COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING

Computer-Aided Civil and Infrastructure Engineering **00** (2016) 1–18

Bayesian Networks-Based Probabilistic Safety Analysis for Railway Lines

Enrique Castillo* & Zacarías Grande

Department of Applied Mathematics and Computational Sciences, University of Cantabria, Santander, Spain

&

Aida Calviño

Department of Statistic and Operations Research III, Complutensis University of Madrid, Spain

Abstract: *A Bayesian network model is developed, in which all the items or elements encountered when travelling a railway line, such as terrain, infrastructure, light signals, speed limit signs, curves, switches, tunnels, viaducts, rolling stock, and any other element related to its safety are reproduced. Due to the importance of human error in safety, especial attention is given to modeling the driver behavior variables and their time evolution. The sets of conditional probabilities of variables given their parents, which permits quantifying the Bayesian network joint probability, are given by means of closed formulas, which allow us to identify the particular contribution of each variable and facilitate a sensitivity analysis. The probabilities of incidents affecting safety are calculated so that a probabilistic safety assessment of the line can be done and its most critical elements can be identified and sorted by importance. This permits improving the line safety and saving time and money in the maintenance program by concentrating on the most critical elements. To reduce the complexity of the problem, an original method is given that permits dividing the Bayesian network in to small parts such that the complexity of the problem becomes linear in the number of items and subnetworks. This is crucial to deal with real lines in which the number of variables can be measured in thousands. In addition, when an accident occurs the Bayesian network allows us to identify its causes by means of a backward inference process. The case of the real Palencia–Santander line is commented on and some examples of how the model works are discussed.*

1 INTRODUCTION AND MOTIVATION

Because of the appearance of high-speed trains, transportation systems have undergone a great transformation. One of the key factors in this change is because high speed trains, due to their high speed features, can compete successfully with air transportation for distances ranging from 300 to 800 km (see Peterman et al., 2009; Todorovich and Hagler, 2011).

Another important change relates to safety requirements, which are becoming more strict not only for high-speed but for conventional lines too. Some advances that have arisen to improve safety levels are: (1) ATP (automatic train protection) systems, which make the existence of high-speed trains possible by helping drivers to avoid human errors and stopping trains when safety violations occur, (2) natural disaster early warning systems, which produce alarms when some risky situations arise (see, e.g., Veneziano and Papadimitriou, 2001), and (3) probabilistic safety analyses, which permit us to quantify the probabilities of undesired events, to improve safety by detecting events or sequences of events that lead to accidents and to force the necessary corrections (see Lahrech, 1999; Miyashita, 2010; Fukuyamaet al., 2008; Sussman, 1996).

Thus, modern railway line design involves not only a classical design in which safe operations are guaranteed, but a probabilistic safety analysis in which

*To whom correspondence should be addressed. E-mail: *castie@unican.es*.

DOI: 10.1111/mice.12195

all sequences of events leading to undesired events are identified and more important, their probabilities of occurrence are estimated and guaranteed to be below a small enough threshold value.

In this article, a Bayesian network model for probabilistic risk analysis (PRA) of railway lines is presented. Bayesian approaches have been used in Civil Engineering when randomness evaluation becomes necessary. For example, Wang et al. (2015), Sun and Betti (2015), and Yuen and Mu (2015) show some interesting applications of these techniques to various fields. In particular, Bayesian networks have shown to be a powerful technique to reproduce multivariate random variables with complex structure (see, e.g., Castillo et al., 2008; Spackova and Straub, 2013). We note that Bayesian networks are much more powerful than fault and event trees, especially for reproducing common causes of failures, which are very important in PRA studies.

PRA techniques are very common and have a long tradition in air transport and nuclear power industries, where powerful techniques for quantitative probabilistic risk assessments (see, e.g., Henley and Kumamoto, 1992), such as fault and event tree analysis, Petri and Bayesian networks (see Castillo et al., 1997b, 1999), or discrete and continuous Markov models are used. However, it must be recognized that this technique is not common in the railway industry and that progress in this direction is slow in many countries. Nevertheless, we can encounter some works, as those of Bearfield and Marsh (2005) and Flammini et al. (2006), where some of the reliability and safety analysis methods used in the nuclear field have been extrapolated to railway lines. Moreover, some countries, such as Japan, the United States, and the United Kingdom, make use of models such as the Safety Risk Model (SRM) of the Railway Safety Standard Board (RSSB), where fault and event tree analysis are used and have been adapted to the railway problem. SRM is an important tool in a series of those recently developed for railway safety analysis. The SRM provides very valuable risk information required in risk assessments. In addition, it allows us to understand the contribution of the different elements or failure modes to risk and to identify causes and consequences of potential accidents that arise in the regular and nonregular railway operation and maintenance. For the particular case of risk assessments of railway lines we must say that PRAs are not common but some exist (see Lahrech, 1999; Castillo et al., 2016) and that though SRM provides some important information for the assessment of such risk, it does not give risk profiles for specific lines.

Some examples of the risk assessment methodology in project management are given in Mokkapati et al. (2009) who provide a methodology for risk assessment, or in Kawakami (2014), who discusses the risk analysis of high-speed rail project managements in the United States. Unfortunately, in many countries, PRAs are rare, not compulsory, and are not regularly used to assess the safety of railway lines.

When performing probabilistic safety analyses, the use of guides is important to quantify the risks and consequences with some precision (see, e.g., Beales, 2002), but unfortunately in many cases only a qualitative risk assessment is done. In countries where these analyses are not common, other qualitative evaluations of railway safety are used, but neither detailed nor rigorous quantification of risks is carried out, so that the risk evaluation can be classified as purely qualitative in nature (see, e.g., Dirección de Seguridad en la Circulación, ADIF, 2009; Ministerio de Trabajo y Asuntos Sociales, 2010; Instituto Nacional de Seguridad e Higiene en el Trabajo, 1992). However, railway line designers should be aware that a qualitative evaluation of railway risks is not sufficient (see Masanori and Fumiaki, 2008).

Human error is probably the most important factor to be considered in any PRA but quantification of human error probabilities is probably one of the most difficult problems. However, this problem can be solved with the help of various groups of professionals (operators, conductors, PRA experts, statisticians, etc.) (see, e.g., Wreathall et al., 2003; Dadashi et al., 2013; Zeilstra and van del Weide, 2013). In this work we pay special attention to the driver attention and how it changes along the line as a function of the encountered elements and the help of ATP systems.

This and previous papers of the authors were motivated by some recent railway accidents which occurred in Spain and in the United States, where exceeding the speed limit in curves and lack of control of this human error was the main cause. An extra motivation comes from the appearance of the low cost and maintenance alternate double-single track lines that were suggested for low demand areas (see Castillo et al., 2011; Castillo et al., 2015b), which require a more careful risk analysis to be done, because of the presence of high-speed single tracks.

One important problem when performing a PRA is the parameter assessment, estimation, and calibration. They are crucial steps that must be done with care and with participation of various specialists if the probabilistic safety assessments are to be interpreted quantitatively. In this context, the works of Kokkings and Snyder (1997), Muttram (2002), and Evans (2011), who report important railway data and a serious statistical analysis of the railway accidents which occurred in Europe during the period 1980–2009, become relevant. However, once the models have been used, a sensitivity

analysis must be done to test the validity of the assumptions and how sensitive the results are to these assumptions (see Castillo et al., 1997a).

A new Bayesian network model was already used by Castillo et al. (2015a), where sequences of small Bayesian networks were used to permit a reasonable model size, that implies a simplification of the real dependencies among the variables. This model permits quantifying the risks of events and sequences of events to assess the safety of railway lines, and allows us to optimize the maintenance programs. However, it has some limitations and cannot be used for large real lines because of its complexity. The present model solves these problems and can be considered as one step further where important improvements have been made including a better representation of the variables involved and a method to deal with the high complexity of the models.

The main original contributions of this article are:

1. The Bayesian network method has the potential to catch mistakes in the signal system design and to quantify the associated probabilities of severe incidents.
2. The Bayesian network method provides a way to account for the incremental impacts of driver's error, given that the signal system has been properly designed (which other signal design practices cannot do) and to quantify the corresponding probabilities.
3. The partitioning procedure provides a way to make real-world assessments feasible and practical.
4. The application of the method to the Palencia–Santander real line with 709 items and 7,820 variables demonstrates the power of the method.

The model proposed in this article is valid for both high-speed and conventional lines and allows us to analyze existing or potential lines and to identify where some actions (corrections) are required. In Section 7, a real example of a conventional line is given.

The article is organized as follows. In Section 2, the Bayesian network variables are described in detail. In Section 3, the most relevant variables in our model: the driver's behavior, including driver's tiredness and driver's attention, and the speed variables are discussed. In Section 4, the components of our Bayesian network model are analyzed, that is, how different items contribute with a subnetwork and how the conditional probability tables are defined. In Section 5, how the Bayesian network can be partitioned into small subnetworks to reduce the complexity without affecting the final results is shown. In Section 6, the data and information to be supplied to the model are given. In Section 7, a real case is used to illustrate the power of the proposed methods. In Section 8, how a backward inference propagation can be done is illustrated. Finally, in Section 9, some conclusions are given.

2 VARIABLES AND ITEMS INVOLVED IN THE MODEL

In this section the variables and items used in our model are described.

2.1 Variables involved in the model

One important step in the building process of a Bayesian network is the identification of variables to be modelled. In the proposed model, the following variables are assumed to be relevant from a safety point of view:

1. *A: Incident.* It refers to the incident occurrence at the actual location and can take the following values: none, minor, medium, and severe.
2. *ATP: Automatic Train Protection System*. This variable refers to the supervising or driving assistance system operating at the considered point of the line. It takes the values: "ERTMS," "ERTMS–ASFA," "ASFA–dig," "ASFA–AV," "ASFA–Conv," "ASFA–anal," "SR" (staff responsible).
3. *AS: Light signal decision*. It refers to the possible driver's decisions at a light signal: correct, Error I (stop announcement signal), Error II (signal at red).
4. a_t*: Driver's tiredness.* Because the driver's tiredness has an important contribution to human error and increases with driving time, a variable is needed to analyze how it changes when travelling along the line.
5. *D: Driver's attention*. It refers to the driver's attention level. Because this variable cannot be measured, in our model it is simplified to three states: distracted, attentive, and alert. By *distracted* we understand a situation in which the driver lacks the necessary attention to correctly react when an action is required and leads to a no action at all. By *attentive* we refer to the case in which the driver is able to react adequately to the required actions with a small probability of error. Finally, *alert* refers to the case where the driver is ready to take an action and knows that he/she has to act immediately (e.g., after seeing a warning signal or consulting the railway driver's guide, etc.). We assume that the alert situation always leads to

a correct decision. Consequently, any driver's erroneous decision must be included in the attentive level.

6. *DA: Driver's decision at signal*. It refers to driver's decisions when the train encounters a signal, and includes the following levels: correct, error (incorrect action of the driver).
7. *DE: Driver's decision on speed control.* It refers to the decision made by the driver when speed must be controlled, and includes the following levels: correct, Error I (speed remains unchanged when it must be changed), Error II (selected speed does not coincide with required speed).
8. *DS: Driver's decision made at a speed limit sign*. It refers to the possible decisions made by the driver at a speed limit signal: correct or Error I (fail to reduce speed).
9. *Inf: Infrastructure*. With this variable the infrastructure state (rails, sleepers, ballast, plate, maintenance standards, etc.) is considered, and it includes the damage levels: none, minor, medium, and severe. Note that knowledge of the infrastructure state is needed to determine the probabilities of related incidents.
10. *RS: Rolling stock*. It refers to the rolling stock conditions and includes the damage levels: none, minor, medium, and severe. Similarly, knowledge of the rolling stock state is needed to determine the probabilities of related incidents.
11. *S: Speed*. It refers to the train speed at the corresponding location and can take a discrete list V of values, which in this article is simplified to a set V starting from 0 and ending with 280 km/hour with increments of 20 km/hour. However, if at given locations, some particular values are of interest, they will replace the closest values in the list.
12. *SS: Light signal state*. It refers to the light signal state: free, stop announcement or stop.
13. *T: Terrain.* This variable is used to consider the risk associated with falling stones on the infrastructure or slope slidings in cuttings and embankments and takes values: stable, small, medium, and high instability.
14. *TF: Technical failure.* It refers to the possibility of a technical failure: yes or no. For example, a brake failure.

Nowadays trains are controlled by train controllers (operators, dispatchers) in operation control centers and interlocking towers. For the sake of simplicity we do not include in this article the train control variables, but this can be done in a form analogous to the rest of variables.

2.2 Items analyzed in the model

By item we refer to an element that exists somewhere or appears along the line and has influence on its safety, such as grade crossings, curves, viaducts, tunnels, light signals, driver booklets, end stations, speed limit signs, switches, ATP systems, warning signs, parameter changes, start station, under and overpasses, and especial locations where a danger exists. In Table 1, we give a list of these items together with their graphical representations (to be used in the figures of this article).

3 RELEVANT VARIABLES

Due to its relevance in safety, we devote this section to the two most relevant variables in the network: the driver's behavior and the speed variable. Other variables are treated similarly.

3.1 Driver's behavior

We devote this section to explain how the driver's behavior is reproduced in our model, including two subsections dedicated to driver's tiredness and driver's attention.

3.1.1 Driver's tiredness variable. Though the Bayesian model normally considers proper random variables it allows us to select some variables as deterministic (a particular case of random variable). This makes sense when the variable has small variations and its random variation has no great effect on the network safety, or when its randomness can be included in other model parameters, as it is the case.

For the sake of simplicity, we have assumed an amplifying factor $a_t \geq 1$ f or the probability of making wrong decisions due to tiredness that changes with time as follows:

$$a_t = e^{-\delta t^2}, t \geq 0 \tag{1}$$

where δ is a parameter, which in our examples we have assumed to be $\delta = 0.02$ hour^{-2}. This exponential decay curve has been selected because it is increasing and with a positive second derivative, two properties that this function must have, that is, increase with tiredness and at a larger rate as time increases. However, further research is needed to justify this type of curve and to have a reasonable estimate of its parameter δ in practical cases. This means that the probability of making wrong decisions by the driver after a driving time t must be increased (multiplied) due to tiredness by the corresponding amplifying factor a_t.

Table 1
Frequency of the different items in the Palencia–Santander line

Item	Frequency	Plot	Item	Frequency	Plot
AnnouncementGradeCrossing	43		SignalP (permanent)	26	
GradeCrossing	56		SignalFP (final permanent)	25	
CurveIn	22		PreannouncementT	0	
CurveOut	22		AnnouncementT	14	
ViaductIn	13		SignalT (temporal)	14	
ViaductOut	13		SignalFT (final temporal)	13	
TunnelIn	30		AnnouncementSwitch	0	
TunnelOut	30		Switch	23	
SignalA (advanced)	24		Supervisor	0	
SignalC	36		Screen1	20	
SignalE (entry)	24		Screen2	20	
DriverBookletAnnouncement	10		Screen3	20	
Driverbooklet	22		Continuous	7	
StopStationAnnouncement	1		ContinuousOFF	7	
StopStation	1		Start	1	
AnnouncementP	25				

The driver's tiredness amplifying factor ranges from 1 (no tiredness) to ∞ (unlimited tiredness) and t is given in driving hours. In practice, its range of application is between $t = 0$ and a few hours (e.g., $t = 6$ hours), because due to safety reasons the continuous and daily driving time is limited. We have considered that after driving six hours the amplifying factor takes value 2, but this needs to be confirmed.

3.1.2 M variable for driver's attention. To reproduce the driver behavior in a segment (a fraction of the railway line) without signals and signs (no light, speed limit, acoustic or other signals or signs are present), we use a continuous Markov process (see references Benjamin and Cornell, 1970; Doob, 1953; Kijima, 1997) and the M node, which refers to the driver's attention. The Markov assumption seems to be a simplification and a reasonable assumption, because what is important with respect to the driver behavior is whether or not the driver is distracted, attentive, or alert, but not how he/she has attained such state.

Apart from the first node M without parents, all other M nodes have only one parent, which is another M node (a subindex p is used to refer to previous nodes of the same type. For example, M_p means the previous driver's attention node M).

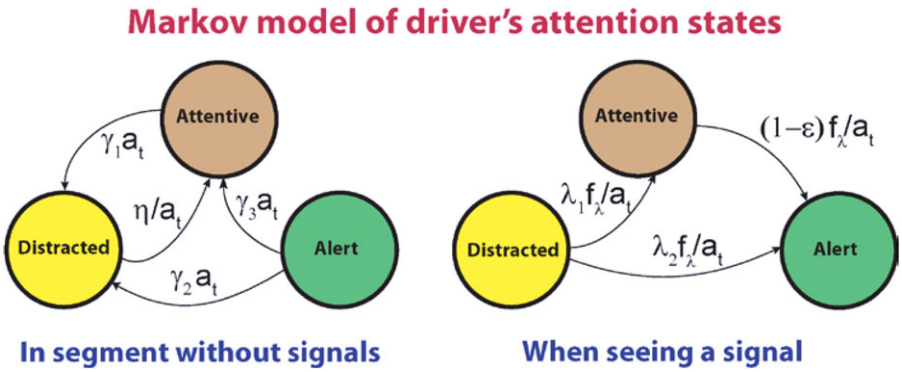


Fig. 1. Markov model illustrating the transitions among different attention levels in segments without signals (left figure) and when visualizing a signal or sign (right figure).

In this case we have the differential equation associated with a standard Markov model (see the left graph in Figure 1):

$$\begin{pmatrix} p_0'(t^*, t-t_0) \\ p_1'(t^*, t-t_0) \\ p_2'(t^*, t-t_0) \end{pmatrix} = M \begin{pmatrix} p_0(t^*, t-t_0) \\ p_1(t^*, t-t_0) \\ p_2(t^*, t-t_0) \end{pmatrix} \quad (2)$$

where

$$M = \begin{pmatrix} -\eta/a_{t^*} & \gamma_1 a_{t^*} & \gamma_2 a_{t^*} \\ \eta/a_{t^*} & -\gamma_1 a_{t^*} & \gamma_3 a_{t^*} \\ 0 & 0 & (\gamma_2 + \gamma_3) a_{t^*} \end{pmatrix} \quad (3)$$

where we have used the notation M for this matrix because this is a matrix (not a transition matrix) associated with the driver attention variable M and $p_0(t^*, t)$, $p_1(t^*, t)$, and $p_2(t^*, t)$ are the probabilities associated with the three driver's states (distracted, attentive, and alert), t, t_0, and t^* are a given time, the segment starting time, and the train passing time associated with the segment center, respectively, and η/a_{t^*} is the time rate of recovering attention when the driver is distracted after a time t^* from the beginning of the trip, $\gamma_i a_{t^*}, i = 1, 2$ are the time rates of becoming distracted when being attentive and alert, respectively, and $\gamma_3 a_{t^*}$ is the time rate of becoming attentive when the driver is alert, which leads to the conditional probabilities $p(a, b) = P(M = a|M_p = b)$ for node M given node M_p given by:

$$\begin{pmatrix} \delta_{a,1} \\ \delta_{a,2} \\ \delta_{a,3} \end{pmatrix}^T M_l(t^*, t_{\text{end}} - t_0; \eta, \gamma_1, \gamma_2, \gamma_3) \begin{pmatrix} \delta_{b,1} \\ \delta_{b,2} \\ \delta_{b,3} \end{pmatrix} \quad (4)$$

where t_{end} is the segment exit time and the deltas are the Kronecker's deltas. They permit us to isolate the elements in matrix M_l.

In the case of the driver seeing a warning signal, the Markov model leads to the conditional probabilities:

$$P(M = a|M_p = b) = \begin{pmatrix} \delta_{a,1} \\ \delta_{a,2} \\ \delta_{a,3} \end{pmatrix}^t M_1 \begin{pmatrix} \delta_{b,1} \\ \delta_{b,2} \\ \delta_{b,3} \end{pmatrix} \quad (5)$$

where t refers to transpose,

$$M_1 = \begin{pmatrix} 1 - (\lambda_1 + \lambda_2) f_\lambda/a_t & 0 & 0 \\ \lambda_1 f_\lambda/a_t & 1 - (1-\varepsilon) f_\lambda/a_t & 0 \\ \lambda_2 f_\lambda/a_t & (1-\varepsilon) f_\lambda/a_t & 1 \end{pmatrix} \quad (6)$$

and, as shown in the right graph in Figure 1, $0 \leq \lambda_1, \lambda_2 \leq 1$ are the probabilities of recovering the attentive or alert levels of attention, respectively, when seeing the signal if the driver is distracted, $0 \leq \varepsilon \leq 1$ is the probability of the driver being unaware of the signal presence when circulating under an attentive attention level and $0 \leq f_\lambda \leq 1$ is a factor that takes into account the cabin warning sounds and others aspects, and takes value 1 when in the most favorable case. Note that, in addition, the reaction probabilities have been reduced using a reduction factor a_t due to tiredness.

The probability ε is assumed very small. However, the value of λ_1 and λ_2, are much higher than ε, that is, $\lambda_1 >> \varepsilon$.

Equation (5) permits updating the probabilities of the driver's attention states when passing a signal or sign. Given the transition matrix structure, it is clear that this always produces a statistical improvement of driver's attention.

3.2 The speed V node

In this section we describe how the speeds are treated in our model. We need to understand that speeds are crucial to safety. Thus, a detailed analysis of how speeds must be controlled is necessary.

Assume that the train is at a given location (the actual location) and that it circulates at a speed v_p. Assume also that the target speed at the next location, at a distance l from the actual one, is v_{target}. It is clear that this target speed will be attainable or not attainable depending on the actual speed v_p, the distance l between the two locations (actual and next), and the possible maximum acceleration $a_{\max} > 0$ and deceleration $a_{\min} < 0$, respectively. Thus, we need to calculate first what are the minimum and maximum attainable speeds after this distance l assuming the actual speed v_p. If the target speed is attainable, it will be the final speed. Otherwise the final speed v will be below or above the target speed v_{target} depending on the initial speed v_p. Taking all this into account, we have that the final attainable speed value v_{at} (v_p, $v_{\text{target}}la_{\min}$, $a_{\max}$) is given by (see the top plot in Figure 2):

$$v_{\text{at}} \begin{cases} \sqrt{v_p^2 + 2a_{\max}l} & \text{if } v_{\text{target}} > \sqrt{v_p^2 + 2a_{\max}l} \\ \sqrt{v_p^2 + 2a_{\min}l} & \text{if } v_{\text{target}} < \sqrt{v_p^2 + 2a_{\min}l} \\ v_{\text{target}} & \text{otherwise} \end{cases}$$

Thus, given the previous speed v_p of the train, the final speed when the target speed is v_{target} will be given by the thick three-piece line in the top graph of Figure 2, where the horizontal dashed line corresponds to $v = v_{\text{target}}$. The three plots in Figure 2 show how the attainable speeds change when the initial speed is v_p and the target speed v_{target} in a light signal at green, at stop announcement (orange), and at stop (red), respectively.

Because this attainable speed v_{at} is relevant to this location, it must be located in the list V of discrete speed values. To this end, we find the index a of the closest nonzero value v_a to v_{at} in the list V and replace this value by v_{at}. Let's denote $a = h(v_{\text{at}})$ to the indicated index value a in list V associated with v_{at}. We note that v_a should be the speed to be attained by the driver when driver's action is correct. However, we must also consider that this action can be erroneous. In this case we simplify and consider two types of error: "Error I" and "Error II." "Error I" means that the driver does not change the speed as required and then, $v = v_p$. To simplify the case of "Error II," in the model we have assumed that the driver instead of v_a can select a speed smaller $v_{\max(0,\, a-1)}$ or larger $v_{\min(n,\, a+1)}$ than the one required, with probabilities κ_1 and κ_2, respectively, where n is the number of possible values of the speed variable v.

Once we know what the attainable speeds are, we can analyze different cases of conditional probabilities, as shown in the following sections.

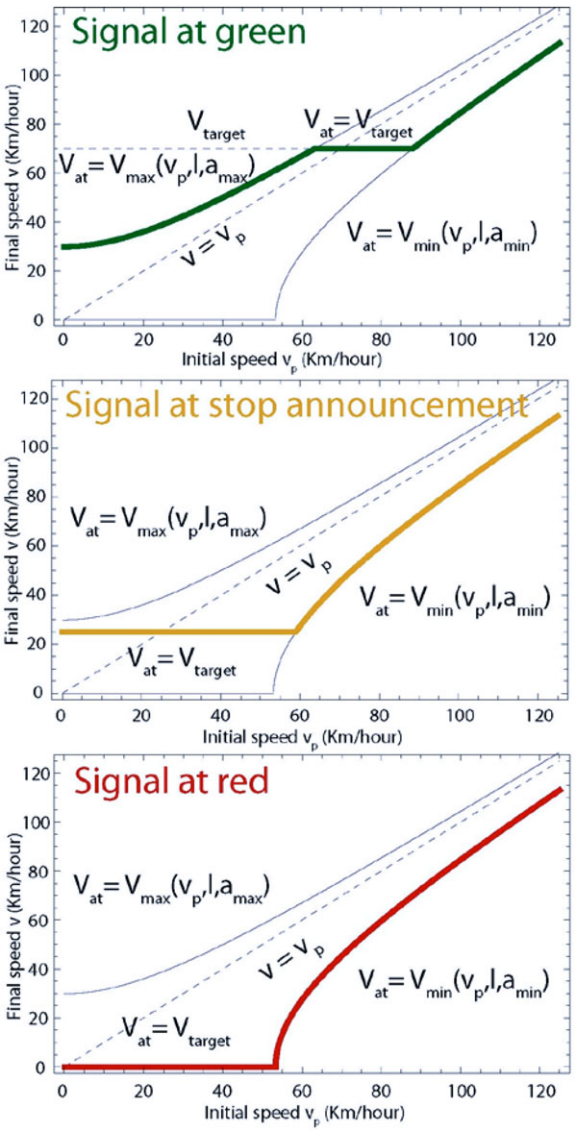


Fig. 2. Illustration of the attainable speed when the initial speed is v_p and the target speed is v_{target} in the cases of light signals at green, at stop announcement (orange), and at stop (red).

4 COMPONENTS OF THE BAYESIAN NETWORK

The structure of the Bayesian network depends on the railway line being studied and tries to reproduce all the elements the driver encounters when travelling along the line.

A Bayesian network has two elements: an acyclic graph and a set of conditional probabilities. The nodes of the graph reproduce the variables and its links reproduce the direct statistical dependencies or relations (causal or not) among them. Each node (son) is connected with its parent. The graph alone permits understanding the qualitative relations among the

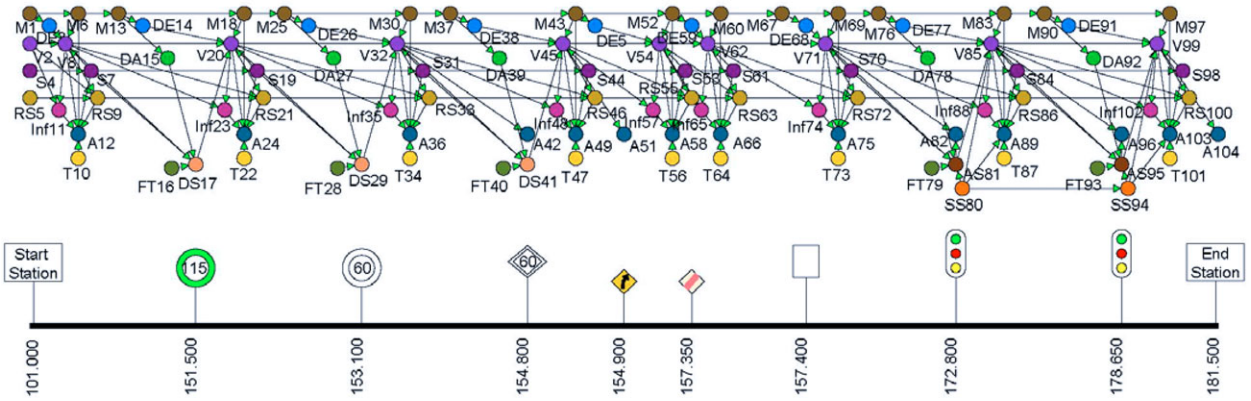


Fig. 3. Proposed Bayesian network. Illustration of the Bayesian model showing the subnetworks for the departure station, the segments without signals, the speed limit signals, curves, light signals, and the end station.

variables (see Figure 3). To complete the definition of the Bayesian network (see Castillo et al., 1997b) tables of conditional probabilities *sons* conditioned to their *parents* need to be incorporated. This allows us to incorporate the quantitative information of the Bayesian network.

In our case, there are some especial variables that must be monitored all along the line. They are the driver's attention, the running speed, the ATP and the rolling stock variables, which are associated with different locations. Analysis of these variables in sequence permits following their evolution along the line (see Figure 3).

In summary, the model is a large Bayesian network in which a subnetwork, made of a group of nodes (variables), is incorporated when a new item is encountered, such as: (a) segments without signals, whose parameters depend on the train and segment characteristics (rolling stock state, tracks, cuttings, embankments, etc.), (b) locations where driver's attention is improved (warning signs, acoustic signs, etc.), (c) concentrated risks (switches and viaducts or tunnel entries and exits), (d) locations where some decision subject to error must be made (e.g., light signals), etc. These items are sorted in the line start to end direction and their nodes are connected to reproduce the statistical dependencies of the variables (see Figure 3).

In this section we describe some of the subnetworks generated by the different items, but due to space limitations, we describe only four of them: the segment without signals, the speed limit sign, the entry light signal, and the curve subnetworks. All other subnetworks are similar to these ones.

Apart from the dependence graph, we need to define the conditional probability of each node given its parents. We indicate below how this can be done for some of them.

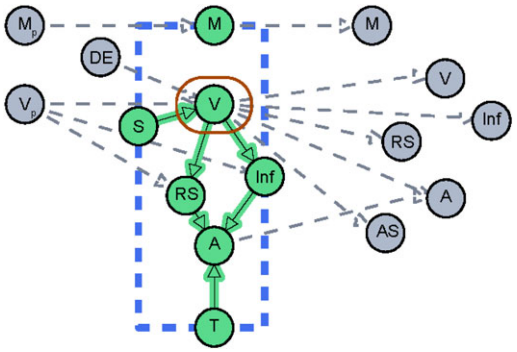


Fig. 4. Illustration of a segment without signals subnetwork showing its seven nodes $\{M, S, V, RS, T, Inf, A\}$ and links together with the dashed links connecting this subnetwork to the previous and the following ones.

4.1 Segment without signals subnetwork

A subnetwork to reproduce a railway line segment without signals can incorporate four different sets of nodes depending on what its subnetworks neighbors are. Due to lack of space, in this article we comment on only the subnetwork with nodes $\{M, S, V, RS, T, Inf, A\}$, which is shown in Figure 4, where the dependencies among the different variables are shown using arrows (links). For example, incidents (node A) can be due only to failures in the rolling stock RS, infrastructure Inf or terrain T. In this case, the subnetwork aims at describing how the driver's attention M, the speed V, the rolling stock RS, the ATP system S, and the incident occurrence A interact and evolve with time (or location) when the train circulates in segments without signals. In this case, the driver's attention changes with travelled length, the driver makes decisions that can be correct and incorrect DE, these decisions can be corrected or not by the ATP system S and derailments or accidents can occur due to failures in the rolling stock RS, the infrastructure

Inf or terrain (slope slides in cuttings or embankments, settlements, etc.) T, or other events jeopardizing safety. Thus, incidences at this type of segment must be taken into consideration. Given the especial character of these events, we can assume a given failure rate per unit length or per unit time (both rates are related by a speed factor). This rate must include all the above-mentioned risk factors and will depend on the characteristics of the particular segment being analyzed, on maintenance quality, weather conditions, etc.

4.1.1 Speed control. V node with three parents V_p, DE, and S. The V node can have different number of nodes depending on the case. In this case, shown in Figure 4, we consider only the case of three nodes: previous speed v_p, driver's decision DE, and supervisor system S (ATP) and the corresponding conditional probability $P(V|V_p, DE, S)$ has elements $p_{a,b,c,d}(s) = P(V = a|V_p = b, DE = c, S = d)$ whose values are:

$$\begin{aligned} p_{a,b,c,d}(s) = {} & \delta_{c,1}\delta_{a,s} + \delta_{c,2}\left((1-\rho_d)\delta_{a,s} + \rho_d\delta_{a,b}\right) \\ & + \delta_{c,3}\left(\kappa_1\rho_d\delta_{a,\max(1,s-1)} + (1-\rho_d(\kappa_1+\kappa_2))\delta_{a,s}\right. \\ & \left. + \kappa_2\rho_d\delta_{a,\min(n,s+1)}\right) \end{aligned} \quad (7)$$

where ρ_d is the failure probability of the ATP system and the argument s has been used to indicate that the conditional probability table depends on the speed level $s = h(v_{at})$. A detailed description of the terms in Equation (7) is given below.

The three terms in Equation (7) correspond to the cases of *correct decision* (term with factor $\delta_{c,1}$), in which case the driver will select as target speed the correct one (term $\delta_{a,s}$), which corresponds to value v_s in the list V, to *Error I* (term with factor $\delta_{c,2}$), in which case the driver will omit actions and unless the ATP system corrects this and selects the correct speed v_s (term $(1 - \rho_d)\ \delta_{a,s}$), the driver will do nothing (term $\rho_d\delta_{a,s}$), and *Error II* (term with factor $\delta_{c,3}$) in which case the driver will try to attain the speed v_s but with probabilities k_1 and k_2 the driver will select speeds $v_{\max(0,s-1)}$ (term with $d_{a,\max(1,s-1)}$) or $v_{\min(n,\ s+1)}$ (term with $d_{a,\min(n,\ s+1)}$), respectively, unless an ATP system correction takes place (term ρ_d). Note that the Kronecker's deltas permit ignoring or considering the different terms depending on the case being considered.

One important original contribution of this article is that the conditional probabilities of the proposed model are given as closed formulas, such as (7). These formulas permit their quick and safe implementation into a computer program and provide a very clear representation of the conditional probabilities.

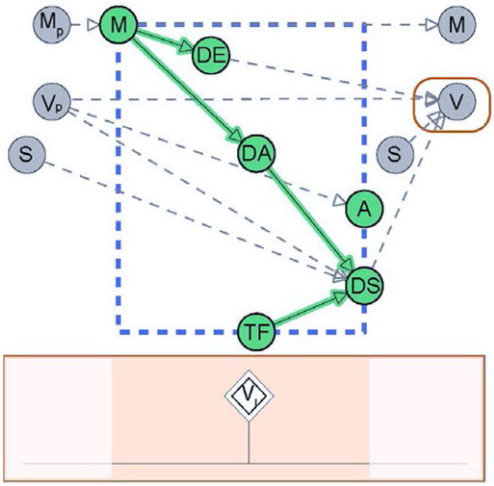


Fig. 5. Illustration of a speed limit sign Bayesian subnetwork showing its six nodes $\{M, DE, DA, TF, DS, A\}$ and links together with the dashed links connecting this subnetwork to the previous and the following ones.

4.2 Speed limit signal subnetwork

Each speed limit sign contributes to the general Bayesian network with a subnetwork with six nodes $\{M, DE, DA, TF, DS, A\}$, as shown in Figure 5, where the subnetwork links together with other links connecting to the subnetwork neighbors are shown.

In this case, the driver changes attention M when seeing the signal, and must make a decision on the speed. This decision implies some action DA that can be altered by a technical failure TF, leading to a final action DS on the target speed, which modifies the speed forward; in this case the incident A is related with the driving speed in that point v_p. The node and link structure in Figure 5 reproduces this behavior.

4.2.1 Speed limit signals. V node with four parents and V_P, DE, DS, and S. In the second type of V node with four parents, we need the conditional probability $P(V|V_P, DE, DS, S)$ which is similar to $P(V|V_P, DE, S)$, but with node DS added (see Figure 5). The DS node can take value *correct*, in which case the conditional probability becomes equal to $P(V|V_p, DE, S)$, given in Equation (7), or take value *error*, in which case the conditional probability of this case is equal to the second term in (7).

Additionally, as shown in Figure 6, the v_{target} depends on the speed limit signals, as indicated.

It must be noted that the type of limit signal (*permanent* or *temporary*), does not modify the v_{target} calculation.

$$v_{target} = \begin{cases} \sqrt{v_0^2 + \dfrac{v_p^2 - v_0^2}{(s_a - s_p)}(s - s_p)} & \text{if} \quad s \in [s_p, s_a) \\ \sqrt{v_p^2 + \dfrac{v_p^2 - v_l^2}{(s_l - s_a)}(s - s_l)} & \text{if} \quad s \in [s_a, s_l) \\ v_l & \text{if} \quad s \in [s_l, s_{fl}) \end{cases}$$

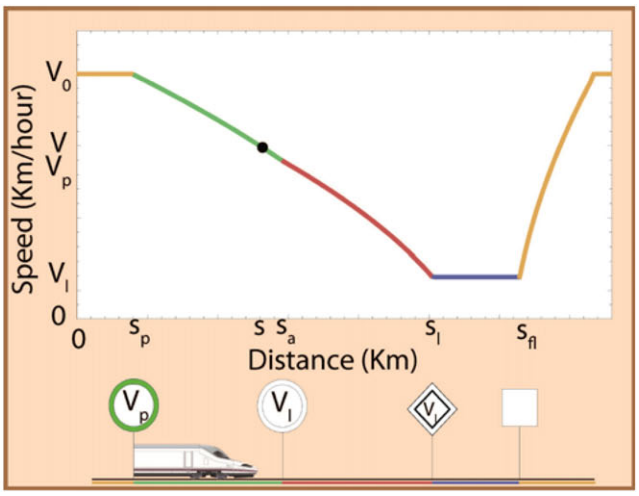


Fig. 6. Illustration of the speed trend between speed limit signs, where v_0, v_p, v_l, are the train speeds at the first light signal, the preannouncement, and the mandatory speed limit, respectively and s_p, s_a, s_l, s_{fl} are the distances to the preannouncement, announcement, mandatory, and end of speed limit sign, respectively.

Once v_{target} has been calculated, the attainable speed v_{at} (depicted in Figure 2) is obtained and then $s = h(v_{\text{at}})$ must be evaluated.

If an item between the signals exists, the target speed associated with the item location must be calculated according to the curves in the figure, depending on the previous and next limit signs.

Thus, the conditional probability $P(V = a|\ V_p = b, DE = c, DS = d, S = e)$ becomes (in closed form):

$$\begin{aligned} p_{a,b,c,d,e}(s) = \ & \delta_{d,1}[\delta_{c,1}\delta_{a,s} + \delta_{c,2}((1-\rho_e)\delta_{a,s} + \ \rho_e\delta_{a,b}) \\ & + \delta_{c,3}\ (\kappa_1\rho_e\delta_{a,\max(1,s-1)} + (1 - \rho_e(\kappa_1 + \kappa_2))\delta_{a,s} \\ & + \kappa_2\rho_e\delta_{a,\min(n,s+1)})] \\ & + \delta_{d,2}((1-\rho_e)\delta_{a,s} + \ \rho_e\delta_{a,b}) \end{aligned} \tag{8}$$

This conditional probability considers node DS, that is, a correct or erroneous decision resulting after the intervention of the driver, the ATP system and the possibility of a technical failure. The first and second terms provide the conditional probabilities in the cases of a correct or erroneous decision, respectively.

4.3 Entry light signal subnetwork

Each entry light signal contributes to the general Bayesian network with a subnetwork with seven nodes $\{M, DE, DA, TF, SS, AS, A\}$, where the subnetwork links together with other links connecting to the subnetwork neighbors.

In this case, the driver changes attention M when seeing the signal, and must make a decision on the speed. This decision implies some action DA that can be altered by a technical failure TF, leading to a final action AS on the target speed, which can lead to an accident A.

4.3.1 Light signals. V node with four parents and V_p, DE, S, and SS. In this case, the node V is connected to the four nodes: previous speed V_p, driver's decision DE, supervisor S, and light signal state SS, and the corresponding conditional probability becomes $P(V|V_p, DE, S, SS)$.

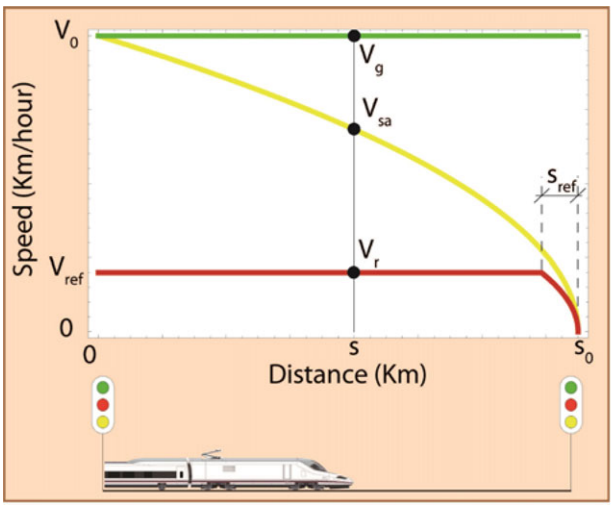
Consequently, we need to define the elements $p_{a,b,c,d,e} = P(V = a|V_p = b, DE = c, S = d, SS = e)$ of the conditional probability table. This case is similar to the case of three parents $P(V = a|V_p = b, DE = c, S = d)$, but now we have added a new variable SS. The SS node value e gives the light signal state, which can take values in the list {*free*, *stop; announcement*, *stop*}, so the signal state will modify the attainable speed v_{target} as follows:

$$v^*_{\text{target}}(e) = \delta_{e,1}v_g + \delta_{e,2}v_{sa} + \delta_{e,3}v_s \tag{9}$$

where v^*_{target} is the attainable speed, and v_g, v_{sa}, and v_s are the target speeds when the signals are at green, at stop announcement, or at stop, respectively.

Figure 7 illustrates the speeds between two light signals depending on the first signal state and the corresponding equations, where v_0 is the train speed at the first light signal, v_{ref} is the red trespassing speed, s_0 is the distance between the two light signals, s_{ref} is the required distance to stop the train when its speed is v_{ref}.

If there exists an item between the two signals, the target speed associated with the item location must be calculated according to the curves in the figure, depending on the first signal state, which can be "at green," "at



1) Light signal at green:

$$v_g = v_0 \quad \text{if } 0 \le s \le s_0.$$

2) Light signal at stop announcement:

$$v_{sa} = \sqrt{v_0^2\left(1 - \frac{s}{s_0}\right)} \quad \text{if } 0 \le s \le s_0.$$

3) Light signal at surpassable red: $v_r =$

$$\begin{cases} v_{ref} & \text{if} \quad s \le s_0 - \dfrac{v_{ref}^2}{2a_{min}} \\ \sqrt{v_{ref}^2 - 2a_{min}(s - s_{ref})} & \text{if} \quad s > s_0 - \dfrac{v_{ref}^2}{2a_{min}} \end{cases}$$

Fig. 7. Illustration of the speed profile between two light signals, depending on the first signal state, which can be "at green," "at stop announcement," or "at surpassable red" (green, yellow, and red-surpassable, respectively).

stop announcement," or "at surpassable red" (green, yellow, and red-surpassable, respectively).

Once v^*_{target} has been calculated in (9), we must calculate the attainable speed v_{at} and then evaluate $s = h(v_{\text{at}})$ and use (7) to obtain the conditional probability $p_{a,b,c,d,e}(s)$.

4.4 Curve sub-Bayesian network

The curve subnetwork contains only nodes A and V, that is, the incident (derailment at the curve) level depends only on speed. Obviously, we have assumed in our model that the curve radius has an effect on the probability of derailment.

A node: Incident at a curve. This node has one parent V variable and its conditional probability can be written as

$$P(A = a|V = b) = \delta_{a,1} q_1^s(b) + g(a,b)\, q_2^s(b) \quad (10)$$

where

$$\begin{aligned} g(a,b) = {} & \delta_{a,1}\left(1 - F_{N(1.2v_{\text{lim}}, v_{\text{lim}}/10)}(v(b))\right) \\ & + \delta_{a,2}\left(F_{N(1.2v_{\text{lim}}, v_{\text{lim}}/10)}(v(b)) \right. \\ & \left. - F_{N(1.35v_{\text{lim}}, v_{\text{lim}}/10)}(v(b))\right) \\ & + \delta_{a,3}\left(F_{N(1.35v_{\text{lim}}, v_{\text{lim}}/10)}(v(b)) \right. \\ & \left. - F_{N(1.5v_{\text{lim}}, v_{\text{lim}}/10)}(v(b))\right) \\ & + \delta_{a,4}\left(F_{N(1.5v_{\text{lim}}, v_{\text{lim}}/10)}(v(b))\right) \end{aligned} \quad (11)$$

is the maximum speed which does not produce an accident, that is, is a function that gives the probability of the incident level to be a when the speed level is b, the subindices of the $F_{N(\mu,\sigma)}$ functions refer to the normal

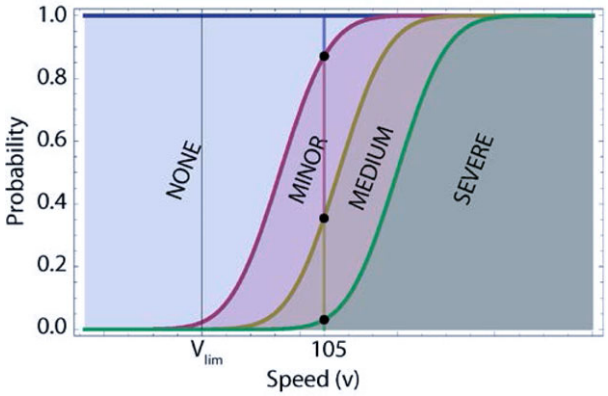


Fig. 8. Illustration of the $g(a, b)$ function, showing how the probabilities associated with the none, minor, medium, and severe incidents depend on the speed v, and the v_{lim}.

distributions used in Figure 8 and v_{lim} is the maximum speed which does not produce an accident, that is,

$$\frac{V_{\text{lim}}}{V_0} = 1 + \alpha_{11}\frac{R}{R_0} + \alpha_{12}\frac{R}{R_0}^2 \quad (12)$$

where v_0 and R_0 are reference values for the speed and the curve radius, respectively, α_{11} and α_{12} are the model parameters, and $q^s{}_i(b)$ are the elements of the $Q^s(b)$ matrix whose two elements refer to the fact that the speed limit is already satisfied at the signal location or to the fact that the speed limit is attainable, respectively.

$$Q^s(b) = (v(b) \le v_{\text{lim}}\, v(b) > v_{\text{lim}}) \quad (13)$$

Figure 8 illustrates the $g(a, b)$ function, showing how the probabilities associated with the none, minor, medium, and severe incidents depend on the speed v,

and the $v_{\lim}$. For example, for a given speed, say $v =$ 105 km/h, the probabilities correspond to the lengths of the vertical segments of the intersection with the vertical line and the corresponding regions (see the vertical line in Figure 8). We note that this way of defining the conditional probabilities is an original contribution.

5 NETWORK PARTITION

When dealing with real lines, the proposed Bayesian network model leads to a very high number of variables. For example, the case of the Palencia–Santander line, which was treated in Section 9 of this article, contains 7,820 variables, for which memory and CPU problems are expected if conventional Bayesian network packages are used. Thus, something must be done to solve this problem.

To reduce memory and CPU requirements and complexity, the partition of the Bayesian network into a sequence of several subnetworks, as small as possible, without altering the results is convenient. This is the aim of the following idea.

Let $\{B_1, B_2, \ldots, B_n\}$ be one of such subnetwork sequences and let $\{B_k^1, B_k^2\}$ be a partition of the nodes in subnetwork B_k. The most convenient partitions are those such that the nodes of B_{k+1} are independent of the nodes $\{B_1, B_2, \ldots, B_{k-1}$ and $B_k^1\}$ given the nodes of B_k^2. This implies that all the information that the nodes in $B_1, B_2, \ldots, B_{k-1}$ and B_k^1 have on the nodes in B_{k+1} is already contained in the nodes of B_k^2. We call the set of nodes in B_k^2. *a separator* of the initial Bayesian network.

One example is given in Figure 9, where a Bayesian network associated with a piece of a railway line is given. It corresponds to the signals or signs indicated in the lower part of the figure. Note that the separator nodes have been duplicated because they must belong to both subnetworks. The head of the node name refers to the node type and the tail refers to the type of item (*s* for signal, VO for viaduct out, etc.).

In the middle plot, the corresponding Bayesian network is shown with all its nodes and links. In particular, a partition B_1, B_2, B_3, B_4, B_5 contain nodes 13–21, 22–35, 36–45, 46–59, and 60–61, respectively.

The upper plot corresponds to the five subnetworks B^*_1, B^*_2, B^*_3, B^*_4, B^*_5 associated with the partition. Note that four artificial nodes (duplicated from their previous subnetworks) have been added to the second subnetwork and five artificial nodes (duplicated from their previous subnetworks) have been added to the third to fifth subnetworks, but no addition has been done to the first one. The duplicated nodes are indicated by using the same colors.

It can be easily shown that Bayesian subnetwork B_{k+1} is independent of the nodes in Bayesian network B_1, $B_2, \ldots, B_{k-1}$ without the nodes in the separator, given the nodes in the separator, for $k = 1, 2, \ldots, n-1$ (see Castillo et al., 1997b).

The selected partition is not arbitrary at all. The key property for a partition to be valid is to contain a set of separators (subsets of nodes) such that the conditional probability of the posterior nodes becomes independent on the previous nodes given the separator subset. Consequently, the separator subset and the partitions have been selected to satisfy this condition.

In addition, some artificial links have been added to the separator for it to become a clique.

Using this partition and its important independence properties, the marginal densities of the nodes in the initial Bayesian network can be obtained using the following process:

1. Step 1. *Initialize counter*. Let $k = 1$ (first partition).
2. Step 2. *Build the Bayesian network* B_k. The Bayesian network B_k is built based on its nodes, links, and probability tables (conditional probabilities of the sons given their parents). To this end a computer program in Matlab developed by the authors automatically builds the Bayesian network acyclic graph from the list of items encountered when travelling the railway line. Similarly, the closed form formulas of each case are used by the Matlab program to build the conditional probability tables automatically.
3. Step 3. *Obtain the marginal densities of the nodes.* The marginal densities of the nodes in Bayesian subnetwork B_k are obtained using the usual inference methods in Bayesian networks. To this end, we have used the JavaBayes and BNT software packages.
4. Step 4. *Check for last partition*. If k is the last partition, stop. Otherwise continue.
5. Step 5. *Obtain the joint density of the separator*. The joint density of the separator B_k^2 is obtained using some standard methods for Bayesian networks. Because the separator has been forced to be a clique artificially, its joint density function can be easily obtained. This is not true for joint densities of nodes not in a clique.
   Note that in Figure 9 some artificial links (*M–V*, *M–S*, and *M–RS*) have been added to the first partition to obtain a clique permitting to evaluate the marginal probabilities of (*M*, *S*, *V*, *RS*) without a high computational cost.
6. Step 6. *Obtain the conditional probabilities of the nodes in the separator set.* For the separator with

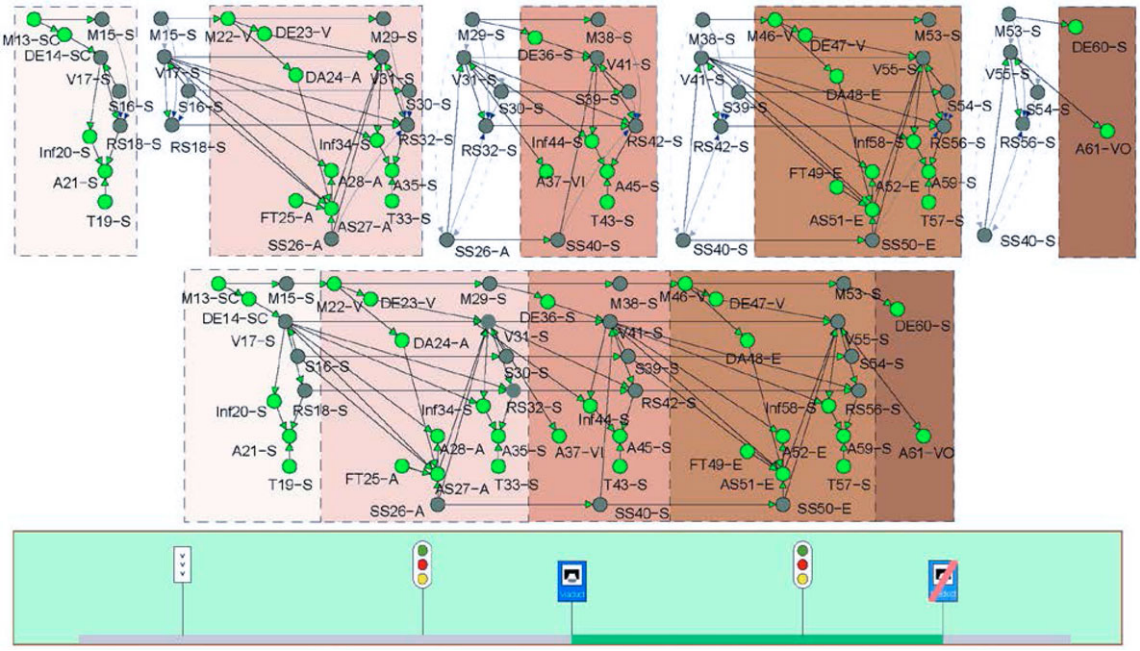


Fig. 9. Illustration of how a Bayesian network can be partitioned into a sequence of Bayesian subnetworks to obtain the marginal probabilities (forward process).

four nodes *M, S, V,* and *RS*, the conditional probabilities *P(M)*, *P(S|M)*, *P(V|S, M)* and *P(RS|V, S, M)* associated with a saturated network are calculated based on the previous joint density using the conditional probability definition, that is,

$$P\ (S|M) = \frac{P\,(S, M)}{P\,(M)};$$

$$P\ (V|S, M) = \frac{P\,(V, S, M)}{P\,(S, M)}. \qquad (14)$$

These probabilities are needed to be transferred to the next partitions.

Similarly, for the separator with six nodes *SS*, *M*, *S*, *SS*, *V*, and *RS*, we can obtain the conditional probabilities *P*(*SS*), *P*(*M*|*SS*), *P*(*S*|*M*, *SS*), *P*(*V*|*S*, *M*, *SS*), and *P*(*RS*|*V*, *S*, *M*, *SS*). In this case, we have added the artificial links (*M–SS*, *M–V*, *M–S*, *M–RS*, and *S–SS*) due to the same reasons.

7. Step 7. *Increase counter*. The counter k is increased.
8. Step 8. *Build the Bayesian network* B^*_k. The Bayesian network B_k is built based on its nodes, links, separator and probability tables (conditional probabilities of the sons given their parents). Note that to the nodes in B_k we have added the nodes in the separator set and the necessary links to convert the separator into a saturated set so that any joint probability, that is, including any dependence structure for these variables can be incorporated. In other words, the variables involved in a required joint probability are all included in a clique. Go to Step 3.

The above process based on partitions reduces substantially the computation time, which becomes linear in the number of nodes or linear in the number of subnetworks.

6 DATA AND INFORMATION GIVEN BY THE MODEL

In this section we describe the necessary information and the output supplied by our model.

The required information consists of the following:

- *Railway regulations to be applied*.
- *Line description*. Containing a detailed description of the location and characteristics of switches, signals, level crossings, tunnels, viaducts, curves, etc.
- *Driver's booklets.* With the characteristics safety regulations for each train and line including detailed maximum speeds, timing, etc.
- *Train characteristics*. Power, maximum speeds, lengths, accelerations, decelerations, etc.

Table 2
Sorted list of the 20 items associated with the largest equivalent severe incident probabilities of occurrence

					Probability	
Rank	*Item*	*Item name*	*PK*	*Node*	*Actual*	*Improved*
1	196	SignalT	390.350	A2230-ST	0.0604825	5.1756e-10
2	197	Blackspot	390.400	A2240-Bs	6.17337e-05	3.57557e-05
3	199	Blackspot	390.605	A2259-Bs	1.27699e-05	9.91016e-06
4	712	StopStation	514.500	A7854-Sta	3.47954e-06	9.41827e-08
5	582	SignalP	487.302	A6431-SP	2.39236e-06	4.52084e-07
6	581	DriverBookletEnd	487.300	A6418-LHE	1.06821e-06	9.3543e-08
7	708	SignalP	513.425	A7816-SP	5.27282e-07	–
8	506	SignalT	469.600	A5581-ST	3.17288e-07	–
9	707	SignalE	513.241	A7803-E	2.59124e-07	–
10	408	SignalT	451.540	A4529-ST	1.59965e-07	–
11	320	SignalT	428.300	A3583-ST	1.4682e-07	–
12	293	SignalP	420.650	A3300-SP	1.14658e-07	–
13	276	SignalE	415.200	A3114-E	5.98791e-08	–
14	70	GradeCrossing	331.600	A790-GC	4.59993e-08	–
15	227	GradeCrossing	396.950	A2562-GC	4.59993e-08	–
16	254	GradeCrossing	407.400	A2866-GC	3.95881e-08	–
17	324	SignalT	429.600	A3628-ST	3.6097e-08	–
18	229	GradeCrossing	397.650	A2586-GC	2.96232e-08	–
19	632	GradeCrossing	495.185	A6964-GC	2.80584e-08	–
20	24	SignalE	308.650	A266-E	2.25965e-08	–

- *A video taken from the cabin in both directions.* These two videos are very important to make decisions about the line safety.

We have developed a Matlab computer program in which the proposed methods have been implemented. With the information above we can prepare the input to our computer program. Some examples are given below.

The output of our model consists of:

1. A list with the marginal probabilities of all incident nodes (A nodes). These probabilities are obtained easily because they are the marginal probabilities of single nodes. The software packages JavaBayes and BNT provide these probabilities directly. This permits identifying the risks associated with each possible incident.
2. A sorted list of the probabilities of a severe equivalent incident (the levels small, medium, and severe have been weighted conveniently with weights 1/500, 1/10, and 1, respectively). This list, obtained directly from the incident node A marginal probabilities by using adequate weights, permits identifying the locations (items) where the safety must be improved first. There is no reason to spend money and time to improve safety at some locations when other locations have associated larger probabilities of severe incidents.
3. Plots of the structure of the subnetworks of all items and accumulated probabilities of (equivalent) severe incidents to identify the most dangerous items visually.
4. Causal analysis information if needed.

The following examples illustrate how this output information is reported.

7 EXAMPLES

In this section we present three artificial examples and the case of the real Palencia–Santander line.

7.1 The Palencia–Santander line

The conventional single track Palencia–Santander line has been chosen to be the first one to be used to test the proposed model. In fact, we aim at testing the model with several more before using it to assess the safety of all Spanish lines. We think that it is noteworthy to report some interesting results obtained from this line, which starts in Palencia (PK 297.0) and ends at Santander (PK 514.5), with a length of 217.5 km.

Our study consists of a PRA of all the elements in the line but stations (stations are very important elements

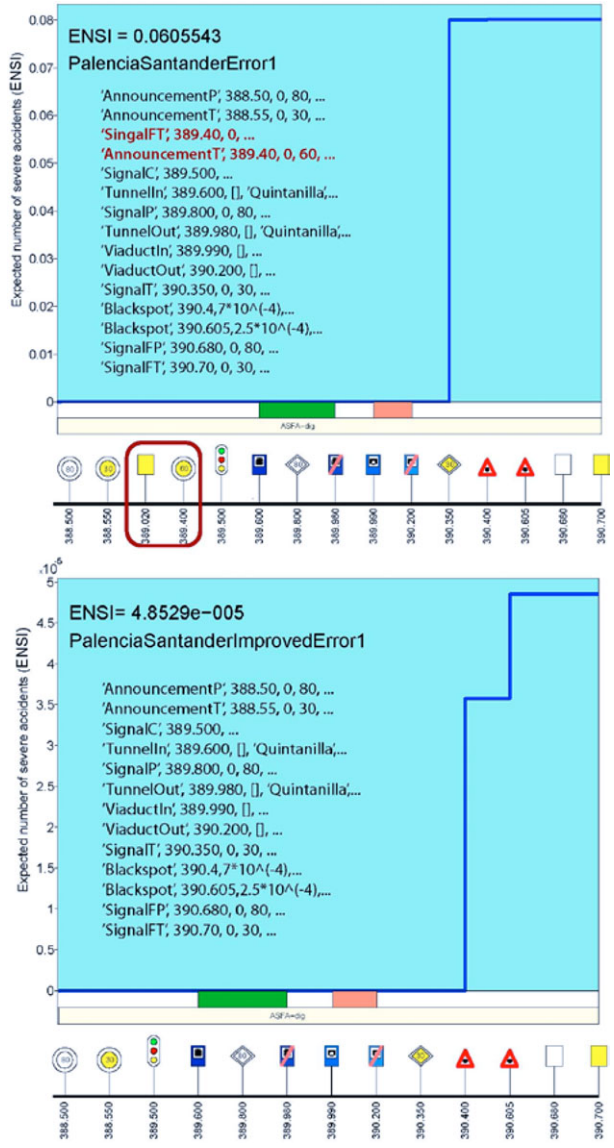


Fig. 10. Case 1. Nested permanent and temporal speed limit signals: actual (upper figure) and corrected situations (lower figure).

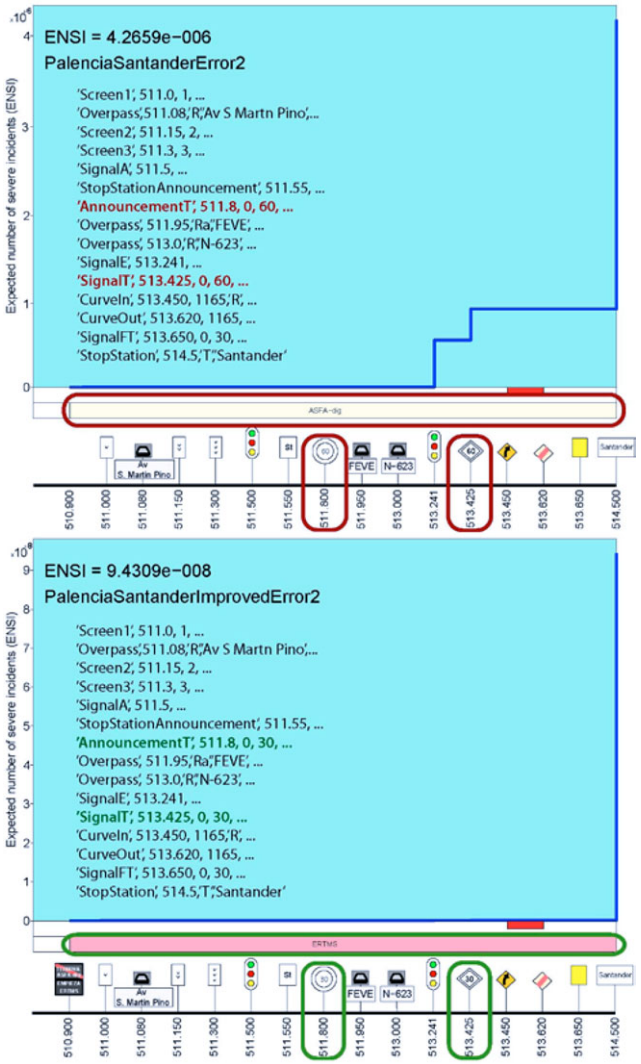


Fig. 11. Case 2. Buffer stop at a terminal station: actual (upper figure) and corrected situations (lower figure).

in a line that need a specialized analysis), that is, we analyze only:

1. A list of 709 items including the elements described in the list in Table 1, which generated a total of 7,820 variables.
2. We also considered the evolution of the driver's tiredness and the driver's attention along the line, because of their relevance on the line safety.

The different items encountered in this line and their frequencies are given in Table 1.

The parameter estimation was done by a mixture of common sense, observed data, recommendations given in the existing literature, discussion among experts, and validation of results (when probabilities of given events were very high or low, the parameters were corrected by trial and error).

The computer time required to analyze this line once the data were given to the computer, was only 11 minutes for the calculations, and 45 minutes including calculations and the generation of all 92 figures reproducing the local structures of the Bayesian network and the incident probability graphs, all the marginal probabilities of the incidents, sorted by item and by associated risks and the report including all figures and tables. It is important to mention that this study could not be done without the partition technique. When the whole Bayesian network was given to the computer program, it worked for more than 12 hours without providing the result. In addition, we indicate that the

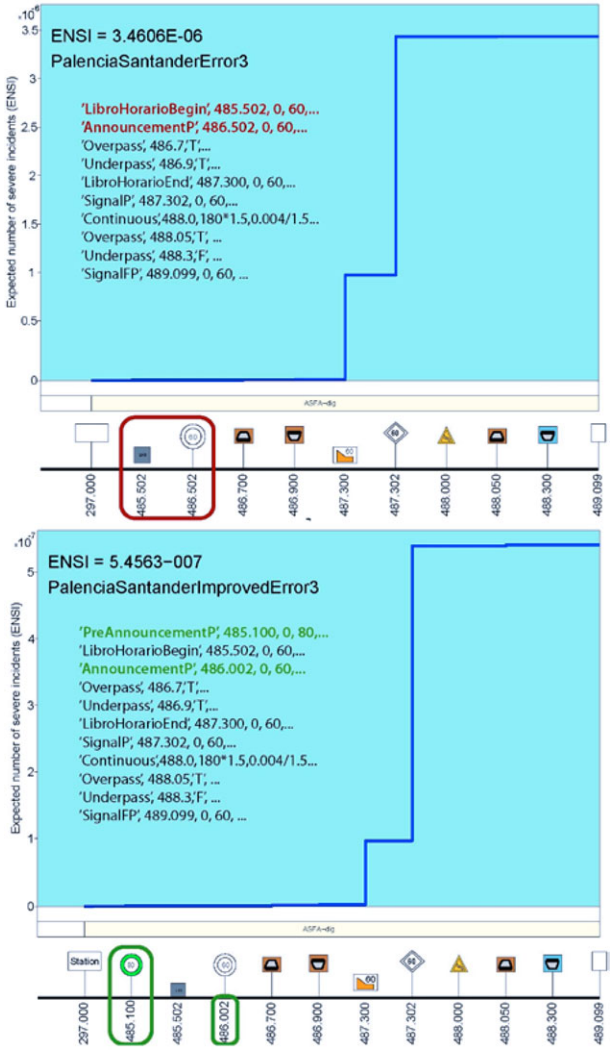


Fig. 12. Case 3. Maximum speed reduction indicated in the Driver's booklet: actual (upper figure) and corrected situations (lower figure).

JavaBayes software cannot deal with this case, unless it is partitioned. In fact, the partitioning technique was developed to solve this problem.

7.2 Results of the computer program

Table 2 shows the sorted list of the 20 items associated with the largest equivalent severe incident probability of occurrence, where it can be seen that the incidents are related to temporal and permanent speed limit signals, light signals, an end station buffer stop, grade crossings and blackspots (they refer to two particular locations where the risk of large falling stones in fractured rock cuts is large).

The use of the Bayesian network model allowed identification of the locations where the probabilities of occurrence were the highest and suggested the order of actions to be considered for safety improvement. It is worthwhile mentioning that the model identified some especial unexpected risks, which would be difficult to identify by another means, an important feature of the model that needs to be emphasized.

Once the computer program is used, we get a list of events sorted by their associated probabilities of a severe accident to occur. Thus, we need to analyze only those above a given threshold value considered as the largest admissible value. We note that an event with associated probability of occurrence smaller or equal to 10^{-9} is considered by the RSSB as deserving no further analysis. In the following subsections we analyze only the largest three.

After presentation of the results, the generalized opinion of the consulted experts in Spain was that this tool provided very valuable information and complements other existing tools for safety assessment. In the next section we discuss in detail three of the most critical cases. We note that the two blackspots indicate that some action is needed on the fracture rock to avoid falling blocks.

1. Case 1: *Nested permanent and temporal speed limit signals*. In this example, we detect the case of a wrong placement of temporary speed limit signs used to protect the train against falling blocks from the fractured rock and nested with permanent speed limit signs. The example shows how the model identifies the problem and how it can be solved.

In the upper plot in Figure 10, where the list of items is shown and the red and salmon colored rectangles refer to tunnel and viaduct, respectively, we show the code to be used in the program for a line segment in which two sets of permanent and temporary speed limit signs have been nested. The permanent speed limit includes an announcement speed limit sign of 80 km/h at PK 388.500, a speed limit signal of 80 km/h at PK 389.800, and an end of speed limit signal at PK 390.680, the temporal speed limit is double, the first includes an announcement speed limit signal of 30 km/h at PK 388.550 and an end of speed limit at PK 389.020 and the second contains an announcement speed limit signal of 60 km/h at PK 389.400, a speed limit signal of 30 km/h at PK 390.350 and an end of speed limit signal at PK 390.700.

Apart from the absence of the preannouncement speed limit signals, the main problem consists of the inconsistency of the temporal speed limit signals at PK 389.400 and PK 390.350, because the first indicated 60 km/h and the second 30 km/h. This makes it impossible to satisfy these constraints because there is not a distance long enough to reduce speed to 30 km/h.

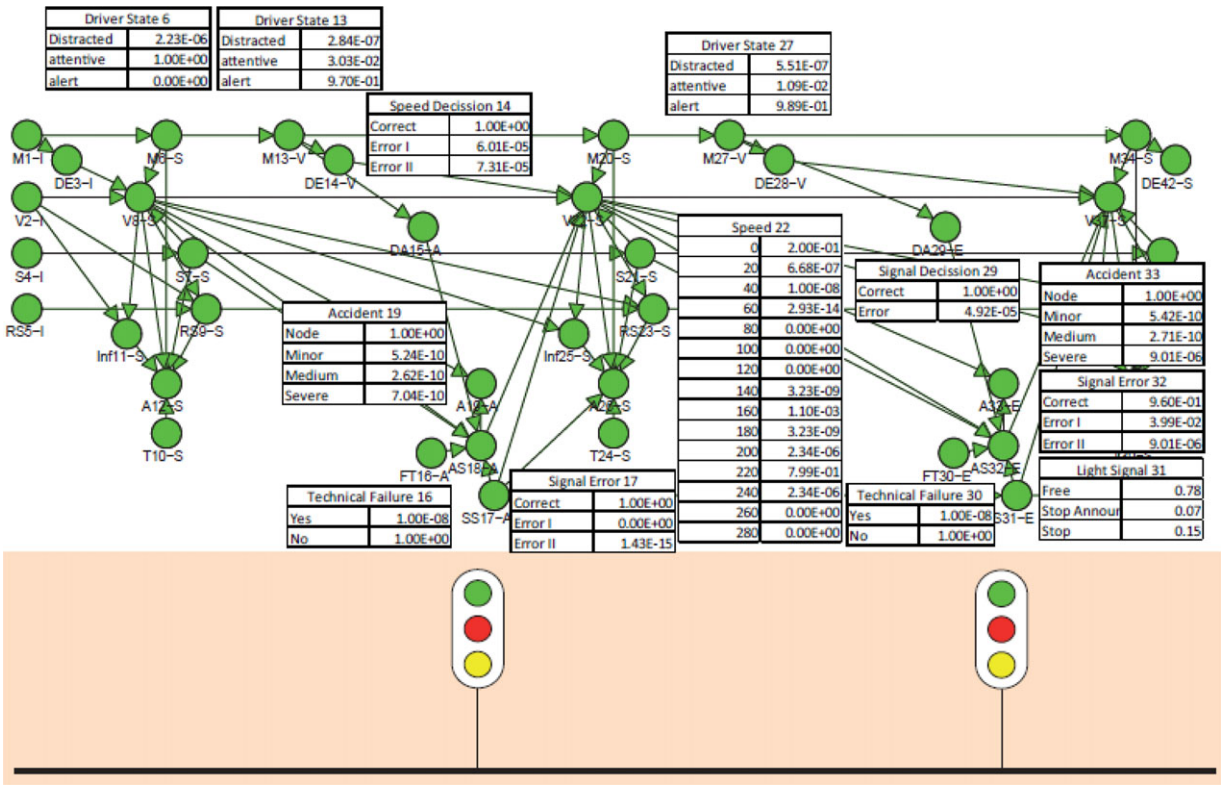


Fig. 13. Bayesian network associated with the example.

First we want to mention that the problem is identified because of a very large probability 0.0604 of occurrence of a severe incident. This gives more information about the risk and possible consequences of this error than a classical analysis will provide. In fact, it indicates not only that the error is present but that the error can produce serious consequences with a high frequency.

One way to solve this problem consists of removing the two signals at PK 389.020 and PK 389.400, as shown in the lower plot in Figure 10. In Table 2, we can see that the probability of occurrence of a severe incident diminishes to 5.1756e-10, which is an important reduction. We can also see that the probability of a severe incident due to falling blocks diminishes too (due to speed reduction) but only by a small amount (see the last column in Table 2).

2. Case 2: *Buffer stop at a terminal station.* In this example our model identifies a high risk associated with the bumper or buffer stop at the end station. Again this is identified because of a relatively large probability 3.47954e-06 of occurrence of an equivalent severe incident.

In the upper plot in Figure 11, where the red rectangle refers to a curve, we present one case in which our model detects a bumper or buffer stop problem at a terminal station. The code to be used in the program is as indicated in the figure. We have a speed limit announcement signal (60 km/h) at PK 511.800, a speed limit sign (60 km/h) at PK 513.425 and an end of speed limit sign at PK 513.650 with a buffer stop at PK 514.500.

The problem here is that first, the speed limit is too high and that the end of speed limit sign suggests increasing the speed after passing the signal. The way to solve this problem, as indicated in the right plot in Figure 11 consists of replacing the speed limitations to 30 km/h, removing the end of speed limit sign and installing an ATP system to stop the train if necessary. This can be checked if the ENSI values for both solutions (initial and improved) are compared.

As can be seen in the right column in Table 2, the probability of occurrence of an equivalent severe incident reduces to 9.41827e-08. This reveals the importance of the quantification of incident probabilities, which allows us to determine if the selected solution to the problem reduces the risk to the desired safety levels.

3 Case 3: *Maximum speed reduction indicated in the Driver's booklet.* In this example, we identify a wrong placement of a maximum speed reduction

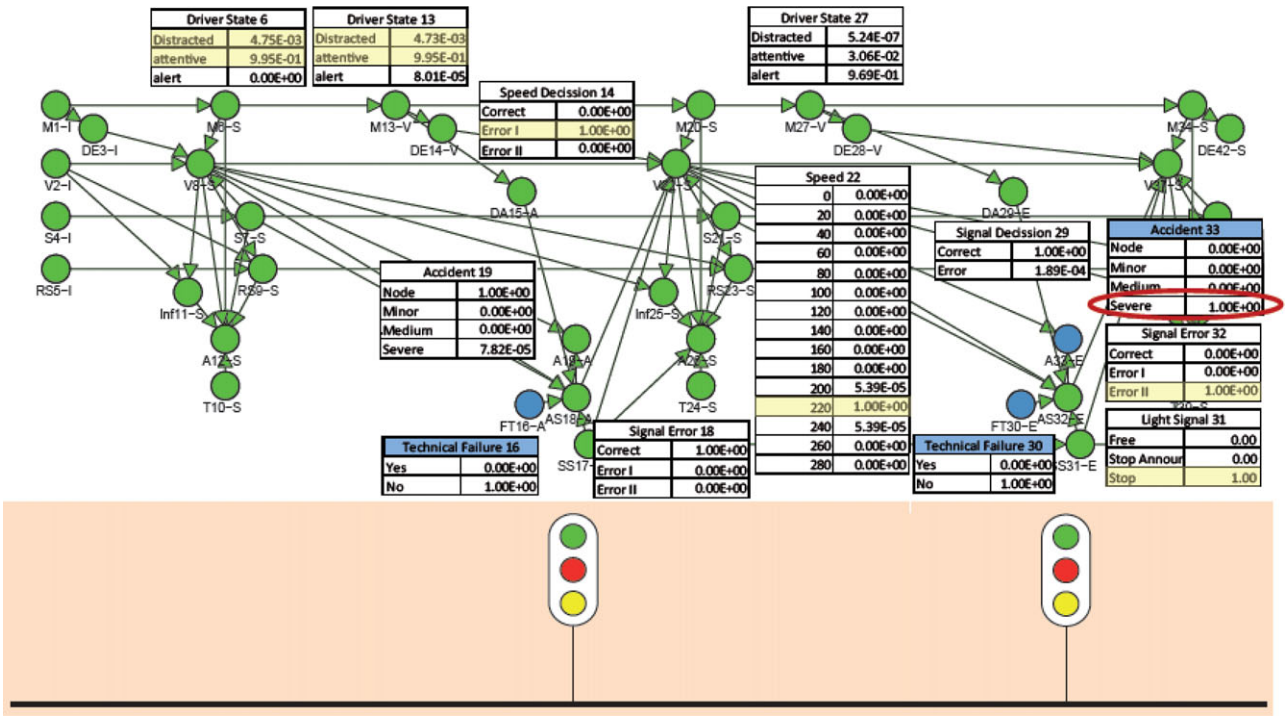


Fig. 14. Illustration of the backward inference based on the occurrence of an accident.

signal to complement the information contained in the Driver's booklet. In fact, we find probabilities of severe incidents 2.39236e-06 and 1.06821e-06 in a permanent speed limit signal and driver's booklet items, respectively.

As shown in the upper plot in Figure 12, where the code to be used in the program is given, we have a maximum speed reduction to 60 km/h at PK 487.302, which is also indicated by means of a speed reduction announcement signal at PK 486.502 and in the driver's booklet. We assume that the driver has decided to fix PK 485.502 as the location where the speed reduction must be initiated.

The problem here is that the train is assumed to run at a low speed at PK 486.502 because there is a station, where all trains stop. However, if one train does not stop, its speed will be high enough to make the distance between the announcement and speed limit sign insufficient.

This problem can be corrected by including a previous announcement signal at PK 485.100 and moving the announcement signal to PK 486.002 to allow for sufficient distance, as indicated in the lower plot in Figure 12. With this change, the probabilities of severe incidents reduce to 4.52084e-07 and 9.3543e-08, respectively, as indicated in the right column of Table 2.

8 A BACKWARD ANALYSIS EXAMPLE

Though the most common and natural use of the proposed model consists of determining by a forward process the marginal probabilities of the nodes and especially the incident nodes, the model can also be used for identifying the most likely causes of some events, especially severe accidents.

In this section we explain how the Bayesian network model can be used to explain the causes of incidents or accidents. Figure 13 shows one example of a Bayesian network whose nodes, links, and some marginal probability tables are shown.

Assume that we have had an accident associated with node 33. This means that the probability of accident at node 30 becomes one. If we add this evidence to the network by fixing this probability to one and if we also know that there was no technical failure at nodes 16 and 30, we can use the proposed evidence backpropagation method to obtain the conditional probabilities of the nodes given the evidence, as shown in Figure 14.

One thing is the probability of having an event and another one is the probability of an event given that other event has already happened. If an accident has happened, its conditional probability given that it has happened is one, however, the probability of a signal to be at red given that an accident has happened is not one,

but it is different from the probability of a signal to be at red without knowing that the accident has happened. If the accident has happened, we can expect that the probability of the accident due to the signal being at red will increase, because this can be the cause of the accident.

We can see that the probabilities of nodes 31, 32, 22, and 14 reveal that the causes of the accident were that the signal was "at red" (stop), there was a driver's Error II, the speed was 220 km/h and in addition the driver made an error in controlling speed. Apart from this, we can observe an important decrease in the driver's attention (compare the probability values in Figures 13 and 14). This proves the efficiency of the back-propagation of evidence.

Finally, we note that the main contribution of this Bayesian network technique to the backward analysis is not identifying causes but quantifying the associated probabilities. Other methods cannot supply these probabilities.

9 CONCLUSION

The following conclusions can be drawn from the content of this article:

1. Bayesian network models provide an important tool to reproduce a railway line to perform a probabilistic safety assessment of it, and are more powerful than the commonly used fault and event trees, especially when common causes are present. This means that Bayesian networks can help to improve the safety of railway lines.
2. The proposed model permits reproducing all the variables involved in the problem, their qualitative dependencies and the quantification of the associated conditional probabilities. This implies reproducing the probabilistic structure of the associated multivariate random variable.
3. The construction of the nodes (variables) and structure of the Bayesian network is very natural because it reproduces all the items encountered when the train travels along the line. A simple list of items can be given for a computer program to build the acyclic graph associated with the Bayesian network automatically.
4. The proposed method for splitting the Bayesian networks in to small pieces without losing its probabilistic representativeness, permits reducing the initial nonlinear complexity to a complexity which is linear with the number of nodes and subnetworks. Consequently, not only the memory and CPU time requirements are reduced substantially, but real cases can be computed.
5. The application of the proposed methodology to the Palencia–Santander line with 709 items and 7,820 variables proves that the method can be applied to real very large lines and shows the power of the method for identifying sequences of events leading to severe incidents and quantifying their probabilities.
6. Some of the particular examples analyzed in this article show that the method is able to identify and quantify relevant incidents and their probabilities of occurrence and that some are difficult to identify by other means. With this methodology the design of a railway line including signal system design will be improved by adding the probabilities of system failures and incidents.
7. The Bayesian network with its backward possibilities is an ideal tool to analyze the causes of incidents and especially those leading to fatal accidents. In particular, the probabilities of alternative causes, not available in other methods, can be obtained because the probability of incidents can be estimated by Bayesian network methods.
8. The most critical part of the proposed model is the parameter estimation and calibration. In this direction the collaboration of various groups of experts is needed to improve the power, the credibility of the results, and the efficiency of the method. A lot of work still needs to be done in the future, but we think it is worthwhile doing it.
9. Some future work improvements are: a discussion of how to estimate or assess the parameters of the model and to incorporate some tools for estimating some parameter values automatically. One example could be the classification of slope stability or tunnel and viaducts entries and exits in several groups or a tool to evaluate slope stability (see, e.g., Revilla and Castillo, 1977).

REFERENCES

Beales, L. (2002), Guidance on the Preparation of Risk Assessments within Railway Safety Cases, Rail Safety and Standards Board, United Kingdom.

Bearfield, G. & Marsh, W. (2005), Generalising event trees using Bayesian networks with a case study of train derailment, *Lecture Notes in Computer Sciences*, **3688**, 52–66.

Benjamin, J. & Cornell, C. A. (1970), *Probability Statistics and Decision for Civil Engineers*, McGraw-Hill, New York.

Castillo, E., Calviño, A., Grande, Z., Sánchez-Cambronero, S., Gallego, I., Rivas, A. & Menéndez, J. M. (2016), A Markovian-Bayesian network for risk analysis of high speed and conventional railway lines integrating human errors, *Computer-Aided Civil and Infrastructure Engineering*, **31**, 3, DOI: 10.1111/mice.12153

Castillo, E., Gallego, I., Sánchez-Cambronero, S., Menéndez, J. M., Rivas, A., Nogal, M. & Grande, Z. (2015), An alternate double-single track proposal for high speed peripheral railway lines, *Computer-Aided Civil And Infrastructure Engineering*, **30**, 181–201.

Castillo, E., Gallego, I., Ureña, J. & Coronado, J. (2011), Timetabling optimization of a mixed double- and single-tracked railway network, *Applied Mathematical Modelling*, 35, 859–78.

Castillo, E., Gutiérrez, J. M. & Hadi, A. (1997a), Sensitivity analysis in discrete Bayesian networks, *IEEE Transactions on Systems, Man and Cybernetics*, **26**(7), 412–23.

Castillo, E., Gutiérrez, J. M. & Hadi, A. S. (1997b), *Expert Systems and Probabilistic Network Models*, Springer Verlag, New York.

Castillo, E., Menéndez, J. M. & Sánchez-Cambronero, S. (2008), Traffic estimation and optimal counting location without path enumeration using Bayesian networks, *Computer-Aided Civil and Infrastructure Engineering*, **23**, 189–207.

Castillo, E., Sarabia, J. M., Solares, C. & Gómez, P. (1999), Uncertainty analyses in fault trees and Bayesian networks using FORM/SORM methods, *Reliability Engineering and System Safety*, **65**, 29–40.

Dadashi, N., Scott, A., Wilson, J. R. & Mills, A. (2013), *Rail Human Factors: Supporting Reliability, Safety and Cost Reduction*, CRC Press, Taylor and Francis, London.

Dirección de Seguridad en la Circulación, ADIF (2009), Sistema de gestión de seguridad en la circulación. Evaluación y gestión de riesgos. Ref: SGSC/EGR, revisión 3, 5/10/2009.

Doob, J. (1953), *Stochastic Processes*, John Wiley and Sons, New York.

Evans, A. W. (2011), Fatal train accidents on Europe's railways: 1980–2009, *Journal of Accident Analysis and Prevention*, **43**(1), 391–401.

Flammini, F., Marrone, S., Mazzocca, N. & Vittorini, V. (2006), Modeling system reliability aspects of ertms/etcs by fault trees and Bayesian networks, in *17th European Safety and Reliability Conference (ESREL)*, 2675–83.

Fukuyama, H., Inutsuka, F., Tachi, M. & Ishige, T. (2008), Application of risk assessment method in railway, *Sociotechnica*, **1**(5), 163–71.

Henley, E. & Kumamoto, H. (1992), *Probabilistic Risk Assessment; Reliability Engineering, Design, and Analysis*, IEEE Press, New York.

Instituto Nacional de Seguridad e Higiene en el Trabajo (1992), Norma CENELEC 50126.

Kawakami, S. (2014), Application of a systems theoretic approach to risk analysis of high-speed rail project management in the US. Master's thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Kijima, M. (1997), *Markov Processes for Stochastic Modeling*, 1st edn., Chapman & Hall, Cambridge.

Kokkings, S. J. & Snyder, E. A. (1997), *Case Studies in Collision Safety*, Report DOT/FRA/ORD-96/01, Federal Railroad Administration, Washington, D.C.

Lahrech, Y. (1999), Development and application of a probabilistic risk assessment model for evaluating advanced train control technologies. Master thesis, Massachussetts Institute of Technology, Cambridge, Massachusetts.

Masanori, T. & Fumiaki, F. (2008), A study about risk evaluation of JR East, in *Eighth World Congress on Railway Research*, 1–9, Seoul, Korea. Union Internationale des Chemins de Fer.

Ministerio de Trabajo y Asuntos Sociales (2010), Notas Técnicas de Prevención (NTP) 330: Sistema simplificado de evaluación de riesgos de accidente.

Miyashita, N. (2010), 2013 Safety vision, *JR EAST Technical Review*, **15**(1), 163–71.

Mokkapati, C., Tse, T. & Rao, A. (2009), *A Practical Risk Assessment Methodology for Safetycritical Train Control Systems*, Technical Report DOT/FRA/ORD-09/15, U.S. Department of Transportation, Washington DC.

Muttram, R. I. (2002), Railway safety's safety risk model, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, **216**(2), 71–79.

Peterman, D. R., Frittelli, J. & Mallet, W. (2009), High speed rail (HSR) in the United States, CSR Report for Congress R40973, Congress of the United States.

Revilla, J. & Castillo, E. (1977), The calculus of variations applied to stability of slopes, *Geotechnique*, **27**(1), 1–11.

Spackova, S. & Straub, D. (2013), Dynamic Bayesian network for probabilistic modeling of tunnel excavation processes, *Computer-Aided Civil and Infrastructure Engineering*, **28**, 1–21.

Sun, H. & Betti, R. (2015), A hybrid optimization algorithm with Bayesian inference for probabilistic model updating, *Computer-Aided Civil and Infrastructure Engineering*, **30**, 602–19.

Sussman, J. M. (1996), Industry/academic cooperation in transportation: the partnership of JR East and MIT, *Japan Railway & Transport Review*, **7**, 26–33.

Todorovich, H. P. & Hagler, Y. (2011), *High Speed Rail in America*, Technical Report, America 2050.

Veneziano, D. & Papadimitriou, A. G. (2001), *Optimizing the Seismic Early Warning System for the Tohoku Shinkansen*, Springer Verlag, New York.

Wang, H., Yajima, A., Liang, R. Y. & Castaneda, H. (2015), Bayesian modeling of external corrosion in underground pipelines based on the integration of Markov chain Monte Carlo techniques and clustered inspection data, *Computer-Aided Civil and Infrastructure Engineering*, **30**, 300–16.

Wreathall, J., Roth, E., Bley, D. & Multer, J. (2003), *Human Reliability Analysis in Support of Risk Assessment for Positive Train Control*. Technical Report DOT/FRA/ORD-03/15, U.S. Department of Transportation, Cambridge, MA.

Yuen, K. V. & Mu, H. Q. (2015), Real-time system identification: an algorithm for simultaneous model class selection and parametric identification, *Computer-Aided Civil and Infrastructure Engineering*, **30**, 785–801.

Zeilstra, M. P. & van del Weide, R. (2013), Human as an asset in a system consideration on the contribution of humans to system performance and system safety, in N. Dadashi, A. Scott, J. R. Wilson and A. Mills (eds.), *Rail Human Factors: Supporting Reliability, Safety and Cost Reduction*, CRC Press, Taylor & Francis Group, New York, 473–82.