Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P2545-cd

Data-driven Bayesian network for risk analysis of urban hydrogen refueling station accident

Jinduo Xing, Jiaqi Qian, Yanxue Wang

School of Mechanical-Electronic and Vehicle Engineering, Beijing University of Civil Engineering and Architecture, Beijing 102600, China. E-mail: jinduoxing@hotmail.com

Hydrogen refueling station (HRS) safety is receiving increasing attention with the growth of hydrogen energy application. Existing risk assessment methods of HRS are primarily based on expert knowledge to develop failure processes. It may lead to insufficient accuracy due to potential subjectivity. This paper aims to conduct a new hybrid risk assessment method by incorporating the latest HRS accident data and physical knowledge into a Bayesian network (BN) model to analyze the key risk influencing factors (RIFs). In this paper, the latest HRS accident data in HIAD 2.1 from 1980 to 2023 is collected. 30 RIFs are identified based on the accident report and physical knowledge. Use Bayesian Search (BS) for structure learning. The expectation maximation algorithm is designed in the parameter learning stage to obtain the data-driven BN model. Additionally, K-fold cross validation is dedicated to test the performance of different BN models. With these developments, new findings and implications are revealed beyond the state-of-the art of HRS risk analysis.

Keywords: Hydrogen refueling stations, Risk analysis, Data-driven Bayesian network, CTGAN, K-fold cross validation, Hydrogen energy.

1. Introduction

As an important part of the hydrogen energy infrastructure, urban HRSs play a key role in the transition of urban transport systems to decarbonization (2024). However, due to the special characteristics of hydrogen, such as low minimum ignition energy (~0.019 mJ), wide explosive limit (18.3%-59%), fast flame speed (~2.79 m/s), and a wide flammability range (4%-75%) (Deng et al. 2023), HRS are exposed to multiple risk scenarios during operation. Without effective risk management, HRSs can lead to serious accidents, endangering the safety of people and environmental stability. For example, in May 2019, an explosion occurred at a hydrogen fuel storage tank located in South Korea, which resulted in two instant deaths and six injuries (China 2019). Therefore, it is particularly important to prevent multiple HRS accidents to ensure operation safety (Hoseyni et al. 2024). For this, it is necessary to carry out risk analyses to discover the key risk factors.

Hydrogen production, storage, and transportation still face numerous challenges. Currently, hydrogen production mostly relies on fossil fuels (gray hydrogen), which accounts for about 96% of global hydrogen production, significantly increasing its environmental impact (Odoi-Yorke et al. 2025). To advance the production of green hydrogen, the research by Hossain et al. focuses on solar energy for hydrogen production. With advances in concentrated solar power (CSP) systems, the efficiency of solar thermal hydrogen production has reached 45%, much higher than that of conventional electrolysis methods. At this stage, storing hydrogen at low temperatures (-253°C) or high pressures (700 bar) remains a significant challenge, both storage efficiency and safety need to be improved. Additionally, the efficiency of electrolysis technology still ranges from 60% to 80%, and production costs remain relative high (Bokde 2025).

Existing methods for risk assessment of hydrogen refueling stations are mainly based on expert knowledge to develop failure processes. This manually designed structure is likely to be subjective, leading to a lack of accuracy. Inadequate data is another challenge for conducting a purely data-driven method. It may lead to inaccurate reasoning when missing data situations are encountered. Regarding the study of HRS accidents, historical data is usually insufficient. To address the missing data, this paper builds a physical-data-driven model by incorporating the latest HRS accident data and physical knowledge into a Bayesian Networks (BN) model, combining the physical model with a data-driven approach.

Hydrogen Incidents and Accidents Database (HIAD 2.1) is the most authoritative on hydrogen accidents. It includes the updating of more than a thousand hydrogen accidents worldwide (as of September 2023), but only 104 accidents are related to HRSs. To address the problem of too few data, in this paper we use Generative Adversarial Network (GAN).

The core concept of GAN is to generate data through an adversarial process that involves two models: a Generator and a Discriminator. Initially, GAN was mainly applied to image synthesis (Tan et al. 2025), art creation, drug discovery (Lu et al. 2022), etc. In these fields, data is often protected by privacy regulations, making it difficult to obtain (Jiangzhou et al. 2024). Both In the foreign exchange market, Kexin Peng and others used GANs to predict exchange rate returns, effectively improving trading decisions and risk assessment capabilities (Peng et al. 2025). Biao He and others used CTGANs to generate synthetic over-sampling datasets, increasing the diversity and quantity of data. It has been shown that the synthetic datasets generated by CTGANs can effectively retain the features of real data and solve the problem of data shortage and imbalance (He et al. 2024).

Leveraging all the above, in this work a physics-informed data-driven BN modeling approach by HRS risk analysis is developed.

The main contributions can be summarized as follows:

(i) Develop a new physics-informed and data-driven risk assessment method for HRSs.

(ii) Use of the CTGAN method for addressing the issue of limited accident data.

(iii) Identification of risk factors for HRS.

2. Methodology

2.1 *Bayesian Networks*

BN is a graphical network based on probabilistic reasoning, which is a combination of probability theory and graph theory. The topology of a BN is a directed acyclic graph, where the nodes represent random variables, which can be observable variables or hidden variables, unknown parameters, etc. The fundamentals of BN are given in formula (1) and (2) (Liu et al. 2022). By considering the conditional dependencies of n random variables *A1, A2, …, An,* a directed acyclic graph with n nodes depicts the joint probability P(U) of variables *U = {A1, A2, …, An}* (Hoseyni, Mesbah Mostafa 2024).

$$P(U) = \prod_{i=1}^{n} P(A_i|P_a(A_i)) \quad (1)$$

where $Pa\ (A_i)$ denotes the parent node of variable A_i in BN.

Based on BN's theorem, given new observation or evidence E, BN can update the prior probabilities of variables with rendering posterior probabilities (Xing et al. 2022).

$$P(U|E) = \frac{P(E|U)P(U)}{P(E)} = \frac{P(U,E)}{\sum_U P(U,E)} \quad (2)$$

3. Data Collection and Processing

3.1 *Data collection and pre-processing*

The initial data for this study was obtained from the HIAD 2.1 database (Tools 2024). This database has a wide range of data sources that are covered globally. We have collated accident data from 1980 to September 2023 from it. And, to avoid duplication of generalized data, only one database is referred to in this study.

The first step was to extract hydrogen accidents related to HRS from HIAD 2.1, and by sorting and categorizing the events and summarizing the previous literature, leakage, fire and explosion can be identified as the main types of HRS accidents (Xing, Wu 2022). Leakage refers to the accidental escape of hydrogen from storage, transportation, or use equipment, leading to an increase in the hydrogen concentration in the air, which raises the risk of fire and explosion. Fire occurs when hydrogen gas, mixed with air, combusts in the presence of an ignition source, potentially releasing heat and posing dangers to the surrounding environment. Explosion, on the other hand, is a violent reaction triggered by an external ignition source or other stimuli when the hydrogen-air mixture reaches its explosive limits in an enclosed or confined space, resulting in a strong shockwave and instantaneous release of energy.

After screening the events in the HIAD 2.1 database, we have compiled 108 events that can be used for reference. By analyzing the causes, passages, and results of these events, 51 RIFs were obtained. However, some of the nodes appeared less frequently, and fewer occurrences may lead to the lack of influence of these nodes in the model, thus affecting the overall structure learning. Therefore, it is necessary to merge these nodes. After a series of adjustments and integrations, we finally obtained 30 new RIFs, as shown in Table 1.

These new RIFs more centrally reflect the characteristics of various potential risks and cover different types of events and factors, making subsequent analysis and research more efficient.

Table 1 RIFs for HRS accidents

Symbol	Event	Symbol	Event
X1	Unexpected source of ignition	X2	Inadequate maintenance
X3	Vehicle collision	X4	Uneven flange preload
X5	Insufficient screw torque value	X6	Non-compliance with emergency procedures
X7	Mishandling	X8	Lack of risk assessment

X9	Poor system design	X10	Lack of training or experience
X11	Inadequate organizational systems	X12	Unreasonable provisions
X13	Emergency management deficiencies	X14	Pipe joint seal failure
X15	Failure of environmental hydrogen detection device	X16	Failure of pressure detection device
X17	Unreasonable detection range or layout	X18	Failure of the emergency response system
X19	Electromagnetic threshold fault (physics)	X20	Pressure relief device failure
X21	Filter failure	X22	Electrical short circuit or overload
X23	Hose fitting rupture	X24	Hydrogen embrittlement
X25	Weld cracking	X26	Inadequate material performance
X27	Abnormal heating of hydrogen due to throttling effect	X28	Natural disaster
X29	Third-party impact	X30	Radioactive isotope of hydrogen

4.2 *Data augmentation*

Insufficient data samples may lead to overfitting of the model during training, thus affecting its generalization ability and prediction accuracy. In this case, data augmentation techniques are particularly important and valuable (Yoo et al. 2024).

GAN is an innovative deep learning model consisting of two neural networks: a Generator and a Discriminator. These two networks are trained in an adversarial way to form a dynamic game process. During the training process, the generator and the discriminator are updated by alternating optimization.

To develop a data-driven BN model, the size of the dataset needs to be further extended to enhance the model's expressive capability. To this end, in this paper, a synthetic dataset containing 1180 samples is generated using CTGAN.

After extending data size, assessing the quality of the generated data is a crucial step. This process not only ensures that the generated data has practical application value but also improves the effectiveness of subsequent model construction. The expert review can judge the quality of the generated data by judging the loss function of the generator and discriminator that generate the data to ensure that the generated data meets the practical application requirements (Li et al. 2023a). The loss curves of the two need to tend to be in equilibrium, as shown in Figure 1, which plots the loss functions of the generator and the discriminator for the data of this study.

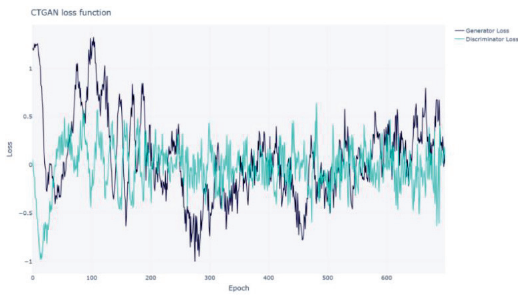

Figure 1 CTGAN loss function

4. Case Study

4.1 *Structure learning and parameter learning*

4.1.1 Structure learning

In this paper, a physical model is combined with a data-driven approach to form a physical-data-driven model that can handle a large amount of uncertainty. The HRSs system is a complex multilevel and multifactor system with many uncertainties, such as equipment failures, fluctuations in energy supply, and changes in the external environment. BN can handle these uncertainties and provide feasible decision support based on data-driven reasoning.

This study adopts a physics-informed and data-driven approach to structure learning with GeNIe 5.0 as a means of automatically identifying relationships and dependency structures between variables. In this section, the processed data information is first input into the system, and then the physical information is used as an additional input to define the coercive relationship between different nodes.

To better compare the advantages and disadvantages of each model, three different types of structure learning algorithms were selected for systematic analysis and comparison. These three algorithms are constraint-based structure learning (PC), score-based structure learning (BS), and probabilistic and graph theory-based structure learning (GTT).

4.1.2 Parameter learning

Parameter learning is the process of estimating unknown parameters in a model from data given by the model structure. It is concerned with optimizing the parameters of the model so that the model can best describe the data. In this section, EM algorithm is utilized to launch the parameter learning of the BN model. The EM algorithm consists of two steps, the E-step and the M-step (Hoseyni, Mesbah Mostafa 2024).

In the E-step, the posterior distribution of the hidden variables is computed given the observed data and the current parameters. Using the structure of BN, the expectation of a hidden variable can be computed by methods such as forward-backward algorithms or variational inference, as shown in formula (3) (Li et al. 2023b):

$$Q\left(\theta|\theta^{(t)}\right) = E_{\left(Z|X,\theta^{(t)}\right)} LnP\left(Z,X|\theta^{(t)}\right) \qquad (3)$$

where $\theta^{(t)}$ is the current parameter. X is the given observation. $Q\ (\theta|\ \theta^{(t)})$ refers to the function that computes the expected value of the hidden variable Z under the current parameter $\theta^{(t)}$. $P\ (Z,\ X|\theta^{(t)})$ refers to the posterior probability distribution of the hidden variable Z.

In the M-step, the parameters of the BN are updated using the expectation of the hidden variables computed in the E-step. For CPT, the parameters can be updated by maximizing the likelihood function using the current hidden variable expectations, as shown in formula (4) (Li, Ren 2023b):

$$\theta^{(t+1)} = arg\max_{\theta} Q\left(\theta|\theta^{(t)}\right) \qquad (4)$$

where $\theta^{(t+1)}$ denotes the updated model parameters in the $t+1$st iteration.

If the change in the log-likelihood function is less than a set threshold ϵ in successive iterations, the algorithm is considered to have converged , as shown in formula (5) .

$$\left|L\left(\theta^{(t+1)}\right) - L\left(\theta^{(t)}\right) < \epsilon \qquad (5)$$

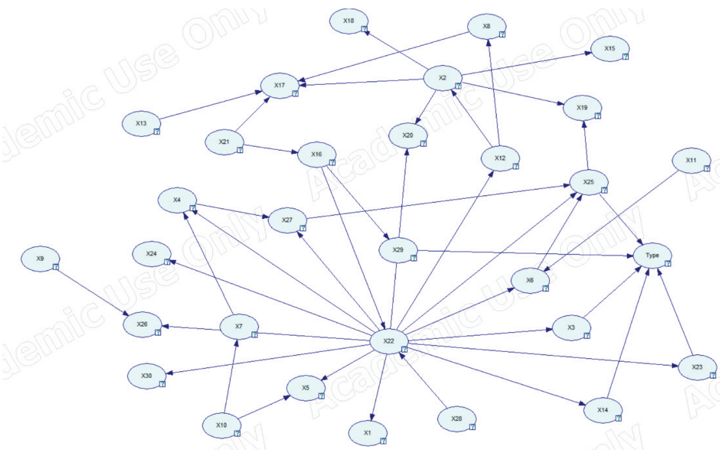

Figure 2 Structure learning developed by BS algorithm

4.2 *Model validation*

Model validation aims at evaluating the validity and reliability of the constructed BN model to ensure that the model accurately represents the underlying connections in the data and may provide reliable predictions or inferences. In this case, the most appropriate method for model evaluation is K cross-validation. In machine learning practice, many studies and experiments have shown that cross-validation using k=10 usually yields better generalization performance.

The outputs of the K-fold cross-validation include Overall Accuracy (OA), Precision, Recall, and F1 score (Li, Ren 2023a). The predictions made by the BS algorithm were generally greater than those made by the GTT and PC algorithms. This shows the good performance of the network structure based on the BS algorithm. The specific results of the calculations are in Table 2.

Table 2 Predictive performance metrics of the BS algorithms

	OA	Precision	Recall	F1 score
BS	91.8%	96.3%	99.4%	95.1%
GTT	90.2%	96.4%	86.4%	92.7%
PC	89.2%	89.4%	97.8%	93.8%

Figure 2 shows the HRS accident network model developed by the BS algorithm.

This paper chooses to use the BN model developed by the BS and EM algorithm for the HRS accident study in the next research.

4.3 *Sensitivity analysis*

Sensitivity analysis is a technique that can help to understand how input parameters affect output parameters, which is a effective means of BN verification. We observe the sensitivity of the model to different nodes by perturbing the data to a certain extent. This reveals which factors have a large impact on the model results. A sensitivity analysis was performed to obtain the impact of different factors on the outcome of the three accident types (Li, Ren 2023b).

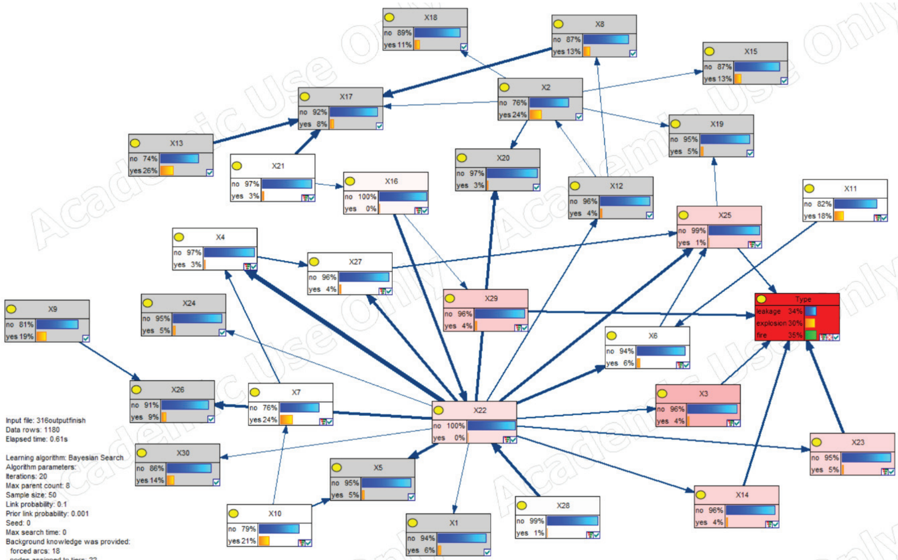

Figure 3 Sensitivity network for HRS accidents

As shown in Figure 3, a detailed sensitivity analysis was conducted for the entire BN model. This analysis takes Type as the target node, and the goal is in exploring the degree of influence and importance of other nodes on this target node. The red nodes (X3, X14, X22, X23, X25, X29) are critical to the posterior probability distribution of the target node. These nodes represent the factors that have the greatest impact on the node Type, and any adjustment to these nodes may significantly change the probability values of the target nodes. Therefore, these red nodes need to be prioritized during the optimization of the model to ensure the validity and accuracy of the model. The pink nodes also play an important role in influencing the target nodes, and although their influence is relatively weak. The gray nodes show a very low probability of influencing the target nodes throughout the analysis (Odoi-Yorke, Agyekum 2025). The factors represented by these nodes have almost no significant effect on the posterior probability distribution of Type. Therefore, the adjustment and optimization of these grey nodes can be ranked as the last consideration in the model optimization process to reduce unnecessary complexity and computational cost (Odoi-Yorke, Agyekum 2025).

The width of the edges in the network indicates the sensitivity of the influence to the target variable, the wider the edge, the greater the influence (Liu, Yu 2022). There are 40 connecting lines in the model, and there are five connecting lines pointing to the target node, all of which have the largest influence on the target node.

4.3.1 *Sensitivity analysis of leakage accident*

Figure 4 shows the comparison of the prior probability and posterior probability of different nodes under hydrogen leakage accident. Figure 5 shows the tornado diagram of the target node (Type) when the accident type is set to leakage. The horizontal axis of the bar graph represents the different parameters while the vertical axis represents the range of variation in the state of the target node. The red bar indicates a negative change in the state of the target when the parameter is varied, which means that the change in that parameter needs lead to deterioration in the state of the target node, while the green bar indicates a positive change, which means that the change in the parameter needs lead to improvement in the state of the target node.

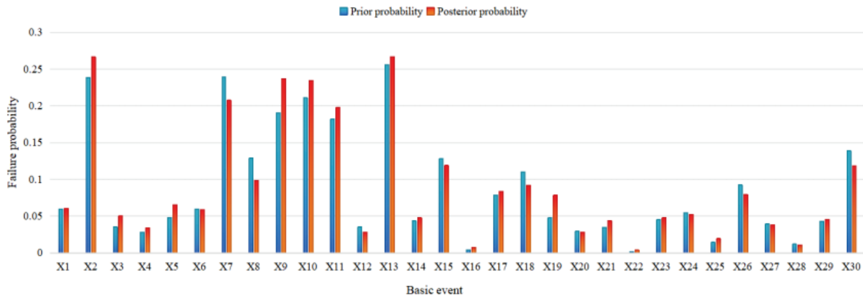

Figure 4 Comparison of the prior and posterior probability under leakage accident

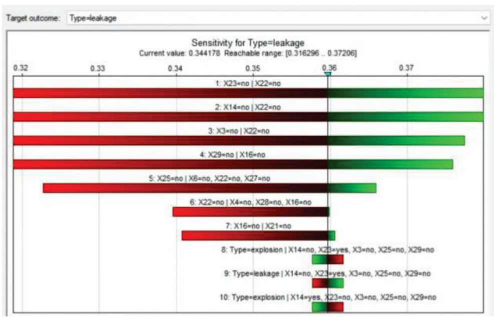

Figure 5 Sensitivity analysis tornado diagram for leakage state

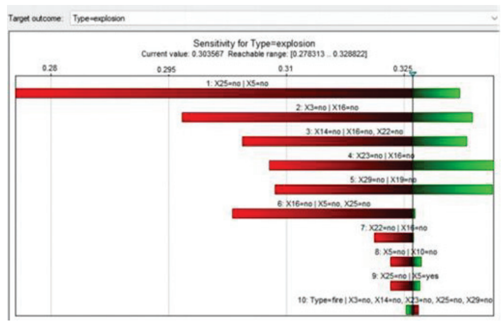

Figure 6 Tornado diagram for explosion accident

Figure 5 shows the 10 factors that have the greatest impact on the leakage accident. The most significant events affecting the state of the target node in the event of a leakage at the HRS are, in descending order of importance, X23, X14, X3, X29 and X 25. Therefore, corresponding measures can be taken for the above nodes to avoid leakage.

Based on the BN diagnosis results, X23, X14 and X3 are affected by node X22, highlighting the severe effects of an electrical short circuit. Similarly, X29 is affected by X16, reflecting problems with regular maintenance, environmental control, and overload protection (Xing, Wu 2022).

4.3.2 *Sensitivity analysis of explosion accident*

In the same way, Figure 7 shows the comparison of the prior probability and posterior probability of different nodes under hydrogen leakage accident. Figure 6 is a tornado diagram of the full range of sensitivity analysis of the BN model when an explosion accident occurs. It is obvious that under the explosion accident, the most influential ones are X25, X3, X14, X23, X29, and X16. X25 (Weld cracking) is the node with the highest sensitivity. X25 is affected by parent node x5 in the explosion state, X3, X14, X23 by X16 and X14 by X22 at the same time.

In addition, compared with the tornado graph in the leakage state, it is found that the nodes that have an impact on the target node remain basically the same. This stability reflects the intrinsic connection between the nodes and suggests that the influence of specific nodes on the system dynamics is continuous and reliable under multiple states (Xing, Wu 2022).

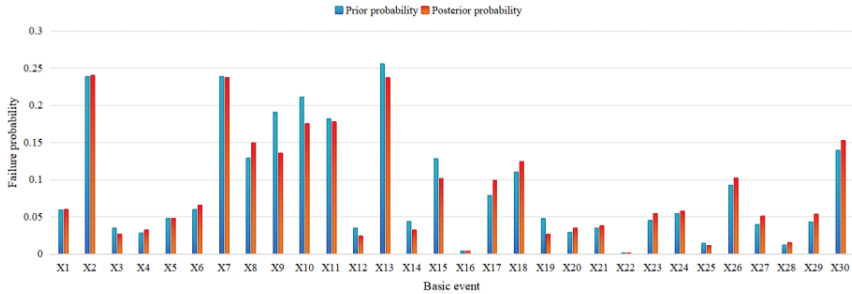

Figure 7 Comparison of the prior and posterior probability under explosion accident

4.3.3 *Sensitivity analysis of fire accident*

When the fire accident occurs, the posterior probability of each node is compared with the prior probability, as shown in Figure 8. Figure 9 is a tornado diagram of the full range of sensitivity analysis of the BN model when the fire accident is set as the target variable of the sensitivity analysis.

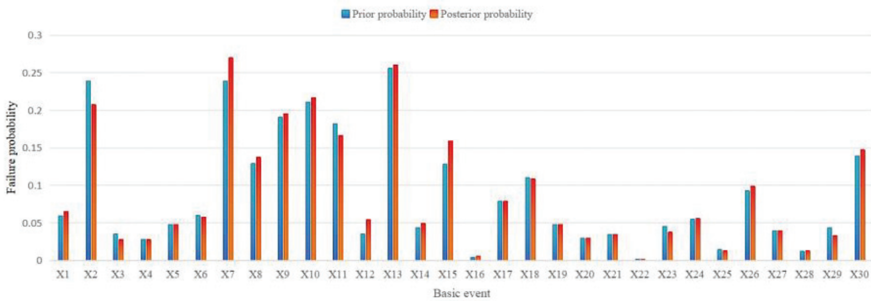

Figure 8 Comparison of the prior and posterior probabilities of nodes under fire accident

Sensitivity order from high to low is as follows: X25, X3, X23, X29, X14. The order of these nodes indicates the degree of their impact on the HRS at the time of the fire event. It is noteworthy that this result is generally consistent with previous analyses of leakage and explosion hazards. This consistency suggests that the effects of the main causal factors are similar in hydrogen station-related accidents, despite the different types of accidents.

In this study, a physics-informed data-driven BN for risk analysis of HRS is proposed. 30 RIFs are identified by applying the HAZID accident data and related literatures. To solve the problem of insufficient data, physics information is integrated to define the coercive relationship of BN nodes.

Additionally, to improve the precision of the results CTGAN is applied to augment the pre-processing data. The BS and EM algorithms are used for structure learning and parameter learning of BN, respectively. Through the results, we found that the critical node with the greatest impact on leakage accidents is X23, and the critical node with the greatest impact on both explosion and fire accidents is X25.

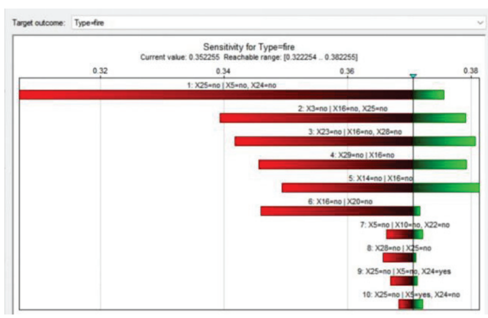

Figure 9 Sensitivity analysis tornado diagram for fire accident

Although the critical node with the greatest impact on these accidents are different, we found that the critical nodes with a greater impact on them are essentially the same in the sensitivity analysis. These critical nodes include X3 (vehicle collision), X14 (pipe joint seal failure), X16 (pressure detection device failure), X23 (hose joint rupture), X25 (weld cracking), and X29 (third party influence). These nodes show high impacts in different accident scenarios, indicating that they are non-negligible factors in the safety management of HRS.

The sensitivity analysis in the paper reveals the key factors influencing the occurrence of accidents at hydrogen refueling stations, and this analytical approach is also applicable to other industries. Hydrogen station managers can embed the BN model into the monitoring system to simulate potential fault scenarios, predict outcomes, and determine the priority of mitigation measures. In future work, dynamic risk assessment of HRS is needed to capture the degradation performance of critical components(Xing et al. 2024). Meanwhile, the dynamic risk assessment method can help predict the time point of equipment failure and optimize the maintenance strategy to avoid unexpected accidents.

References

H2stations.org. 2024 28 Oct 2024]; Available from: https://www.h2stations.org/.

Deng, S., F. Li, H. Luo, et al. 2023 *Lumped Parameter Modeling of SAE J2601 Hydrogen Fueling Tests*. Sustainability. **15**, DOI: 10.3390/su15021448.

China, B. *Cause of Norwegian hydrogen station explosion identified, giving China's hydrogen industry a piece of mind*. 2019 28 Oct 2024]; Available from: http://www.caam.org.cn/chn/38/cate_430/con_5224848.html.

Hoseyni, S.M., M.O. Mesbah Mostafa and J. Cordiner (2024). Mitigating risks in hydrogen-powered transportation: A comprehensive risk assessment for hydrogen refuelling stations, vehicles, and garages. *International Journal of Hydrogen Energy 91*, 1025-1044.

Odoi-Yorke, F., E.B. Agyekum, F.L. Rashid, et al. (2025). Trends and determinates of hydrogen energy acceptance, or adoption research: A review of two decades of research. *Sustainable Energy Technologies and Assessments 73*, 104159.

Bokde, N.D. (2025). A unified framework for multi-type hydrogen production and storage in renewable energy systems. *Energy Conversion and Management: X 25*, 100847.

Tan, H., B. Ma, Y. Xu, et al. (2025). An innovative variant based on generative adversarial network (GAN): Regression GAN combined with hyperspectral imaging to predict pesticide residue content of Hami melon. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 325*, 125086.

Lu, Y., D. Chen, E. Olaniyi, et al. (2022). Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review. *Computers and Electronics in Agriculture 200*, 107208.

Jiangzhou, D., W. Songli, Y. Jianmei, et al. (2024). DGRM: Diffusion-GAN recommendation model to alleviate the mode collapse problem in sparse environments. *Pattern Recognition 155*, 110692.

Peng, K., H. Iima and Y. Kitamura (2025). Predicting FX market movements using GAN with limit order event data. *Finance Research Letters 72*, 106527.

He, B., D. Jahed Armaghani, S. Hin Lai, et al. (2024). Applying data augmentation technique on blast-induced overbreak prediction: Resolving the problem of data shortage and data imbalance. *Expert Systems with Applications 237*, 121616.

Liu, K., Q. Yu, Z. Yang, et al. (2022). BN-based port state control inspection for Paris MoU: New risk factors and probability training using big data. *Reliability Engineering & System Safety 224*.

Xing, Y., J. Wu, Y. Bai, et al. (2022). All-process risk modelling of typical accidents in urban hydrogen refueling stations. *Process Safety and Environmental Protection 166*, 414-429.

Tools, H. *HIAD - hydrogen incident and accident database*. 2024 28 Oct 2014]; Available from: https://h2tools.org/bibliography/hiad-hydrogen-incident-and-accident-database.

Yoo, J.D., H. Kim and H.K. Kim (2024). GUIDE: GAN-based UAV IDS Enhancement. *Computers & Security 147*, 104073.

Li, H., X. Ren and Z. Yang (2023a). Data-driven Bayesian network for risk analysis of global maritime accidents. *Reliability Engineering & System Safety 230*, 108938.

Li, H., X. Ren and Z. Yang (2023b). Data-driven Bayesian network for risk analysis of global maritime accidents. *Reliability Engineering & System Safety 230*.

Xing, J., J. Qian, R. Peng, et al. (2024). Physics-informed data-driven Bayesian network for the risk analysis of hydrogen refueling stations. *International Journal of Hydrogen Energy 110*, 371-385.