of the professional qualifications is 0.5, the weight value of the years of work experience is 0.4, and the weight value of the education background is 0.1.

Table 2 Descriptions of fuzzy probability interval

| Intervals(k) | Fuzzy expressions | Fuzzy numbers |
|--------------|-------------------|---------------|
| 1            | very low          | a₁, c₁, b₁    |
| 2            | low               | 0.10, 0.20, 0.30 |
| 3            | the lower         | 0.20, 0.35, 0.50 |
| 4            | medium            | 0.40, 0.50, 0.60 |
| 5            | the higher        | 0.50, 0.65, 0.80 |
| 6            | high              | 0.70, 0.80, 0.90 |
| 7            | very high         | 0.80, 0.90, 1.00 |

Table 3 The level of expert credibility

| level | Years of work experience | Professional qualifications | Education background | q   |
|-------|--------------------------|-----------------------------|----------------------|-----|
| I     | more than 30 years       | Senior engineers            | Graduate degree or above | 1.0 |
| II    | 20-30 years              | Engineers                   | bachelor degree      | 0.9 |
| III   | 10-20 years              | Assistant engineers         | College degree       | 0.8 |
| IV    | 5-10 years               | Skilled worker              | Secondary education  | 0.7 |
| V     | 1-5 years                | Ordinary worker             | The following secondary education | 0.6 |
Step 3 Probability fuzzification

The distribution of the probability which one expert considers a basic event lying in the \( i \)th fuzzy probability interval is presented, as seen in Eq. (2). Continually, According to the “3σ criterion”, the characteristic values is calculated by the Eq. (3) - (6).

\[
p_i = \frac{1}{\sum_{j=1}^{\omega} \left( a_j - a_{j+1} \right)} \times \frac{1 - \omega}{2}, 1 \leq i \leq k - 1
\]

\[
m = E(P) = \sum_{i=1}^{7} (c_i \times P_i).
\]

\[
s = \sqrt{D(P)} = \left( \sum_{i=1}^{7} (c_i - E(P))^2 \right) \times P_i.
\]

\[
a = m - 3\sigma, b = m + 3\sigma.
\]

Where \( P_i \) refers to the probability which one expert considers a basic event lying in the \( i \)th fuzzy probability interval; \( \omega \) refers to the expert confidence indicator, \( a_i \), \( a_j \), \( a_{j+1} \) and \( a_{j+k} \) refer to the values of lower boundary as shown in Table 1; \( i, j, k \) and \( k \) refer to the level of fuzzy probability interval.

\[
P = \sum_{i=1}^{7} P_i / n.
\]

\[
m = \frac{1}{n} \sum_{i=1}^{7} (c_i \times P_i).
\]

\[
s = \sqrt{\frac{1}{n} \sum_{i=1}^{7} (c_i - \frac{1}{n} \sum_{i=1}^{7} c_i)^2} \times \frac{1}{n} \sum_{i=1}^{7} P_i.
\]

\[
a = m - 3s, b = m + 3s.
\]

Table 4 The fuzzy probabilities of root nodes in FBN

| Root nodes | The fuzzy probabilities \( \alpha, \beta, \gamma \) | Val(F) |
|------------|------------------------------------------------|--------|
| X1         | 0.033, 0.045, 0.057 | 0.045 |
| X2         | 0.055, 0.068, 0.081 | 0.068 |
| X3         | 0.019, 0.031, 0.043 | 0.031 |
| X4         | 0.021, 0.037, 0.053 | 0.037 |
| X5         | 0.043, 0.061, 0.079 | 0.061 |
| X6         | 0.063, 0.084, 0.105 | 0.084 |
| X7         | 0.012, 0.018, 0.024 | 0.018 |
| X8         | 0.037, 0.048, 0.059 | 0.048 |
| X9         | 0.072, 0.095, 0.118 | 0.095 |
| X10        | 0.045, 0.059, 0.073 | 0.059 |
| X11        | 0.003, 0.007, 0.011 | 0.007 |
| X12        | 0.028, 0.037, 0.046 | 0.037 |

4 Results and Discussion

4.1 Model Validation

Before making railway passenger safety risk analysis, firstly, it needs to verify the validity of the model by comparing the occurrence frequency of the accident statistics with the conditional probability calculated by the fuzzy Bayesian model of railway passenger safety risk. The reliability of the model is considered to be accepted when the relative error is within a certain range.

The comparison of the analytical results (i.e. probabilities of the four basic risk factors) of the BN model to the safety investigation statistics records (i.e. accident statistics shown in Fig. 2) is shown in Table 5.

From the table, it can be seen that the relative error rate of the posterior probability and the actual statistical probability for the four basic risk factors are 12.11%, 13.40%, 16.20% and 7.99% respectively, and the average error is 12.42%. Because both the relative error and average error are less than twenty percentages, it can be considered that the FBN-based safety
risk analysis method had certain accuracy and practicability for railway passenger transport safety risk assessment. It is worth noting that only major railway accident statistics records are using to the validation due to the lack of the available data. Moreover, it is foreseeable the error rate of the model will be further reduced as long as there is enough railway accident statistics, thereafter, the FBN-based safety risk analysis method can be more precisely.

Table 5 Comparison between BN risk assessments and real statistics records

| Nodes | Descriptions | The proportion of the posterior probabilities | Actual statistics | Relative error rate |
|-------|--------------|---------------------------------------------|-------------------|--------------------|
| A₁    | Personnel risk | 29.90%                                      | 26.67%            | 12.11%             |
| A₂    | Management risk | 32.76%                                      | 28.89%            | 13.40%             |
| A₃    | Equipment risk  | 15.49%                                      | 13.33%            | 16.20%             |
| A₄    | Environmental risk | 16.80%                                      | 15.56%            | 7.99%              |
| Average error rate |                |                                             |                   | 12.42%             |

4.2 Safety Risk Analysis

In the premise of meeting the accuracy requirement, the developed method can be used for safety risk analysis in RPTS during the stage of security risk prevention.

4.2.1 Scenario Analysis

In fact, the railway passenger transport safety risk in the risk prevention phase is related to the basic influencing factors, namely Personnel risk, Management risk, Equipment risk and Environmental risk. The purpose of the scenario analysis is to calculate the conditional probability of the top event (i.e. the railway passenger transportation safety risk). Thus, it is successful to obtain the results (i.e. posterior conditional probability) which are used for comparison with entering the state variable parameters of each scenario (yes or no) into the BN as evidence. Obviously, it is the best result without any basic risk in RPTS (i.e. scenario A) from the perspective of the administrators, according to the inference probabilities in Table 6. Furthermore, the highest posterior probability is obtained when the railway passenger transport management risk occurs (i.e. scenario C), compared to the personnel risk (i.e. scenario B), the equipment risk (i.e. scenario D) and the environmental risk (i.e. scenario E). The railway passenger transport managers with safety responsibilities can optimize continuously the risk prevention plans by this way until the high risk factors are under control.

4.2.2 Sensitivity analysis

Sensitivity analysis is to identify which factors most affect railway passenger transport accident, and therefore according which the managers of railway enterprise could take appropriate measures to effectively improve the operating performance of railway passenger transport. Adopted by the merit of applying the sensitivity analysis technique with entering the fuzzy probability of the root nodes which have been calculated, three key performance indicators (KPIs) by using Eq. (11)-(13) are shown in Fig. 6.

\[
I^{REF}(X_i) = \frac{\text{Max} \left\{ P(T=1 | X_i = x_i) \right\} - P(T=1)}{P(T=1)}, i = 1, 2, \ldots, n. \tag{11}
\]

\[
I^{REV}(X_i) = \frac{P(T=1) - \text{Min} \left\{ P(T=1 | X_i = x_i) \right\}}{P(T=1)}, i = 1, 2, \ldots, n. \tag{12}
\]

\[
I^{AVG}(X_i) = \frac{1}{2} [I^{REF}(X_i) + I^{REV}(X_i)], i = 1, 2, \ldots, n. \tag{13}
\]

Where \(X_i\) refers to the \(i\)th risk factor; \(I^{REF}(X_i)\) refers to risk expansion sensitivity of \(X_i\); \(I^{REV}(X_i)\) refers to risk reduction sensitivity of \(X_i\); \(I^{AVG}(X_i)\) refers to risk average sensitivity of \(X_i\); \(i\) and \(n\) refer to the number of risk factors.

Table 6 Probabilistic results of the scenario analysis of railway passenger safety risk (T)

| Scenario | Descriptions | P(T=1) |
|----------|--------------|--------|
| A: A₁=no, A₂=no, A₃=no, A₄=no | No basic risk | 1.82% |
| B: A₁=yes, A₂=no, A₃=no, A₄=no | Only personnel risk | 8.88% |
| C: A₁=no, A₂=yes, A₃=no, A₄=no | Only management risk | 9.42% |
| D: A₁=no, A₂=no, A₃=yes, A₄=no | Only equipment risk | 6.15% |
| E: A₁=no, A₂=no, A₃=no, A₄=yes | Only environmental risk | 2.24% |

As can be seen from the Fig. 6, the ranking results of these three KPIs enjoy high consistency. Moreover, \(X_1\) (i.e. personnel composition), \(X_2\) (i.e. safety production process) and \(X_3\) (i.e. emergency rescue system) are the most important factors affecting the occurrence of the top event (i.e. railway passenger transport safety) due to the ranking results calculating the KPIs. Therefore, corresponding measures should be taken to ensure that the three key factors are in a reasonable state at the risk control stage.
5 Conclusions
This study developed an effective FBN-based safety risk analysis method for railway passenger transport. The inference results were validated against the four basic risk factors which derived from the statistical data of major railway passenger accidents over the years. It is showed that the posterior probability of the basic risk factors calculated by BN inference were within a certain error range compared with the actual railway passenger transport accident rates. It is suggested that the proposed fuzzy-Bayesian-network-based safety risk analysis method have a certain accuracy and practicality in railway passenger transport safety risk analysis. Accordingly, on the basis of the scenario simulation and sensitivity analysis, railway passenger transport governors can take preventive and protective measures in advance to effectively control the safety risks of railway passenger transportation.

Although the fuzzy probability assessment method reduces the uncertainty of expert information to some extent, a more rational structure of BN can be explored and established with abundant reliable safety data in the future research. Moreover, the other approach for expert elicitation is worth studying in the future researches. Finally, the situations applied in this study to assessment the safety risk of railway passenger transport still need more detail parameters to define more comprehensive condition in the future.

Acknowledgement
The project presented in this article is supported by the National Natural Science Foundation of China (71571011) and the Fundamental Research Funds for the Central Universities (2017YJS097).

References
An, M., Chen, Y., Baker, C. J. (2011). A fuzzy reasoning and fuzzy-analytical hierarchy process based approach to the process of railway risk information: a railway risk management system. Information Sciences. 181(18), pp. 3946-3966. https://doi.org/10.1016/j.ins.2011.04.051
Aneacelo, O., Queen, C., Albors, C. J. (2013). Multivariate forecasting of road traffic flows in the presence of heteroscedasticity and measurement errors. Journal of the Royal Statistical Society. 62(2), pp. 251-270. https://doi.org/10.1111/j.1467-9876.2012.01059.x
Braae, M., Rutherford, D. A. (1978). Fuzzy relations in a control setting. Kybernetes. 7(3), pp. 185-188. https://doi.org/10.1108/eb005482
Castillo, E., Menéndez, J. M., Sánchez-Cambronero, S. (2008). Predicting traffic flow using bayesian networks. Transportation Research Part B Methodological. 42(5), pp. 482-509. https://doi.org/10.1016/j.trb.2007.10.003
Chen, T. T., Lee, S. S. (2014). Fall risk assessment of cantilever bridge projects using bayesian network. Safety Science. 70(70), pp. 161-171. https://doi.org/10.1016/j.ssci.2014.05.011
Detynecki, M., Yager, R. R. (2000). Ranking fuzzy numbers using α-weighted valuations. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. 8(05), pp. 573-591. https://doi.org/10.1142/S021848850000040X
Eleyedatubo, A. G., Wall, A., Wang, J. (2008) Marine and offshore safety assessment by incorportative risk modelling in a fuzzy Bayesian network of an induced mass assignment paradigm. Risk Analysis. 28(1), pp. 95–112. https://doi.org/10.1111/j.1539-6924.2008.01004.x
Fink, O., Zio, E., Weidmann, U. (2014). Predicting component reliability and level of degradation with complex-valued neural networks. Reliability Engineering & System Safety. 121(1), pp. 198-206. https://doi.org/10.1016/j.ress.2013.08.004
Halliwell, J., Shen, Q. (2009). Linguistic probabilities: theory and application. Soft Computing. 13(2), pp. 169-183. https://doi.org/10.1007/s00500-008-0304-1
Helge, L., Verner, J. F. (2007). Bayesian networks in reliability. Reliability Engineering & System Safety. 92(1), pp. 92-108. https://doi.org/10.1016/j.ress.2005.11.037
Holický, M., Marková, J., Sykora, M. (2013). Forensic assessment of a bridge downfall using bayesian networks. Engineering Failure Analysis. 30(2), pp. 1-9. https://doi.org/10.1016/j.engfailanal.2012.12.014
Horčík, R. (2008). Solution of a system of linear equations with fuzzy numbers. Fuzzy Sets & Systems. 159(14), pp. 1788-1810. https://doi.org/10.1016/j.fss.2008.01.018
Johnston, G. (2000). Reliability for technology, engineering, and management, by paul kales. Techinometrics. 42(2), pp. 207-207. https://doi.org/10.2307/1271456
Khakzad, N., Khan, F., Amyotte, P. (2011). Safety analysis in process facilities: comparison of fault tree and bayesian network approaches. Reliability Engineering & System Safety. 96(8), pp. 925-932. https://doi.org/10.1016/j.ress.2011.03.012
Khakzad, N., Khan, F., Paltrinieri, N. (2014). On the application of near accident data to risk analysis of major accidents. Reliability Engineering & System Safety. 126(5), pp. 116-125. https://doi.org/10.1016/j.ress.2014.01.015
Lee, S. J., Man, C. K., Seong, P. H. (2008). An analytical approach to quantitative effect estimation of operation advisory system based on human cognitive process using the bayesian belief network. Reliability Engineering & System Safety. 93(4), pp. 567-577. https://doi.org/10.1016/j.ress.2007.02.004
Leu, S. S., Chang, C. M. (2013). Bayesian-network-based safety risk assessment for steel construction projects. Accident Analysis & Prevention. 54(2), pp. 122-133. https://doi.org/10.1016/j.aap.2013.02.019
Liu, P., Yang, L., Gao, Z., Li, S., Gao, Y. (2015). Fault tree analysis combined with quantitative analysis for high-speed railway accidents. Safety Science. 79, pp. 344-357. https://doi.org/10.1016/j.ssci.2015.06.017
Mamdani, E. H. (2010). Application of fuzzy algorithms for control of simple dynamic plant. Proceedings of the Institution of Electrical Engineers. 121(121), pp. 1585-1588. https://doi.org/10.1049/piee.1974.0328
Mentes, A., Helvaciglu, I. H. (2011) An application of fuzzy fault tree analysis for spread mooring systems. Ocean Engineering. 38(2-3), pp. 285-294. https://doi.org/10.1016/j.oceaneng.2010.11.003
Muttram, R. I. (2002). Railway safety’s safety risk model. Proceedings of the Institution of Mechanical Engineers Part F Journal of Rail & Rapid Transit. 216(2), pp. 71-79. https://doi.org/10.1243/09544090260082317
Ozbay, K., Noyan, N. (2006). Estimation of incident clearance times using bayesian networks approach. *Accident Analysis and Prevention*. 38(3), pp. 542-55.  
https://doi.org/10.1016/j.aap.2005.11.012

Paul, P. S., Maiti, J. (2007). The role of behavioral factors on safety management in underground mines. *Safety Science*. 45(4), pp. 449-471.  
https://doi.org/10.1016/j.ssci.2006.07.006

Savage, S. L. (2005). Risk and uncertainty in dam safety. Thomas Telfrod PTR, London.  
https://doi.org/10.1680/rauids.2005.32705.0007

Trucco, P., Cagno, E., Ruggeri, F., Grande, O. (2008). A bayesian belief network modelling of organisational factors in risk analysis: a case study in maritime transportation. *Reliability Engineering & System Safety*. 93(6), pp. 845-856.  
https://doi.org/10.1016/j.ress.2007.03.035

Wickens, C. D., Hollands, J. G. (2000). *Engineering psychology and human performance*. 3rd edition. Prentice Hall PTR, New Jersey. URL: http://www.citeulike.org/group/398/article/225322