WILEY | Hindawi

Research Article

Analysis of Factors Affecting Road Transport Accidents of Hazardous Materials Based on PG-BN

Zewen Li[1,2] **and Mengmeng Zhang** [1,2]

[1]*Transportation and Logistics Engineering College, Shandong Jiaotong University, Jinan 250357, China*
[2]*Shandong Key Laboratory of Smart Transportation (Preparation), Jinan 250357, China*

Correspondence should be addressed to Mengmeng Zhang; zhangmengmeng@sdjtu.edu.cn

To analyze the factors affecting road accidents involving hazardous materials, the Bayesian network (BN) model was used to fit the accident data. However, considering the possible overfitting phenomenon of the BN model, the model was optimised by combining Pearson's chi-squared test and Granger causality test (PG) methods. First, the data of hazardous materials accidents were preprocessed, and the index system of factors affecting hazardous materials road transport was constructed from five dimensions of "people, vehicles, hazmat, roads, and environment"; second, Pearson's chi-squared test and the Granger causality test were used to screen the factors affecting hazardous materials road transport accidents and to determine the causal relationship between the factors; finally, the BN model was constructed with accident severity and accident processing time as target nodes, and the results were analyzed and validated. The results show that the overall relative error rate of the model is less than 10% and can be used to explore the risk factors of hazardous materials transport accidents; weather, visibility, lighting, intersection type, road condition, road type, driver condition, vehicle type, etc. are all important factors affecting the severity of hazardous materials transport accidents. The study can serve as a reference for the safety supervision and management of hazardous materials transport enterprises and industrial management departments.

1. Introduction

Since the 21st century, with the acceleration of economic development, the volume of hazardous materials transport in China has shown a trend of continuous growth. Once an accident occurs during the transport of hazardous materials, it will cause great losses to society and the environment [1]. Therefore, it is necessary to analyze the factors affecting hazardous materials road transport accidents to help develop effective risk-prevention measures.

Many theories and methods have been developed for the analysis of factors affecting hazardous materials road transport accidents. When studying the mechanism of hazardous materials road transport accident characteristics, it is necessary to recognize and analyze from the perspective of uncertainty. In order to better understand and predict

hazardous materials road transport accidents, Shen et al. [2, 3] chose to model the factors affecting hazardous materials transport accidents using XGBoost and C5.0 machine learning algorithms by comparing statistical models and machine learning models, and the results showed that the models had better performance. Ma et al. [4] and Xing et al. [5] proposed an ordered logit regression model to account for unobserved heterogeneity among the factors of hazardous materials transport accidents. But ordered Logit models use probability values as the values of the dependent variable, which may increase computational complexity. In contrast, Bayesian network models deal directly with probability values and are able to more accurately describe the uncertainty of random variables. However, statistical models such as the ordered Logit model add computational complexity and algorithms such as XGBoost are less

interpretable, in contrast to BN, which provides more intuitive explanations and visualizations and more accurately describes the uncertainty of random variables.

BN is a probabilistic inference, network model that can represent probabilistic causality, which was proposed by Pearl in 1985 [6] and can be used to express and reason about uncertain problems. Applying this model to the mechanistic characterization of hazardous materials transport accidents, it can effectively identify factors affecting road transport accidents of hazardous materials from the intuitive graphical structure and identify the connection between each affecting factor of hazardous materials transport accidents. Zhao et al. [7] constructed a model of factors affecting hazardous materials transport accidents based on the Bayesian network, taking the hazardous materials transport accidents in China as a case study; Ma et al. [8] used the Genie tool to construct a Bayesian network model of hazardous materials accidents to reveal the causality of accident occurrence; and Li et al. [9] proposed a fuzzy Bayesian network for identifying real-time risk analysis of road tanker traffic. However, the BN structure is more complicated to establish, and the direct use for the analysis of factors affecting hazardous materials road transport accidents can easily confuse the causal relationship between nodes. At present, most scholars combine other methods to preprocess the data before modeling to effectively improve model accuracy. Hashemi et al. [10] proposed a multivariate security analysis model based on copula BN, combined with the copula function to consider the correlation between variables, and illustrated the superior performance of the CBN model compared to the traditional BN model through a case; Huang et al. [11] combined the explanatory structural model (ISM) and the BN model to construct a BN model for predicting hazardous materials road transport accidents after identifying causal relationships between variables; Sun et al. [12] used random forests to rank the importance of risk factors for hazardous materials transport and then developed a BN to provide probabilistic inference, and the results showed that the proposed method was very efficient; Shen et al. [13] used fault tree analysis to first find the direct and indirect causes of hazardous transport accidents and constructed a BN model with strong descriptive power; Ding et al. [14] combined the credal network and IDM methods to construct a model for analyzing the causes of hazardous materials road transport accidents based on a credal network; Pan et al. [15] proposed a risk assessment method based on an improved FBN model, which provides an effective tool for risk management of hazardous material transportation enterprises; Wang et al. [16] first used the grounded theory (GT) to identify influencing factors and then developed a BN-based model for these collected hazardous materials transport accidents; and Cheng et al. [17] used dynamic BN to calculate the likelihood of hazardous chemical spills and explosions. In conclusion, it can be seen that before constructing the BN model, it is important to effectively deal with the relationship among the factors affecting hazardous materials road transport accidents to improve the accuracy and robustness of the model.

Previous studies have mostly used association rule mining [18], structural equation modeling, or the Pearson correlation coefficient method to test the correlation between variables, but the Pearson correlation coefficient method usually requires converting variables into dummy variables when dealing with discrete variables, and the more the variable classifications, the more the dummy variables need to be set up. Pearson's chi-squared test is more suitable for discrete variable correlation analysis than using the Pearson correlation coefficient method, which can reduce the complexity caused by dummy variable settings. After the factor correlation analysis, the causal relationship between each factor and the hazardous materials transport accidents needs to be discerned. Generally, the expert scoring method is more commonly used, but the method is more workload and subjective, which has certain limitations and tends to make the model less reliable. To reduce the workload, Wang [19] used Bayesian truth serum to fuse with expert opinions; although the method requires only a few opinions to conclude, it still has certain subjectivity. The Granger causality test method assesses the causal relationship between relevant variables, which can be used to determine the causal relationship between variables by conducting regression analysis based on data from two variables in different periods.

In general, it is proposed to use the Pearson's chi-squared test and Granger causality test (PG) to identify the causal relationship between variables and improve the BN model to construct the index system, which can not only make up for the deficiency that BN itself cannot screen variables but also prevent the BN structure from overfitting and improve the learning efficiency of the network structure. It effectively prevents the BN model from relying too much on the data set, ensures the accuracy and robustness of the model, and can provide theoretical guidance for the risk assessment of hazardous materials road transport.

The remainder of this paper is organized as follows. The second chapter is concerned with the methodology used for this study. The third section presents the findings of the research.

2. Materials and Methods

2.1. Analysis of Factors Affecting Accidents of Hazardous Materials Road Transport. Road accidents of hazardous materials occur as a result of the combined effect of people, vehicles, hazmat, roads, environment, and other factors. According to the theory of transport system safety, hazardous materials are the first type of risk source; people, vehicles, roads, and the environment are the second type of risk source. Once an accident or risk factors lead to an accident, all of them will cause damage to drivers, vehicles, surrounding people, and the environment. Therefore, this paper analyses the five dimensions of hazmat, people, vehicles, roads, and environment based on screening the factors affecting hazardous materials transport accidents, determining the causal relationship between the factors, and studying the mechanism of hazardous materials road transport accident characteristics. Based on one city's

statistics of 448 accidents between 2018 and 2020, it includes information on the accident processing time, the accident severity, the driver status, the driver age, the driver education level, the type of vehicle, whether the vehicle is overloaded, hazmat categories, quantity of hazmat, the type of intersection section, the condition of the pavement, the road line, the type of road, the weather, and visibility and lighting conditions. The model index system is constructed from five dimensions of "people, vehicles, hazmat, roads, and environment." Considering that the nodes of the BN model are discrete variables, continuous variables need to be disaggregated, and the discrete variables such as driver status, weather, and road condition are more classified, which may make the sample observation too discrete and have poor explanatory power as model variables. Therefore, some of the accident information variables are combined and classified according to the descriptive characteristics to form "synthetic variables," and these "synthetic variables" are used for statistical analysis, and the sample data are discrete variables after processing (see Figure 1). Accident severity is divided into minor, general, major, and serious, and the accident processing time is divided into <30, [30, 60), [60, 120), and ≥120 minutes.

2.2. Correlation Analysis Based on Pearson's Chi-Squared Test.
The Pearson's chi-squared test [20, 21] is used to characterize the degree of association between two factors and is calculated as follows:

$$\chi^2 = \sum \frac{(O\text{-}E)^2}{E}, \tag{1}$$

where χ^2 is the Pearson's chi-squared test value, which measures the degree of deviation between the actual value and the theoretical value. When the value is less than 0.05, it indicates that the two factors are correlated; O represents the actual observed frequency and E represents the expected frequency of the category.

The correlation of variables was determined as follows: the driver status was correlated with road and environment, etc.; the vehicle type was correlated with volume and hazmat categories transport, etc.; and with the exception of driver age, education, and road type, the driver status, vehicle type, and hazmat categories were correlated with accident severity and accident handling time, etc. The relevant variables were used as study variables to determine the nodes of the study BN (see Table 1).

2.3. Causal Analysis Based on the Granger Causality Test.
The Granger causality test was proposed by Granger in 1969 [22] and is used to analyze the causal relationship between variables. The Granger causality test is generally defined in statistics as follows: suppose the data set has two variables and b. If predicting the current value using past data is better than predicting the current value without using past data, then it is Granger causal. The regression model is as follows:

$$y'_t = \sum a'_i x'_{t-1} + \sum b'_i y'_{t-1} + c_1,$$
$$x'_t = \sum a''_i x'_{t-1} + \sum b''_i y'_{t-1} + c_2, \tag{2}$$

where y'_t and x'_t are the study variables; c_1 and c_2 are the error coefficients, which are assumed to be uncorrelated; a'_i, b'_i, a''_i, and b''_i are the coefficients to be determined.

Based on the Granger causality test method to calculate the causal relationship between the variables, the final structure of the BN is built (see Figure 2). The connection between each node is directly causal, from cause to effect.

2.4. BN Model Construction.
BN is a directed acyclic graph consisting of nodes and directed edges, where each node corresponds to a random variable, the edges can represent the dependency between variables, and the nodes marked with arrows indicate that they are causal nodes [23, 24] (see Figure 3).

Figure 3 shows a simple four-node BN, whose joint probability distribution can be expressed by multiplying the conditional probability distributions of each node as follows:

$$p(a, b, c, d) = p(d|b, c)p(b|a)p(c|a)p(a), \tag{3}$$

where p (a, b, c, d) can also be written as $p(a \cap b \cap c \cap d)$, which represents the probability of the event a, b, c, d occurring together; a, b, c, d are random variables; nodes b, c are the parents of the node d; node a is the parents of node b, c; p (a) is the marginal probability, which is the probability of the occurrence of the event a, also known as the a priori probability; and $0 < p$ (a) < 1; p ($c \mid a$), p ($b \mid a$), and p ($d \mid b$, c) are the conditional probabilities, taking p ($c \mid a$) as an example, which is the probability of the occurrence of the event c given the occurrence of the event a. The conditional probability, also known as the posterior probability, is calculated as follows:

$$p(c|a) = \frac{p(ac)}{p(a)}. \tag{4}$$

BN parameter learning is the process of obtaining a model through data analysis. In this paper, BN parameter learning is performed using Netica software [25, 26]. The software can perform probabilistic reasoning, cause diagnosis, and outcome prediction according to different needs and select appropriate algorithms according to the integrity of the data. In this paper, 348 hazardous materials road transport accidents data from 2018 to 2020 are randomly selected as samples, and BN parameter learning is carried out on the sample data to establish the model (see Figure 4).

As can be seen from Figure 4, nodes $S1$ and $S2$ are target nodes, which are used to characterize the level of hazardous materials transport accidents; other nodes indicate the relevant factors that lead to $S1$ and $S2$, and the connecting lines between the nodes are direct cause-effect relationships, from cause to effect.
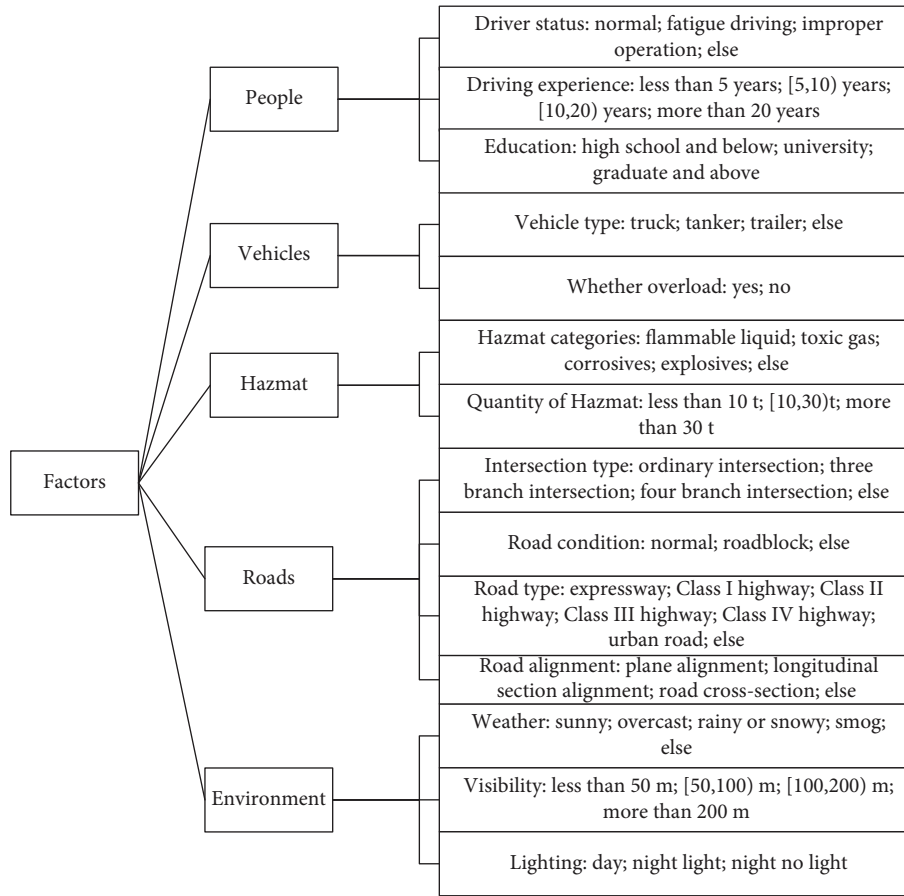
Figure 1: Index system of affecting factors of hazardous materials road transport accidents.

3. Results and Discussion

3.1. Model Validation. In order to validate the effectiveness of the model, additional 100 hazardous materials accidents data from 2018 to 2020 are extracted for analysis in this paper. The validation data are input into the BN model to obtain the probability distribution of the two nodes of accident severity and accident handling time, and the prediction results are compared with the actual results (see Table 2).

As can be seen from Table 2, the overall relative error values of the model in predicting accident severity and accident handling time are only 5% and 8%, respectively, indicating that the model has good applicability.

3.2. Model Results. The BN model can predict the probability of S1 and S2 by the state of each factor. As can be seen from Figure 5, among the road and environmental factors, "sunny," "visibility more than 200," "day," "urban road," "ordinary intersection," "normal," and other node states accounted for the largest proportion, that is, the most normal conditions. To observe the influence of the above-mentioned characteristics, this study set them as "evidence" and can observe that the probability of being "normal" increases from 63.8% to 95.3%; the probability of being "major accident." The probability of being a "major

accident" decreased from 15.1% to 10.2% and the probability of being "more than 120" decreased from 14.8% to 11.9%. This shows that under most normal circumstances, the probability of major accidents in the transport of hazardous materials is low.

The BN model can also analyze the effect of a single factor on the target node [27]. For example, if other factors are held constant and the effect of one factor on the severity of a hazardous materials transport accident is analyzed, the results will be as follows:

(1) The effect of weather and visibility on accident severity

When the weather node is "overcast" and the visibility is "less than 50," the probability of a major accident is 17.6%; and when the weather node is "rainy or snowy" and the visibility is "less than 50," the probability of a major accident is 18.3%. Compared to "sunny" and visibility ">200 m," the probability increases by about 5%. This indicates that the transport of hazardous materials is more likely to lead to major road accidents in bad weather.

(2) The effect of lighting conditions on accident severity

When the lighting condition node is set to "night no light," the probability of a major accident is 15.0%. The probability of a major accident increases by

TABLE 1: Study variables.

| Node	Symbols	Node status
Driver status	R1	Normal Improper_operation Fatigue_driving Else
Vehicle type	C1	Truck Tank Trailer Else
Whether overload	C2	Yes No
Hazmat categories	W1	Toxic_gas Corrosives Flammable_liquid Explosives Else
Quantity of hazmat (t)	W2	Less_than_10 Between_10_and_30 More_than_30
Intersection type	D1	Ordinary_intersection Three_branch_intersection Four_branch_intersection Else
Road condition	D2	Normal Roadblock Else
Road type	D3	Expressway Class_I_Highway Class_II_Highway Class_III_Highway Class_IV_Highway Urban_road Else
Weather	H1	Sunny Overcast Rainy_or_snowy Smog Else
Visibility (m)	H2	Less_than_50 Between_50_and_100 Between_100_and_200 More_than_200
Lighting	H3	Day Night_light Night_no_light
Accident severity	S1	Minor_accident General_accident Serious_accident Major_accident
Accident processing time (min)	S2	Less_than_30 Between_30_and_60 Between_60_and_120 More_than_120
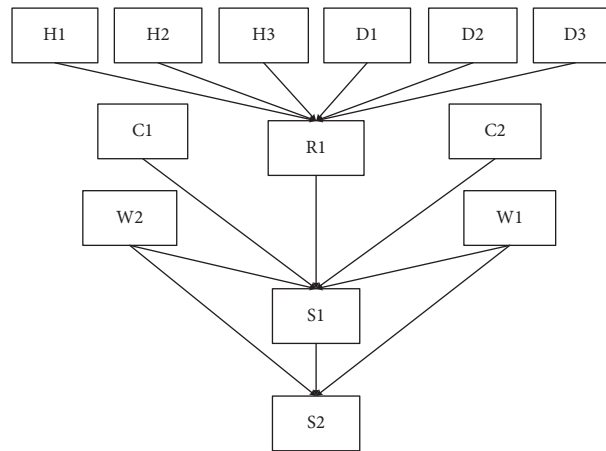
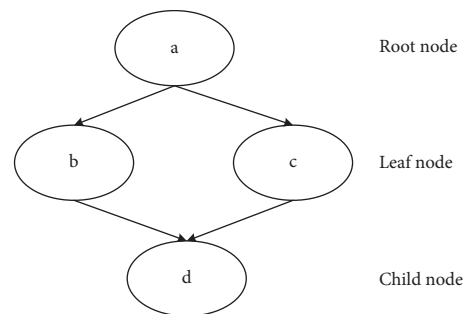FIGURE 2: BN structure among factors of hazardous materials transport accidents.
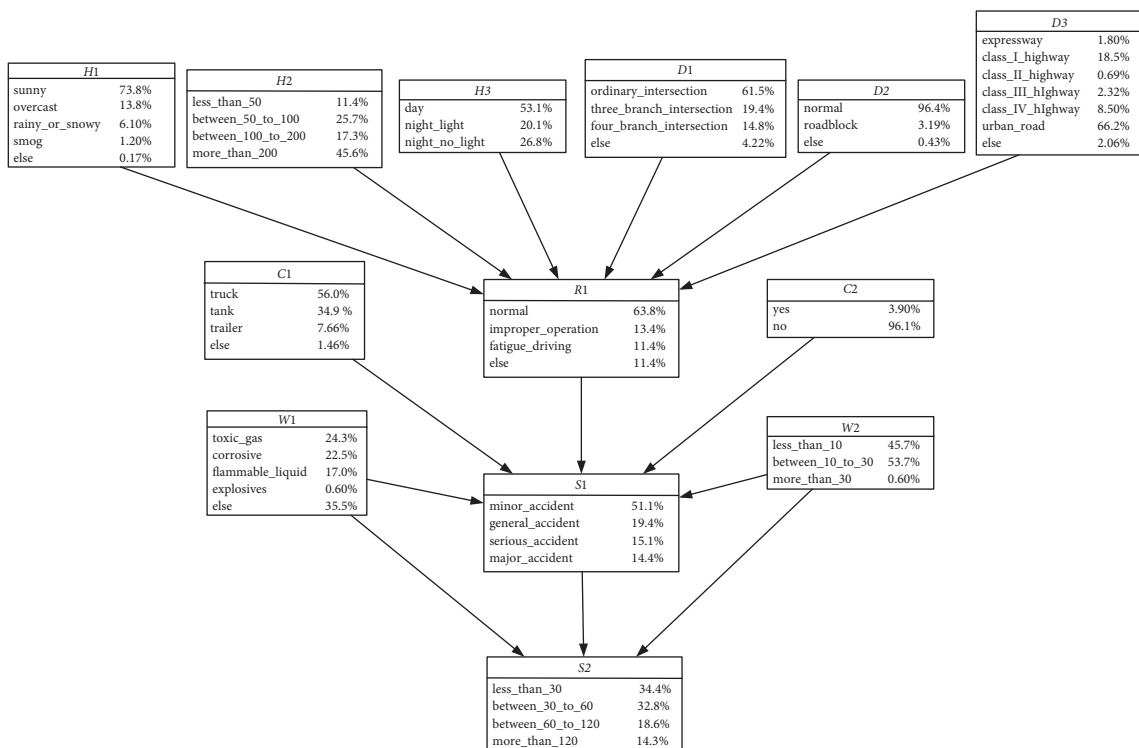

FIGURE 3: BN sketch.

FIGURE 4: Hazardous materials transport accidents BN model.

TABLE 2: Model accuracy.

True	Predict S1			
	Minor_accident	General_accident	Serious_accident	Major_accident
Minor_accident	43	0	1	0
General_accident	1	29	1	0
Serious_accident	0	1	17	0
Major_accident	0	0	1	6
Overall relative error		5%		

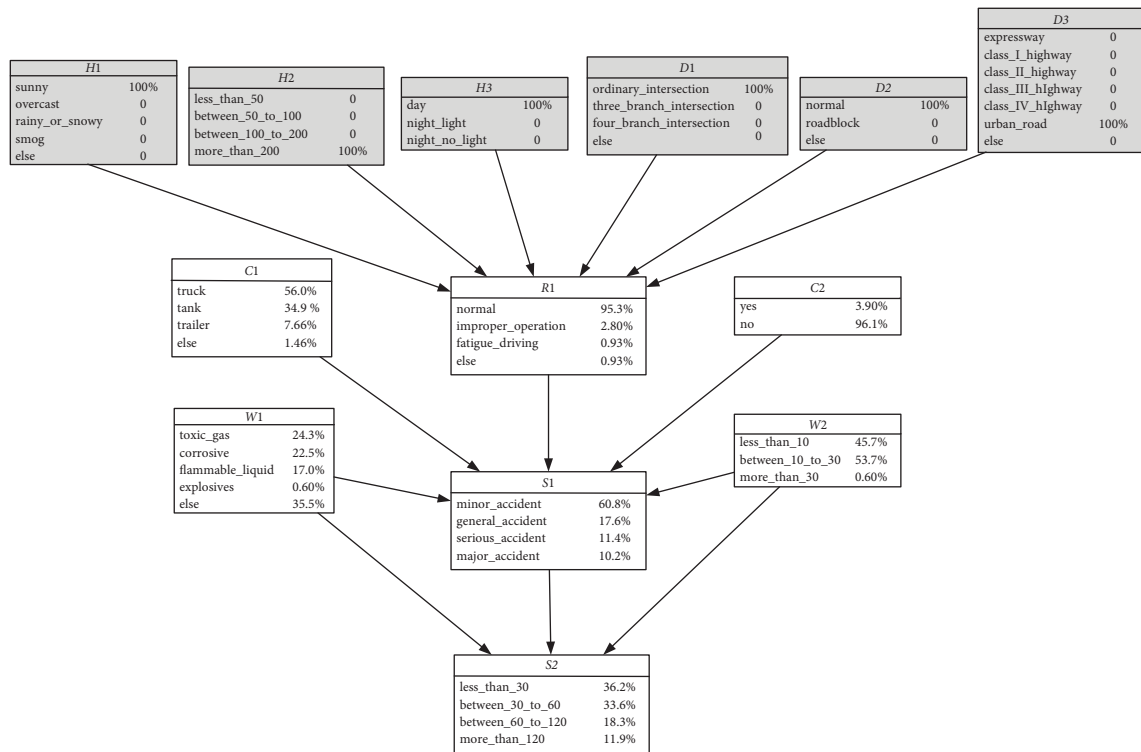| True	Predict S2			
	Less_than_30	Between_30_and_60	Between_60_and_120	More_than_120
Less_than_30	27	0	1	1
Between_30_and_60	1	38	0	1
Between_60_and_120	1	1	15	0
More_than_120	0	1	1	12
Overall relative error		8%		

FIGURE 5: Most normal settings.

about 1.2% compared to the lighting condition "day." This means that it is safer to choose day than night for the transport of hazardous materials.

(3) The effect of intersection type on accident severity

When the intersection type node is set to "three-branch intersection," the probability of a major accident is 15.7%, which is about 2% higher than when the intersection section type is "ordinary intersection." This indicates that the probability of a major accident occurring during the transport of hazardous materials is higher when the hazardous

materials transport vehicle is driving at a three-branch intersection.

(4) The effect of road condition on accident severity

When the road condition node is set to "roadblock," the probability of a major accident is 18.5%, which is about 4% higher than the probability of major accidents when the road surface condition is "normal." This indicates that road condition is the main factor causing major accidents in the transport of hazardous materials.

(5) The effect of the road type on accident severity

When the road type node is set to "expressway," the probability of major accidents is the highest, with a probability value of 18.5%. This means that compared to other road types, there is a higher probability of major accidents when vehicles transport hazardous materials on expressway.

(6) The effect of the driver status on accident severity

When the driver status node is set to "improper operation," the probability of a major accident is 18.7%, which is about twice as high as the probability of a major accident when the driver status is "normal." This indicates that driver status is also a key factor in causing major accidents in the transport of hazardous materials.

(7) The effect of vehicle type on accident severity

When the vehicle type node is set to "tanker," the probability of a major accident is the highest, with a probability value of 16.7%. This means that major accidents are more likely to occur when hazardous materials are transported by tanker.

From the above analysis it can be seen that weather, visibility, lighting, intersection type, road condition, road type, driver status, vehicle type, etc. are all important factors affecting accident severity involving the transport of hazardous materials.

4. Conclusions

(1) This paper constructs a model index system for affecting factors of hazardous materials transport accidents from five dimensions of "people, vehicles, hazmat, roads, and environment" and establishes a high-precision PG-BN model with accident severity and accident processing time as the main nodes. The results show that weather, visibility, lighting, intersection type, road condition, road type, driver status, and vehicle type are all important factors influencing the severity of hazardous materials transport accidents.

(2) The PG-BN model proposed in this paper can effectively avoid the overfitting or underfitting phenomenon of the traditional BN model and improve the accuracy and reliability of prediction. Combined with probabilistic graphs, the model can better handle uncertainty and incomplete information, more accurately infer the causes of hazardous materials transport accidents and predict the severity of accidents, and better guide the development of safety management and response strategies. To better guide the development of safety management and response strategies, some suggestions are as follows: strengthening weather and environmental monitoring, improving driver quality and training, and improving road infrastructure and maintenance. Through data analysis, potential accident risks can be predicted and warned, providing strong support for safety management. Based on the analysis results of the PG-BN model, we can formulate targeted emergency response plans for different accident influencing factors. Establish a rapid response mechanism to ensure that effective action can be taken quickly to mitigate the consequences of accidents when they occur.

(3) In this paper, this study mainly analyzes the accident level of hazardous materials transportation from five aspects, such as people, vehicles, hazmat, roads, and environment, and finally screened 11 factors to establish the prediction model of accident severity and accident processing time BN. However, the affecting factors considered are still not comprehensive enough, and factors such as enterprise management can also be added to future studies. In future research, we should continue to optimize and improve the PG-BN model, and we can also add factors such as enterprise management. Through in-depth analyzes of the safety management system, safety culture, and personnel training of hazardous materials transport enterprises, we can gain a more comprehensive understanding of the various factors influencing the occurrence of hazardous materials transport accidents in order to improve the accuracy and reliability of their predictions and better support the safety management of hazardous materials road transport.

Data Availability

The data used to support the findings of this study are currently under embargo. Requests for data, 12 months after publication of this article, will be considered by the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] F. G. Cordeiro, B. S. Bezerra, A. S. P. Peixoto, and R. A. R. Ramos, "Methodological aspects for modeling the environmental risk of transporting hazardous materials by road," *Transportation Research Part D: Transport and Environment*, vol. 44, pp. 105–121, 2016.

[2] X. Shen and S. Wei, "Application of XGBoost for hazardous material road transport accident severity analysis," *IEEE Access*, vol. 8, pp. 206806–206819, 2020.

[3] X. Shen and S. Wei, "Severity analysis of road transport accidents of hazardous materials with machine learning," *Traffic Injury Prevention*, vol. 22, no. 4, pp. 324–329, 2021.

[4] C. Ma, J. Zhou, and D. Yang, "Causation analysis of hazardous material road transportation accidents based on the ordered logit regression model," *International Journal of Environmental Research and Public Health*, vol. 17, no. 4, p. 1259, 2020.

[5] Y. Xing, S. Chen, S. Zhu, Y. Zhang, and J. Lu, "Exploring risk factors contributing to the severity of hazardous material transportation accidents in China," *International Journal of Environmental Research and Public Health*, vol. 17, no. 4, p. 1344, 2020.

[6] J. Pearl, "Bayesian networks A model of self-activated memory for evidential reasoning," in *Proceedings of the Conference of the Cognitive Science Society*, Nashville, TN, USA, August 1985.

[7] L. Zhao, X. Wang, and Y. Qian, "Analysis of factors that influence hazardous material transportation accidents based on Bayesian networks: a case study in China," *Safety Science*, vol. 50, no. 4, pp. 1049–1055, 2012.

[8] X. Ma, Y. Xing, and J. Lu, "Causation analysis of hazardous material road transportation accidents by bayesian network using genie," *Journal of Advanced Transportation*, vol. 2018, Article ID 6248105, 12 pages, 2018.

[9] Y. Li, D. Xu, and J. Shuai, "Real-time risk analysis of road tanker containing flammable liquid based on fuzzy Bayesian network," *Process Safety and Environmental Protection*, vol. 134, pp. 36–46, 2020.

[10] S. J. Hashemi, F. Khan, and S. Ahmed, "Multivariate probabilistic safety analysis of process facilities using the Copula Bayesian Network model," *Computers & Chemical Engineering*, vol. 93, pp. 128–142, 2016.

[11] W. Huang, Y. Zhang, X. Kou, D. Yin, R. Mi, and L. Li, "Railway dangerous goods transportation system risk analysis: an Interpretive Structural Modeling and Bayesian Network combining approach," *Reliability Engineering & System Safety*, vol. 204, Article ID 107220, 2020.

[12] M. Sun, R. Zhou, C. Jiao, and X. Sun, "Severity analysis of hazardous material road transportation crashes with a Bayesian network using Highway Safety Information System data," *International Journal of Environmental Research and Public Health*, vol. 19, no. 7, p. 4002, 2022.

[13] X. Y. Shen, D. L. Xiao, S. S. Wei, X. Y. Sui, and Z. S. Zhu, "Analysis of road transportation accidents of dangerous goods based on fault tree analysis and Bayesian network," *Journal of Safety and Environment*, vol. 22, pp. 338–346, 2022.

[14] S. Ding, X. Pan, D. Zuo, W. Zhang, and L. Sun, "Uncertainty analysis of accident causality model using Credal Network with IDM method: a case study of hazardous material road transportation accidents," *Process Safety and Environmental Protection*, vol. 158, pp. 461–473, 2022.

[15] L. Pan, Y. Zheng, J. Zheng et al., "Characteristics of chemical accidents and risk assessment method for petrochemical enterprises based on improved FBN," *Sustainability*, vol. 14, no. 19, Article ID 12072, 2022.

[16] H. Wang, Y. Zhang, R. Li, Y. Cui, A. He, and W. Jiang, "Analyzing the risk factors of hazardous chemical road transportation accidents based on grounded theory and a bayesian network," *Sustainability*, vol. 15, no. 24, Article ID 16657, 2023.

[17] J. Cheng, B. Wang, C. Cao, and Z. Lang, "A quantitative risk assessment model for domino accidents of hazardous chemicals transportation," *Processes*, vol. 11, no. 5, p. 1442, 2023.

[18] S. Wei, X. Shen, M. Shao, and L. Sun, "Applying data mining approaches for analyzing hazardous materials transportation accidents on different types of roads," *Sustainability*, vol. 13, no. 22, Article ID 12773, 2021.

[19] X. L. Wang, "Analysis on influencing factors of hazardous chemical materials road transport accidents and safety measures," *Journal of Highway and Transportation Research and Development*, vol. 34, no. 22, pp. 115–121, 2017.

[20] L. Bai, P. Liu, C. Y. Chan, and Z. Li, "Estimating level of service of mid-block bicycle lanes considering mixed traffic flow," *Transportation Research Part A: Policy and Practice*, vol. 101, pp. 203–217, 2017.

[21] Z. Sun, J. Wang, Y. Chen, and H. Lu, "Influence factors on injury severity of traffic accidents and differences in urban functional zones: the empirical analysis of Beijing," *International Journal of Environmental Research and Public Health*, vol. 15, no. 12, p. 2722, 2018.

[22] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, pp. 424–438, 1969.

[23] S. Banerjee and S. Ghosal, "Bayesian structure learning in graphical models," *Journal of Multivariate Analysis*, vol. 136, pp. 147–162, 2015.

[24] Y. Yang, K. He, Y. P. Wang, Z. Z. Yuan, Y. H. Yin, and M. Z. Guo, "Identification of dynamic traffic crash risk for cross-area freeways based on statistical and machine learning methods," *Physica A: Statistical Mechanics and its Applications*, vol. 595, Article ID 127083, 2022.

[25] X. Zou and W. L. Yue, "A bayesian network approach to causation analysis of road accidents using netica," *Journal of Advanced Transportation*, 2017.

[26] J. Wu, Y. Lu, S. Shi, R. Zhou, and Y. Liu, "Research on the prediction model of hazardous chemical road transportation accidents," *Journal of Loss Prevention in the Process Industries*, vol. 84, Article ID 105103, 2023.

[27] Y. Yang, K. Wang, Z. Yuan, and D. Liu, "Predicting freeway traffic crash severity using XGBoost-Bayesian network model with consideration of features interaction," *Journal of Advanced Transportation*, vol. 2022, Article ID 4257865, 16 pages, 2022.