Bayesian Network Modeling Applied on Railway Level Crossing Safety

Ci Liang[1,2,5(✉)], Mohamed Ghazel[1,2,5], Olivier Cazier[1,3], Laurent Bouillaut[4], and El-Miloudi El-Koursi[1,2,5]

[1] FCS Railenium, Valenciennes, France
ci.liang@railenium.eu
[2] IFSTTAR-COSYS/ESTAS, 20 Rue Élisée Reclus, BP 70317, 59666 Lille, Villeneuve d'Ascq, France
[3] SNCF Réseau, Paris, France
[4] IFSTTAR-COSYS/GRETTIA, Paris, Marne-la-Vallée, France
[5] University Lille 1, Lille, Villeneuve d'Ascq, France

Abstract. Nowadays, railway operation is characterized by increasingly high speed and large transport capacity. Safety is the core issue in railway operation, and as witnessed by accident/incident statistics, railway level crossing (LX) safety is one of the most critical points in railways. In the present paper, the causal reasoning analysis of LX accidents is carried out based on Bayesian risk model. The causal reasoning analysis aims to investigate various influential factors which may cause LX accidents, and quantify the contribution of these factors so as to identify the crucial factors which contribute most to the accidents at LXs. A detailed statistical analysis is firstly carried out based on the accident/incident data. Then, a Bayesian risk model is established according to the causal relationships and statistical results. Based on the Bayesian risk model, the prediction of LX accident can be made through forward inference. Moreover, accident cause identification and influential factor evaluation can be performed through reverse inference. The main outputs of our study allow for providing improvement measures to reduce risk and lessen consequences related to LX accidents.

Keywords: Bayesian network modeling · Level crossing safety · Train-car collision · Risk assessment · Statistical analysis

1 Introduction

Railway Level crossings (LXs) are potentially hazardous locations where trains, road vehicles and pedestrians move in close proximity to one another. LX safety remains one of the most critical issues for railways despite an ever-increasing focus on improving design and application practices [1,2]. Accidents at European LXs account for about one-third of the entire railway accidents and result in more than 300 deaths every year in Europe [2]. In France, the railway network shows more than 18,000 LXs for 30,000 km of railway lines, which are crossed

daily by 16 million vehicles on average, and around 13,000 LXs show heavy road and railway traffic [3]. Despite numerous measures already taken to improve the LX safety, SNCF Réseau (the French national railway infrastructure manager) counted 100 collisions at LXs leading to 25 deaths in 2014. This number was half the total number of collisions per year at LXs a decade ago, but still too large [4]. In order to significantly reduce the accidents and lessen their related consequences at LXs, an effective risk assessment means is needed urgently.

Many available studies dealing with LX safety have tended to take a qualitative approach to understand the potential factors causing accidents at LXs. These works employ surveys [5], interviews [6], focus group methods [7] or driving simulators [8], rather than collecting real field data. For example, Lenné et al. [9] examined the effect of installing active controls, flashing lights and traffic signals on vehicle driver behavior. This study was achieved through adopting the driving simulation. Tey et al. [10] conducted an experiment to measure vehicle driver response to LXs equipped with stop signs (passive), flashing lights and half barriers with flashing lights (active), respectively. In this study, both a field survey and a driving simulator have been utilized. Although those aforementioned approaches are beneficial to explore the potential factors causing accidents, they still show some limits. For instance, they do not allow for quantifying the contribution degree of these factors. In addition, the reaction of vehicle drivers in simulation scenarios could differ from that in reality, due to the different levels of feeling of danger. Therefore, quantitative approaches based on real field data are indispensable if we want to understand the impacting factors thoroughly and enable the identification of practical design and improvement recommendations to prevent accidents at LXs.

Nowadays, risk analysis approaches are required to deal with increasingly complex systems with a large number of configuration parameters. Therefore, such approaches should satisfy the following requirements:

– Strong modeling ability;
– Easy to specify a risk scenario or a system;
– High computational efficiency.

In the domain of risk assessment, various approaches are adopted for the modeling and analyzing process. Due to the combination of qualitative and quantitative analysis, the Fault Tree Analysis (FTA) developed by H.A. Watson at Bell Laboratories [11] has been widely used for risk analysis in various contexts. FTA is a deductive and top-down method which aims at analyzing the effects of initiating faults and events on a complex system and offering the designer an intuitive high-level abstraction of the system. Compared with the Failure Mode and Effects Analysis (FMEA), which is an inductive and bottom-up analysis method aimed at analyzing the effects of single component or function failures on equipment or subsystems, FTA is more useful in showing how resistant a system is to single or multiple initiating faults. However, one obvious disadvantage of FTA is that it is not clear on failure mechanism, since the causal

relationship between events is not a simple YES or NO (1 or 0). Therefore, FTA is prone to missing the possible initiating faults. In addition, traditional static fault trees cannot handle the sequential interaction and functional dependencies between components. Consequently, it is necessary to employ dynamic methodologies to overcome these weaknesses. Markov Chains (MCs) and their extensions have been mainly used for modeling complex dynamic system behavior and dependability analysis of dynamic systems. Two-state Markov switching multinomial logit models are introduced by [12] to explain unpredictable, unidentified or unobservable risk factors in road safety. Although MCs can elaborate the statistical state transition of different variables, they cannot formalize causal relationships between the various events.

Afterward, risk analysis based on formal modeling expanded. In order to compare the effectiveness of two main Automatic Protection Systems (APSs) at LXs: two-half-barrier APS and four-half-barrier APS, Generalized Stochastic Petri Nets (GSPNs) were used in [13] to analyze the aleatory fluctuations of various parameters involved in the dynamics within the LX area. Over the last few years, Bayesian network (BN), a method of reasoning using probabilities, has been an increasingly popular method used for risk analysis of safety-critical systems or large and complex dynamic systems [14]. In order to obtain proper and effective risk control, risk planning should be performed based on risk causality, which can provide more information for decision making. In this context, a model using BNs with causality constraints (BNCC) for risk analysis was proposed in [15]. In [16], Bouillaut et al. discussed the development of a decision tool realized by hierarchical Dynamic BNs (DBNs), which is dedicated to the maintenance of metro lines in Paris. This modeling work has comprehensively described the rail degradation process, the different diagnosis actors (devices and staff) and the maintenance actions decision. In [17], Langseth and Portinal introduced the applicability of BNs for reliability analysis and offered an instance of BNs application for preventive maintenance. The advantages behind BNs were discussed in this article: (a) BNs constitute a modeling framework, which is particularly easy to use for interaction with domain experts; (b) the sound mathematical formulation has been utilized in BNs to generate efficient learning methods; and (c) BNs are equipped with an efficient calculation scheme which often makes BNs preferable to traditional tools like Fault Trees (FTs). To sum up, the BN technique offers interesting features: the flexibility of modeling, strong modeling power, high computational efficiency and, most importantly, the outstanding advantages involving causality analysis based on both forward inference and reverse inference [18] and the conjunction of domain expertise.

Therefore, based on the above investigation of risk analysis, an approach of Causal Reasoning Analysis based on Bayesian risk model (CRAB) is presented in this paper to deal with the risk assessment at LXs. Namely, a thorough statistical analysis based on the accident/incident data pertaining to French LXs is firstly performed, and the statistical results are used as the import sources of BN risk model. Then, the BN risk model is developed according to the causal

relationships between the accidents and various influential parameters considered. Through the BN risk model, one can quantify the risk level impacted by various potential factors and identify the factors which contribute most to the accidents at LXs, as well as their combined impact on LX safety.

2   Preliminary Introduction of Bayesian Belief Networks

In railways, potential hazards including equipment failures, human errors and some non-deterministic factors, such as environment aspects, may lead to accidents. In fact, causalities between accidents and these impacting factors exist, as shown in Fig. 1. Identifying such causality relationships is a crucial issue in the process of reasoning. In particular, a functional intelligent identification model should have the ability of making reasoning based on the causal knowledge.

The Bayesian belief network (BN) employed to model causality is a graphical model that can be characterized by its structure and a set of parameters [19]. $BN = (P, G)$, where P represents the parameters of prior probabilities that quantify the arcs, while G defines the model structure. $G = (V, A)$, which is a Directed Acyclic Graph (DAG), is comprised by a finite set of nodes (V) linked by directed arcs (A). The nodes represent random variables (V_i) and directed arcs (A_i) between pairs of nodes represent dependencies between the variables [19].

In our study, the BN works based on the theory of probability for discrete distributions. Assume that there is a set of mutually exclusive events: B_1, B_2, ...,B_n and a given event A, such that, $P(A)$ can be expressed as follows:

$$P(A) = \sum_{i=1}^{n} P(B_i) P(A|B_i) \tag{1}$$

According to Bayes' formula:

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{n} P(B_j)P(A|B_j)} \tag{2}$$
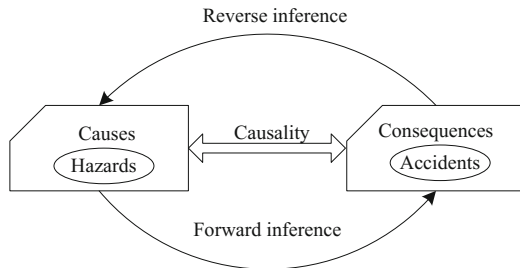
Fig. 1. Reasoning between hazards and accidents.

Equation (2) can be converted into:

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{P(A)} \tag{3}$$

where $P(B_i)$ is the prior probability, $P(B_i|A)$ is the posterior probability.

For any set of random variables in a BN, the joint distribution can be computed through conditional probabilities using the chain rule as shown in Eq. (4):

$$P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{v=1}^{n} P(X_v = x_v | X_{v+1} = x_{v+1}, \ldots, X_n = x_n) \tag{4}$$

Due to the conditional independence, X_v only relates to its parent node $Pa(X_v)$ and is independent of the other nodes. Hence, Eq. (4) can be rewritten as follows:

$$P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{v=1}^{n} P(X_v = x_v | Pa(X_v)) \tag{5}$$

For more details about BN, the reader can refer to the tutorial book on Bayesian networks edited by [20].

3   Methodology

As mentioned before, the present study aims to perform risk assessment at French LXs. The CRAB approach is illustrated to assist our risk assessment based on the accident/incident data collected by SNCF Réseau. Namely, it is applied to assessing the risk level with regard to various impacting factors taken into account and evaluating the contribution degree of these factors. Thus, we pave the way towards identifying the important factors which contribute most to the overall risk.

There are 4 LX types in France [21]:

– SAL4: Automated LXs with four half barriers and flashing lights;
– SAL2: Automated LXs with two half barriers and flashing lights;
– SAL0: Automated LXs with flashing lights but without barriers;
– Crossbuck LXs, without automatic signaling.

As shown in Table 1, SAL2 (more than 10,000) is the most widely used type of LX in France. Moreover, more than 4,000 accidents at SAL2 LXs contributed most to the total number of accidents at LXs from 1974 to 2014. Since the motorized vehicle is the main transport mode causing LX accidents in France [22], considering the train/motorized vehicle (train-MV) collisions, SAL2 LXs also have the most part of LX accidents according to the accident/incident statistics as shown in Fig. 2. Moreover, according to the SNCF statistics, these accidents can be considered as the most representative for LX accidents in general. For all these reasons, our analysis will focus on train-MV accidents occurring at SAL2 LXs.

Table 1. Accidents at different types of LXs in France from 1974 to 2014

Type of LX	Number	#Accident
SAL4	>600	>600
SAL2	>10,000	>4,200
SAL0	>60	>50
Crossbuck LX	>3,500	>700
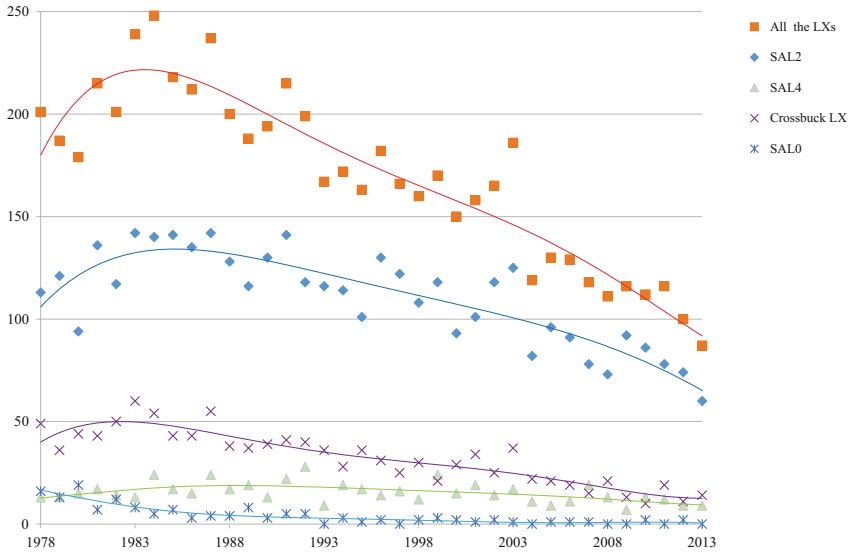
Fig. 2. The number of train-MV collisions at different types of LX from 1978 to 2013

3.1   Data Collection

SNCF Réseau has recorded the detailed elements of each LX accident, including various attributes of LX accidents/incidents, surrounding characteristics of LXs and accident causes, and provides two accident/incident databases to support our study. The first database (D1) records the accident/incident data that cover SAL2 LXs in mainland France from 1990 to 2013.

From D1, the subdataset (SD1) including the data ranging in the decade from 2004 to 2013 is selected, which provides reliable and sufficient information about both LX accidents and static railway, roadway and LX characteristics. Namely, the selected LX inventory presents the LX identification number, the railway line involved, the LX kilometer point, the LX accident timestamp, the average daily railway traffic, the average daily road traffic, the rail speed limit, the LX length and width, the profile and alignment of the entered road and geographic region involved. There are 8,332 public SAL2 LXs included in SD1.

According to the statistics of SNCF Réseau, the majority of train-MV accidents at LXs are caused by motorist violations. Due to the lack of accident causes in SD1, causal relationship analysis cannot be performed with regard to the static factors and motorist behavior. Therefore, we seek another database which records detailed accident causes. Fortunately, the second database (D2) contains the information about SAL2 LX accidents from 2010 to 2013, the LX identification number, the railway line involved and detailed accident causes (including static factors and inappropriate motorist behavior). Thus, using the LX ID and the railway line ID, data merging of these two databases is carried out to create a new database (ND) containing the LX accident information, static railway, roadway and LX characteristics and accident causes related to static factors and motorist behavior. This combined database ND covers LX accidents during a period of 4 years from 2010 to 2013, which forms the basis of our present study.

The detailed accident causes considered in this study are shown in Table 2. Here, a second-level cause is given: corrected moment. The conventional formula of the traffic moment is defined as: Traffic moment = Road traffic frequency × Railway traffic frequency [22]. However, based on the previous analysis of SNCF Réseau, we adopt a variant called "corrected moment" instead (CM for short).

Table 2. Accident causes

Primary causes	Second-level causes	Third-level causes	Explanation
Static factors	Corrected Moment (CM)	Average daily railway traffic (T)	$CM = V^{0.354} \times T^{0.646}$;
		Average daily road traffic (V)	
	Railway speed limit		The maximum permission speed of train within the LX section;
	Alignment		Horizontal road alignment shape: "straight", "curve" or "S";
	Profile		Vertical road profile shape: "normal" or "hump or cavity";
	Width		The width of LX;
	Length		The length of LX that road vehicles need to cross;
	Region risk		Region Risk factor, highlighting the general LX-accident-prone region: *The number of SAL2 accidents over the observation period in the region considered/The number of SAL2 LXs in the region considered*;
Inappropriate motorist behavior	Stall on LX	Blocked on LX	A vehicle is blocked on the SAL2 LX by the external environment;
		Stop on LX	A motorist intentionally stops the vehicle on the SAL2 LX;
	Zigzag violation		A vehicle skirts the half barriers to cross the SAL2 LX;

$CM = V^a \times T^b$, where $b = 1 - a$ and the best value of a in terms of fitting is computed to be $a = 0.354$ according to the statistical analysis performed by SNCF Réseau [23], since railway traffic has a more marked impact on LX accidents than road traffic. Therefore, $(V^{0.354} \times T^{0.646})$ is considered as an integrated parameter that reflects the combined exposure frequency of both railway and road traffic.

3.2 Bayesian Risk Model Establishment

Variable Definition. Based on the combined database ND, the statistical results are organized as input sources which will be imported to the BN risk model. Data discretization is applied on continuous variables. Namely, the continuous variables, i.e., "Average Daily Road Traffic", "Average Daily Railway Traffic", "Railway Speed Limit", "Width", "Length" and "Corrected Moment", are divided into 3 groups and each group has the similar number of samples. As for the "Region Risk" factors corresponding to 21 regions in mainland France, they are divided into 3 groups as well, ranked according to the risk level in descending order, and each group contains 7 region risk factors. As for the finite discrete variables, i.e., "Alignment", "Profile", "Stall on LX", "Zigzag Violation", "Blocked on LX" and "Stop on LX", we allocate an individual state to each value of the variable. The consequence severity of SAL2 accidents [24] is defined according to the number of fatalities and injuries in an SAL2 accident. The definition of consequence severity pertaining to an SAL2 accident is shown in Table 3. Five levels of consequence severity are set according to the number of fatalities, severe injuries and minor injuries caused by the accident, respectively. The consequence severity increases progressively from level 1 to 5. Thus, a summary of states of each node in the BN risk model is offered in Table 4.

Table 3. Consequence severity definition

Consequence severity	Level 1	Level 2	Level 3	Level 4	Level 5
$0 \leq$ fatalities < 5, $0 \leq$ severe injuries < 5, $0 \leq$ minor injuries < 20;	×	–	–	–	–
$0 \leq$ fatalities < 5, $0 \leq$ severe injuries < 5, $20 \leq$ minor injuries;	–	×	–	–	–
$0 \leq$ fatalities < 5, $5 \leq$ severe injuries, $0 \leq$ minor injuries < 20;	–	–	×	–	–
$0 \leq$ fatalities < 5, $5 \leq$ severe injuries, $20 \leq$ minor injuries;	–	–	×	–	–
$5 \leq$ fatalities, $0 \leq$ severe injuries < 5, $0 \leq$ minor injuries < 20;	–	–	–	×	–
$5 \leq$ fatalities, $0 \leq$ severe injuries < 5, $20 \leq$ minor injuries;	–	–	–	×	–
$5 \leq$ fatalities, $5 \leq$ severe injuries, $0 \leq$ minor injuries < 20;	–	–	–	–	×
$5 \leq$ fatalities, $5 \leq$ severe injuries, $20 \leq$ minor injuries;	–	–	–	–	×

Table 4. States of nodes in the BN risk model

Node name	Node property	Node state
Corrected Moment (CM)	Chance node	CM_below_19 ($0 \leq$ CM < 19),
		CM_19_49 ($19 \leq$ CM < 49),
		CM_49_up ($49 \leq$ CM);
Average Daily Railway Traffic (ADRT)	Chance node	ADRT_below_9 ($0 \leq$ ADRT < 9),
		ADRT_9_25 ($9 \leq$ ADRT < 25),
		ADRT_25_up ($25 \leq$ ADRT);
Average Daily Road Vehicle (ADRV)	Chance node	ADRV_below_72 ($0 \leq$ ADRV < 72),
		ADRV_72_403 ($72 \leq$ ADRV < 403),
		ADRV_403_up ($403 \leq$ ADRV);
Railway Speed Limit (RLS)	Chance node	RLS_below_70 ($0 \, \mathrm{km/h} \leq$ RLS $< 70 \, \mathrm{km/h}$),
		RLS_70_110 ($70 \, \mathrm{km/h} \leq$ RLS $< 110 \, \mathrm{km/h}$),
		RLS_110_up ($110 \, \mathrm{km/h} \leq$ RLS);
Alignment	Chance node	Straight, C_shape, S_shape;
Profile	Chance node	Normal, Hump_cavity;
Width (W)	Chance node	W_below_5 ($0 \, \mathrm{m} \leq$ W $< 5 \, \mathrm{m}$),
		W_5_6 ($5 \, \mathrm{m} \leq$ W $< 6 \, \mathrm{m}$),
		W_6_up ($6 \, \mathrm{m} \leq$ W);
Length (L)	Chance node	L_below_7 ($0 \, \mathrm{m} \leq$ L $< 7 \, \mathrm{m}$),
		L_7_11 ($7 \, \mathrm{m} \leq$ L $< 11 \, \mathrm{m}$),
		L_11_up ($11 \, \mathrm{m} \leq$ L);
Region Risk (R)	Chance node	R_low (region with low risk level),
		R_medial (region with medial risk level),
		R_high (region with high risk level);
Stall on LX	Chance node	True, False;
Blocked on LX	Chance node	True, False;
Stop on LX	Chance node	True, False;
Zigzag Violation	Chance node	True, False;
Motorist Behavior Accident	Chance node	True, False;
Static Factor Accident	Chance node	True, False;
SAL2 MV Accident	Chance node	True, False;
Fatalities (F)	Chance node	F_0_5 ($0 \leq$ F < 5), F_5_up ($5 \leq$ F);
Severe Injuries (S)	Chance node	S_0_5 ($0 \leq$ S < 5), S_5_up ($5 \leq$ S);
Minor Injuries (M)	Chance node	M_0_20 ($0 \leq$ M < 20), M_20_up ($20 \leq$ M);
Consequence Severity	Deterministic node	Level_1, Level_2, Level_3, Level_4, Level_5;

Model Structure. Artificial restrictions are adopted to build the model structure, which means the model structure is defined according to the causal relationships between accident occurrence and influential variables based on expert proposes, instead of using general structure learning methods, since the general structure learning methods suggest us unreasonable model structures which

Table 5. Spearman correlation checking

Static factors / Motorist behavior	Railway speed limit	Length	Width	Corrected moment	Alignment	Profile	Region risk factor
Blocked on LX	-0.0444	0.0031	-0.0341	-0.0525	-0.1769	-0.0352	-0.0432
Stop on LX	0.0179	-0.0668	-0.1138	-0.0402	-0.0329	-0.0307	-0.0420
Zigzag violation	0.0347	0.0374	0.1143	0.2118	-0.0462	-0.0221	0.1238
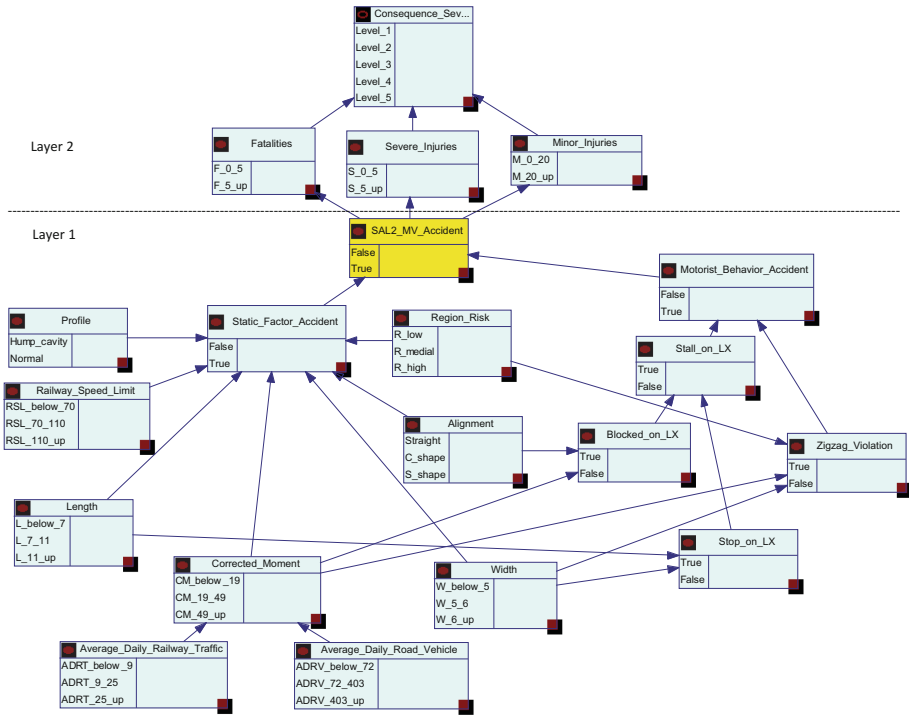
Fig. 3. BN risk model

are inconsistent with the causal relationships in reality and impede identification of important accident causes. It is worth noticing that there are still some potential connections between static factors and motorist behavior. The Spearman correlation checking is adopted to explore important connections and filter off negligible connections between these two kinds of variables. As shown in Table 5, the absolute values of correlation bigger than 0.05 are highlighted (Red color highlights negative values and green color highlights positive values). Their corresponding connections will be considered in our model. Conditional probability parameters are generated based on the real field accident/incident data.

The final model is developed as shown in Fig. 3, which contains 3,132 conditional probabilities.

4   Analysis and Discussion

As shown in Fig. 3, the risk model contains two layers: (1) **Layer 1** is used for predicting accident occurrence and diagnosing influential factors; (2) **Layer 2** is used for evaluating consequences related to LX accidents. The "SAL2 MV Accident" node is the key node connecting the two layers, as well as the target node of accident prediction. Note that the Receiver Operating Characteristic (ROC) curve and the Area Under the ROC Curve (AUC) [25] have already been adopted to ensure that the model performance is sound (the AUC values of key consequence node prediction, i.e., "SAL2 MV Accident", "Fatalities", "Severe injuries" and "Minor injuries", are all bigger than $0.9 > 0.5$: the standard limit); while the detailed validation process is not presented here due to space limitation.

One can estimate the probability of a train-MV accident occurring at an SAL2 LX through forward inference based on the BN risk model. As shown in Fig. 4, the general probability of a train-MV accident influenced by the interaction of all factors considered, is estimated as almost 0.0061. In detail, the probability of a train-MV accident caused by static factors is about 0.0011 and the probability of a train-MV accident caused by inappropriate motorist behavior is about 0.0049. Moreover, fatalities and severe injuries caused by the accident are, to a large extent, fewer than 5. Minor injuries caused by the accident are most likely to be fewer than 20. Thus, the consequence severity level are most likely to be level 1. However, Fig. 5 shows that the probability of a train-MV accident occurring at a SAL2 would increase to 0.0107 if all the second-level causes occur, namely, "Corrected Moment" in the "CM_49_up" group, "Railway Speed Limit" in the "RSL_110_up" group, "Alignment" in the "S_shape" group, "Profile" in the "Hump_cavity" group, "Width" in the "W_6_up" group, "Length" in the "L_11_up" group, "Region Risk" in the "R_high" group, "Stall on LX" being

SAL2_MV_Accident		Static_Factor_Accident		Motorist_Behavior_Accident	
False	0.99390817	False	0.99894278	False	0.99511565
True	0.0060918269	True	0.0010572159	True	0.0048843469

Fatalities		Severe_Injuries		Minor_Injuries	
F_0_5	0.99999427	S_0_5	0.99999323	M_0_20	0.99998072
F_5_up	5.7331767e-006	S_5_up	6.7722109e-006	M_20_up	1.9277586e-005

Consequence_Severity	
Level_1	0.99996826
Level_2	1.9238033e-005
Level_3	6.7658374e-006
Level_4	5.7268032e-006
Level_5	6.3735037e-009

Fig. 4. General prediction results

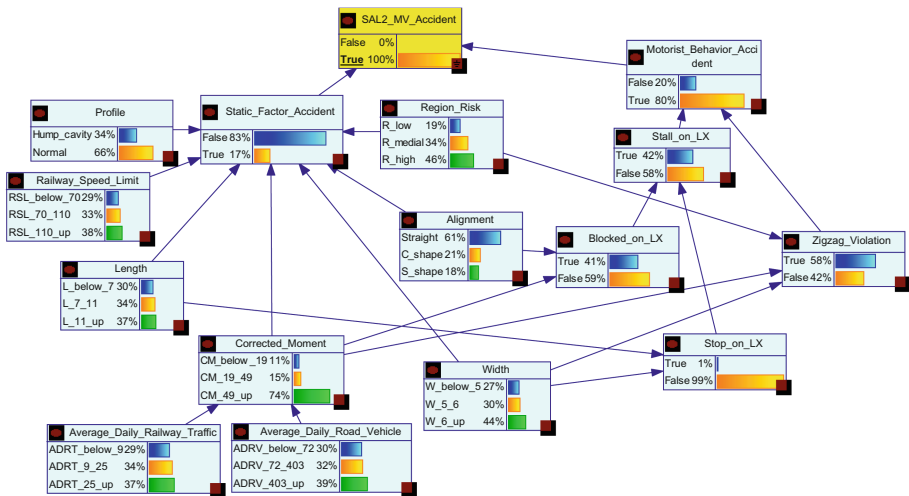**Fig. 5.** Prediction results when second-level causes occur

Fig. 6. Cause diagnosis when a train-MV accident occurs

true and "Zigzag Violation" being true. The related consequences are likely to be severer as well.

Subsequently, the "SAL2 MV Accident = True" state is configured as the targeted state. In this way, one can assess the contribution degree of each influential factor to train-MV accident occurrence through reverse inference. Detailed results are given in Fig. 6. It is worth noticing that accidents caused by inappropriate motorist behavior contribute 80% to the entire train-MV accidents at SAL2 LXs, while accidents caused by static factors contribute only 17%. As for inappropriate motorist behavior, "Zigzag violation" is more significant than "Stall on LX" in terms of causing train-MV accidents, due to the contribution of 58% (compared with 42% contribution of "Stall on LX"). On the other hand, in terms of static factors, when a train-MV accident occurs at a SAL2 LX, this LX

has the probabilities of 74%, 38%, 44%, 37% and 46% respectively involved in the most risky situations that "Corrected Moment" in the "CM_49_up" group, "Railway Speed Limit" in the "RSL_110_up" group, "Width" in the "W_6_up" group, "Length" in the "L_11_up" group and "Region Risk" in the "R_high" group. These results indicates that more attention needs to be paid to LXs having the above risky static characteristics. Moreover, technical solutions need to be implemented to prevent motorist zigzag violations, for example, transforming SAL2 LXs into SAL4 LXs (Four-half barrier systems) or SAL2F (two-full barrier LXs) or installing median separators between opposing lanes of road traffic in front of SAL2 LXs.

5   Conclusions

The contributions of the present study are as follows: the approach of Causal Reasoning Analysis based on Bayesian risk model (CRAB) is proved to be fruitful and practical when analyzing French LX accidents. Although the conditional probabilities of our BN risk model is tailored to SAL2 LX accidents in France, the CRAB approach and the model structure can be applied to different contexts pertaining to LX safety. Based on the CRAB approach, various important static factors pertaining to LX safety, namely, the corrected moment, the rail speed limit, the LX length and width, the profile and alignment of the entered road and geographic region involved, and significant inappropriate motorist behavior, i.e., zigzag violation, blocked on LX and stopping on LX, have been analyzed meticulously. Moreover, the application of CRAB to investigating LX safety allows us to not only predict the probability of accident occurrence, but also evaluate related consequence severity level, quantify the respective contribution degrees of the above influential factors to the overall LX risk and identify the most risky factors, which are rarely achieved in many existing related works. Besides, in our study, expert knowledge is integrated with real field data to optimize the model structure, so as to neglect inappropriate connections to facilitate highlighting the main causes.

In summary, the outcomes of the BN risk model offer a significant perspective on potential parameters causing LX accidents and pave the way for identifying practical design measures and improvement recommendations to prevent accidents at LXs. In future works, a thorough analysis on inappropriate motorist behavior will be carried out due to its significant contribution to LX accident occurrence. In addition, practical solutions will be proposed to improve LX safety according to the analysis results of the BN risk model and the effectiveness of these solutions (e.g., transforming SAL2 LXs into SAL4 LXs or SAL2F or installing median separators) will be investigated.

References

1. Ghazel, M.: Using stochastic Petri nets for level-crossing collision risk assessment. IEEE Trans. Intell. Transp. Syst. **10**(4), 668–677 (2009)
2. Liu, B., Ghazel, M., Toguyeni, A.: Model-based diagnosis of multi-track level crossing plants. IEEE Trans. Intell. Transp. Syst. **17**(2), 546–556 (2016)
3. SNCF Réseau World Conference of Road Safety at Level Crossings (Journée Mondiale de Sécurité Routière aux Passages à Niveau), France (2011). http://www.planetoscope.com/automobile/1271-nombre-de-collisions-aux-passages-a-niveau-en-france.html
4. SNCF Réseau: 8th National Conference of Road Safety at Level Crossings (8ème Journée Nationale de Sécurité Routière aux Passages à Niveau), France (2015). http://www.sncf-reseau.fr/fr/dossier-de-presse-8eme-journee-nationale-de-securite-routiere-aux-passages-a-niveau
5. Wigglesworth, E.C.: A human factors commentary on innovations at railroadhighway. J. Saf. Res. **32**(3), 309–321 (2001)
6. Read, G.J., Salmon, P.M., Lenné, M.G., Stanton, N.A.: Walking the line: understanding pedestrian behaviour and risk at rail level crossings with cognitive work analysis. Appl. Ergon. **53**, 209–227 (2016)
7. Stefanova, T., Burkhardt, J.-M., Filtness, A., Wullems, C., Rakotonirainy, A., Delhomme, P.: Systems-based approach to investigate unsafe pedestrian behaviour at level crossings. Accid. Anal. Prev. **81**, 167–186 (2015)
8. Larue, G.S., Rakotonirainy, A., Haworth, N.L., Darvell, M.: Assessing driver acceptance of intelligent transport systems in the context of railway level crossings. Transp. Res. Part F Traffic Psychol. Behav. **30**, 1–13 (2015)
9. Lenné, M.G., Rudin-Brown, C.M., Navarro, J., Edquist, J., Trotter, M., Tomasevic, N.: Driver behaviour at rail level crossings: responses to flashing lights, traffic signals and stop signs in simulated rural driving. Appl. Ergon. **42**(4), 548–554 (2011)
10. Tey, L.S., Ferreira, L., Wallace, A.: Measuring driver responses at railway level crossings. Accid. Anal. Prev. **43**(6), 2134–2141 (2011)
11. Ericson, C.A., Li, C.: Fault tree analysis. In: Proceedings of 17th International Systems Safety Conference, Orlando, Florida, pp. 1–9 (1999)
12. Malyshkina, N.V., Mannering, F.L.: Markov switching multinomial logit model: an application to accident-injury severities. Accid. Anal. Prev. **41**(4), 829–838 (2009)
13. Ghazel, M., El-Koursi, E.-M.: Two-half-barrier level crossings versus four-half-barrier level crossings: a comparative risk analysis study. IEEE Trans. Intell. Transp. Syst. **15**(3), 1123–1133 (2014)
14. Chemweno, P., Pintelon, L., Van Horenbeek, A., Muchiri, P.: Development of a risk assessment selection methodology for asset maintenance decision making: an analytic network process (ANP) approach. Int. J. Prod. Econ. **170**, 663–676 (2015)
15. Hu, Y., Zhang, X., Ngai, E.W.T., Cai, R., Liu, M.: Software project risk analysis using Bayesian networks with causality constraints. Decis. Support Syst. **56**, 439–449 (2013)
16. Bouillaut, L., Francois, O., Dubois, S.: A Bayesian network to evaluate underground rails maintenance strategies in an automation context. Proc. Inst. Mech. Eng. Part O J. Risk Reliab. **227**(4), 411–424 (2013)
17. Langseth, H., Portinale, L.: Bayesian networks in reliability. Reliab. Eng. Syst. Saf. **92**, 92–108 (2007)

18. Weber, P., Medina-Oliva, G., Simon, C., Iung, B.: Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. Eng. Appl. Artif. Intell. **25**(4), 671–682 (2012)
19. Jensen, F.V.: An Introduction to Bayesian Networks, vol. 210. UCL Press, London (1996)
20. Pourret, O., Naim, P., Marcot, B.: Bayesian Networks: A Practical Guide to Applications. Wiley, Hoboken (2008)
21. SNCF: Research on the material of level crossing in 2014, France (2015)
22. Liang, C., Ghazel, M., Cazier, O., El Koursi, E.M.: Risk analysis on level crossings using a causal Bayesian network based approach. Transp. Res. Procedia **25**, 2172–2186 (2017)
23. SNCF Réseau: Statistical analysis of accidents at LXs, France (2010)
24. EN 50126: Railway applications-The specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS), British Standards Institution (1999)
25. Hanley, J.A., McNeil, B.J.: The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology **143**(1), 29–36 (1982)