Probabilistic Safety Analysis of High Speed and Conventional Lines Using Bayesian Networks

Zacarias Grande
Assistant Professor, University of Cantabria, Spain.
Enrique Castillo
Full Professor, University of Cantabria and Royal Academy of Engineering, Spain.
Maria Nogal and Alan O'Connor
Full Professor, Trinity College of Dublin, Ireland.

SUMMARY

A Bayesian network approach is presented for probabilistic safety analysis (PSA) of railway lines. The idea consists of identifying and reproducing all the elements that the train encounters when circulating along a railway line, such as light and speed limit signals, tunnel or viaduct entries or exits, cuttings and embankments, acoustic sounds received in the cabin, curves, switches, etc. In addition, since the human error is very relevant for safety evaluation, the automatic train protection (ATP) systems and the driver behavior and its time evolution are modelled and taken into account to determine the probabilities of human errors. The nodes of the Bayesian network, their links and the associated probability tables are automatically constructed based on the line data that need to be carefully given. The conditional probability tables are reproduced by closed formulas, which facilitate the modelling and the sensitivity analysis. A sorted list of the most dangerous elements in the line is obtained, which permits making decisions about the line safety and programming maintenance operations in order to optimize them and reduce the maintenance costs substantially. The proposed methodology is illustrated by its application to several cases that include real lines such as the Palencia-Santander and the Dublin-Belfast lines.

1. INTRODUCTION

Increasing speeds have led high speed trains to successfully compete with air transportation (see Peterman et al. (2009), Todorovich and Hagler (2011), Castillo et al. (2015, 2016c), Andersson et al. (2015)), but the associated risks and their consequences have motivated the appearance of automatic train protection systems (ATP) and the need of probability risk assessments (see Lahrech (1999), Miyashita (2010) or Fukuyama et al. (2008)). In this paper we deal with the problem of probability risk assessment (PRA) of a given railway line and operating company.

The Safety Risk Model (SRM) used by the RSSB (Railway Safety Standard Board) enables global risk to be calculated (see Muttram (2002)) and provides information to be used in risk assessments and on the contribution of the different elements to failure. However, line or segment risk assessments are not very common (see Lahrech (1999) for some small

examples).

Nuclear power plants and other industries have a long tradition in using powerful techniques for probability risk assessments (Henley et al. (1992)), such as fault and event tree analysis, Petri and Bayesian networks, discrete and continuous Markov models, etc. However, the railway industry runs slowly in this area (see Bearfield and Marsh (2005)). However, some countries, such as Japan, the US and the UK, use models where fault and event tree analysis are used in the railway problem. However, in many countries PRAs are neither compulsory nor regularly used in testing the safety of railway lines.

Once we have a parametric model for PRA its parameters need to be estimated. In this context, the works of Kokkings and Snyder (1997), Muttram (2002) and Evans (2011), who report important railway data and serious statistical analysis of the railway accidents which occurred in Europe during the period 1980-2009 become relevant.

Human error is an important factor that has to be considered in any PRA. Quantification of human error probabilities is possible with participation of miscellaneous groups of professionals(see Dadashi et al. (2013)) .Among the most important consequences of human errors (for example, a train trespassing a sign at red or failing to comply with a speed limit) are railway incidents. Thus, a valid model must include human error evaluation.

In the following sections we build a Bayesian model (see Castillo et al. (2016a,b)),which permits us to evaluate the probability of derailments, collisions or other types of incidents along the line being analyzed. Markovian models have been frequently used in safety analysis (see Slovak et al. (2007), Fuqua (2007) or Bjorkman (2011)). A Bayesian network has two components: (a) an acyclic graph including the nodes (variables) and the directed links, which informs us about the relevant variables in the problem and the direct relations (causal or not) among then providing a very valuable qualitative information, and (b) the conditional probabilities of each variable given its parents, which complete the quantitative information. Contrary to fault trees (see Lahrech (1999)), Bayesian networks (see Castillo et al. (1997)) allow any joint probability distribution to be reproduced, including common causes.

2. VARIABLES INVOLVED IN THE PROBLEM.

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: *driver's tiredness (D), driver's attention (D_a), speed (S), incident (A), rolling stock (RS), infraestructure (Inf), terrain (T), driver's decision on speed control (DE), driver's decision at sign (DA), automatic train protection system (ATP), light sign decision (AS), driver's decision made at a speed limit sign (DS), light sign state (SS) and technical failure (TF).*

3. BAYESIAN NETWORKS AND SUB-BAYESIAN COMPONENTS.

The Bayesian network structure depends on the railway line being studied because all the elements the driver encounters when travelling along the line are reproduced and different lines have different elements.

In our model there are two types of Bayesian subnetworks, as indicated in Figure 1:

1. The first corresponds to the segments without signs or signals, used to evaluate the risk incurred when driving, due to rolling stock failure, infrastructure state (rails, sleepers, ballast, plate, maintenance standards (see \cite{PengO:14}), etc.) or problems associated with slipping or falling materials in areas of deforestation, landslides in areas of landfill, differential settlement between rails, etc. This subnetwork type considers a Markovian model in continuous time and is shown in Figure 1 to the left of each element, indicated in brown color.

2. The second subnetwork type corresponds to each of the points or locations in which signs are located, decisions subject to errors are to be taken, occasional incidents can occur due to different reasons (step turnouts, entrances and exits of tunnels, sediment accumulation on the track, etc.), viaducts entries and exits. This type of subnetwork is shown in Figure 1 to the right of each element, indicated in blue.
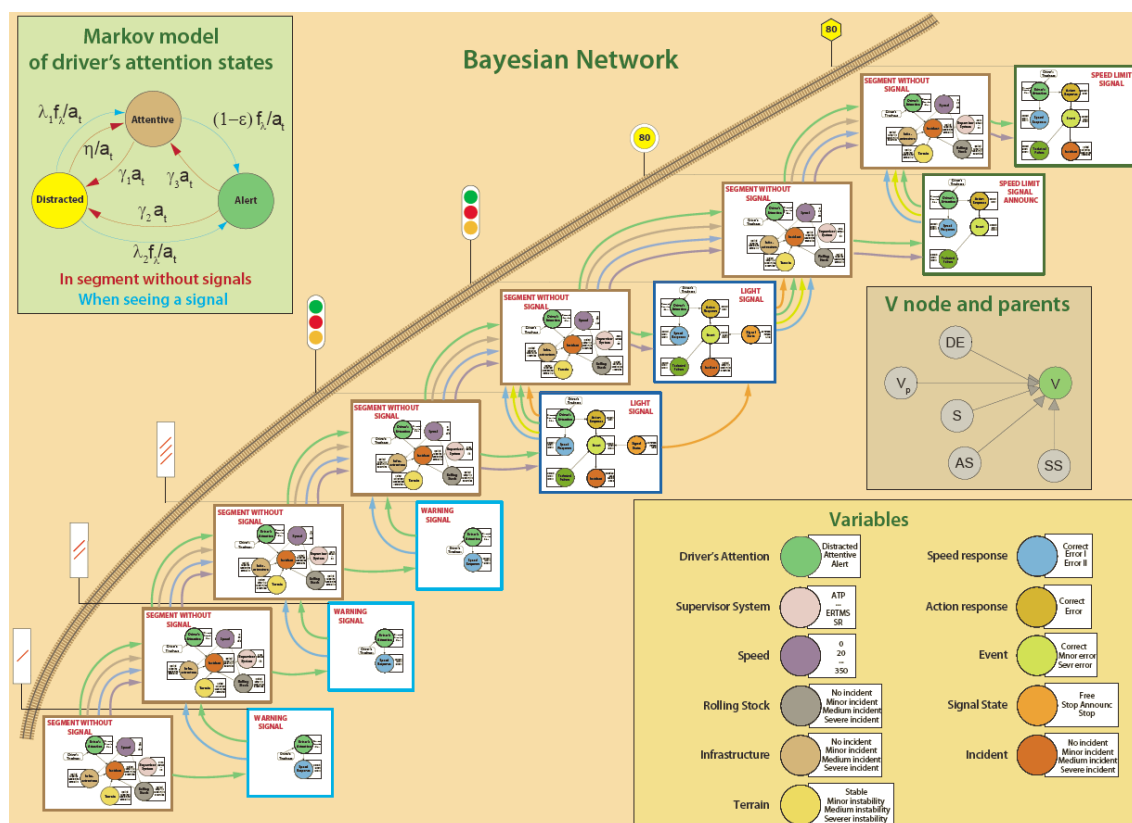

Figure 1 Illustration of (a) the Bayesian network model showing the Bayesian subnetworks for the segments without signs and signals (left subnetworks) and for the segments with signs or discontinuous associated risk (right subnetworks), (b) the

Markov model of driver's attention states, and (c) the *V* node and its parents.

The complete Bayesian network results after joining all sub-networks (see Figure 1)

3.1 Defining conditional probabilities

As indicated, a very important step in the process of building the Bayesian network is the definition of the conditional probabilities, one per node. To this end, we have made a huge effort in obtaining closed formulas in order to facilitate the comprehension and its implementation into the computer programs.

As one example, we show the case of the conditional probability $P(V/V_p,DE,SS,AS,S)$. In this case (see Figure 1), node V is connected to 5 parent nodes and the conditional probability matrix depends on the previous speed V_p the driver's decision on the speed DE and DA at the light sign, the signal state and the operating ATP system. Thus, the final conditional probability matrix $p\_{a,b,c,d,e,f}(s)=P(V=a/V_p=b,DE=c,SS=d,AS=e,S=f)$ is:

$$p_{a,b,c,d,e,f}(s) = \delta_{e,1}\big[\delta_{c,1}\delta_{a,s} + \delta_{c,2}\big((1-\rho_f)\delta_{a,s} + \rho_f\delta_{a,b}\big) +$$
$$+ \delta_{c,3}\left(\kappa_1\rho_f\delta_{a,\max(1,s-1)} + \left(1 - \rho_f(\kappa_1+\kappa_2)\right)\delta_{a,s} + \tag{1}\right.$$
$$\left. + \kappa_2\rho_f\delta_{a,\min(n,s+1)}\right)\big] + (\delta_{d,2}\delta_{e,2} + \delta_{d,3}\delta_{e,3})\big((1-\rho_f)\delta_{a,s} + \rho_f\delta_{a,b}\big),$$

where the v_{target} is defined according to the sign state, determined by SS node, $\delta_{a,b}$ are the Kronecker's deltas, ρ_d is the failure probability of the ATP system and κ_1 and κ_2 are the probabilities of selecting erroneously the speeds $v_{max(1,s-1)}$ and $v_{min(n,s+1)}$, respectively. More details can be seen in Castillo et al. (2016 a,b) and Grande (2015).
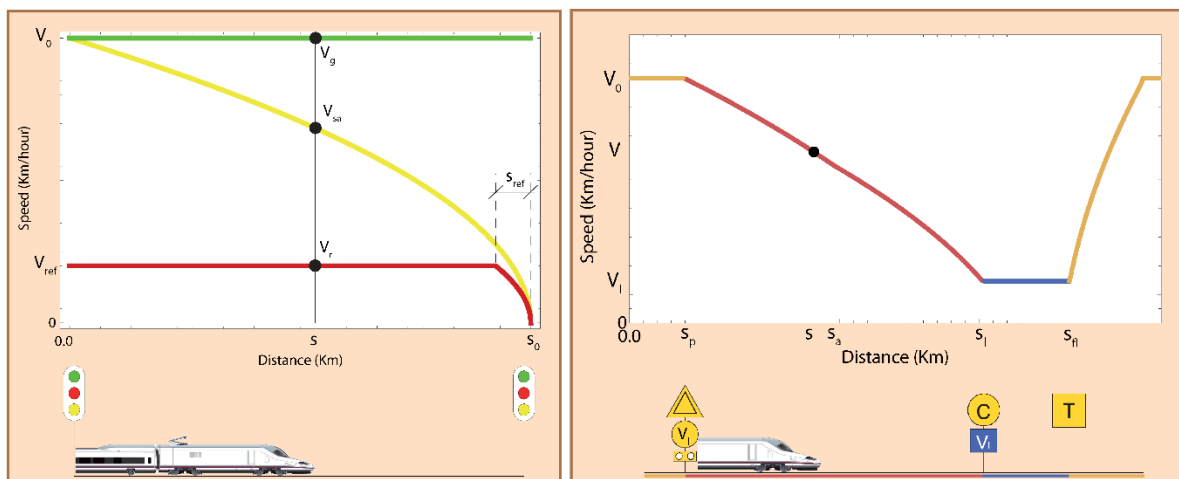

Figure 2 Illustration of different situations in which the train accelerates or decelerates according to: (a) light sign, (b) temporary speed limit sign.

The model needs to take into account the train possibilities in terms of feasible maximum accelerations and decelerations. That is, assuming that the train has a known speed at a given location, the model must check whether or not another speed at another location is attainable. To this end, the whole range of possibilities must be analyzed. As one example, Figures 2

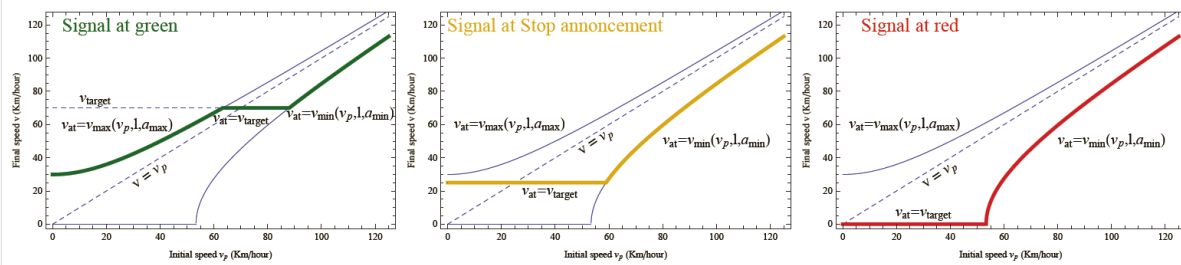and 3 show the final attainable speeds at a location when the train starts with a speed v_p.

Figure 3 Illustration of the attainable speed when the initial speed is v_p and the target speed is v_{target} in the cases of light signs at green, at stop announcement (orange) and at stop (red).

4. PROBABILISTIC SAFETY ANALYSIS.

In this section we present two real examples.

4.1 Palencia-Santander line

The Palencia-Santander line is a 217 Km single-track line that belongs to the Spanish conventional railway network. After using the proposed methodology, 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 is detected. The example shows how the model identifies the problem and how it can be solved.

In the table at the bottom of Figure 4, where the red and salmon colored narrow rectangles refer to curve and viaduct, respectively, the list of most critical items of the line are given sorted by probability of failure in decreasing order of magnitude. They correspond to a line segment in which two sets of permanent and temporary speed limit signs have been nested to protect against falling blocks at a curve. The permanent speed limit includes an announcement speed limit sign of *80* km/h at PK 388.500, a speed limit sign of *80* km/h at PK 389.800 and an end of speed limit sign at PK 390.680 (not shown), the temporary speed limit is double, the first includes an announcement speed limit sign 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 sign of 60 km/h at PK 389.400, a speed limit sign of 30 km/h at PK 390.350 and an end of speed limit sign at PK 390.700 (not shown).

Apart from the absence of the pre-announcement speed limit signs, the main problem consists of the inconsistency of the temporary speed limit signs 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.

The problem is identified because of a very large probability *0.0653* 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 signs at PK 389.020 and PK 389.400, as shown in the lower plot in Figure 4. In the table we can see that the probability of occurrence of a severe incident diminishes to *1.88*10^{-7}*, 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 the table). Thus, it would be convenient to reduce this probability even more.

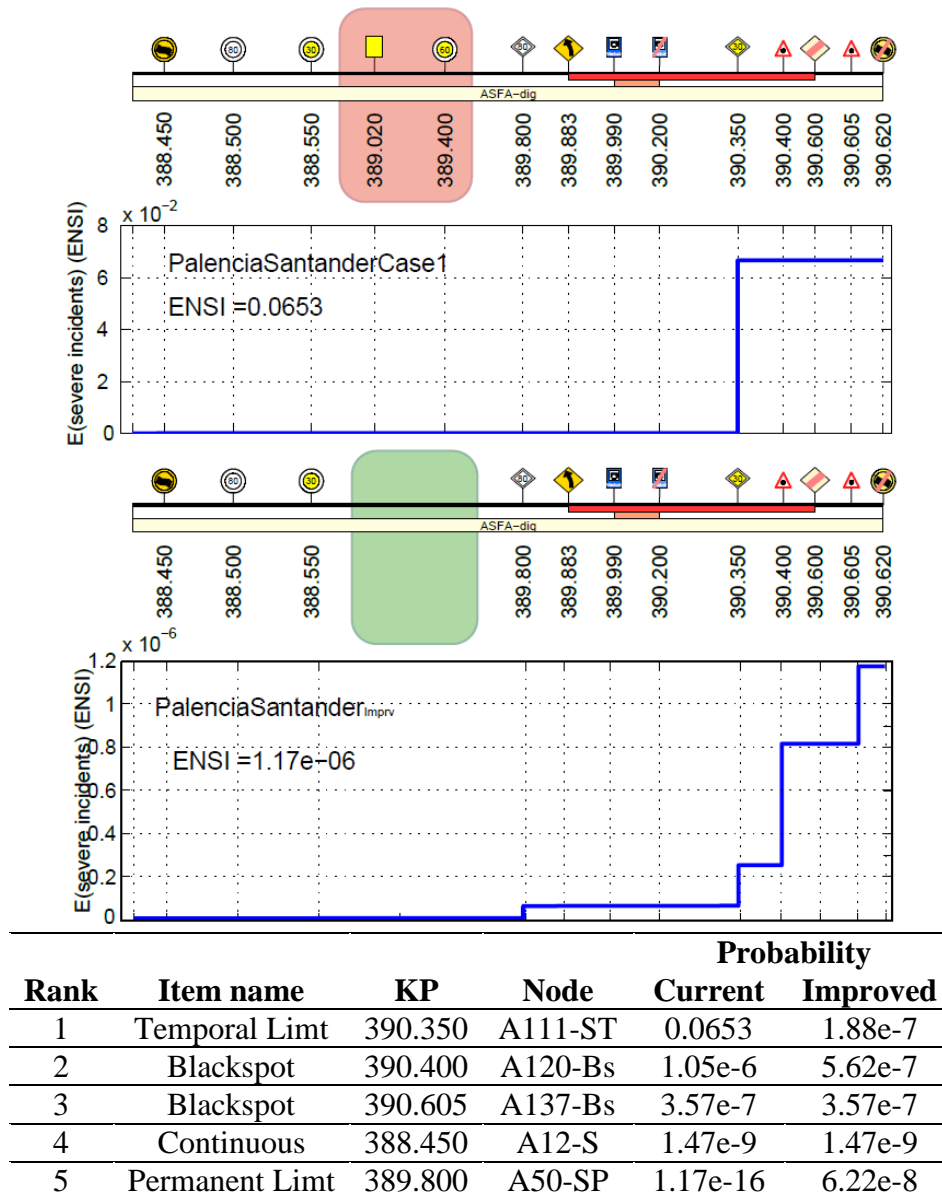| Rank	Item name	KP	Node	Probability Current	Improved
1	Temporal Limt	390.350	A111-ST	0.0653	1.88e-7
2	Blackspot	390.400	A120-Bs	1.05e-6	5.62e-7
3	Blackspot	390.605	A137-Bs	3.57e-7	3.57e-7
4	Continuous	388.450	A12-S	1.47e-9	1.47e-9
5	Permanent Limt	389.800	A50-SP	1.17e-16	6.22e-8

Figure 4 Palencia-Santander line. Illustration of the most hazardous case, before (upper plot) and after (lower plot) the improvement, and their associated probabilities.

3.2 Dublin-Cork and Cork-Dublin lines

A preliminary probabilistic safety analyses of the Dublin-Cork and Cork-Dublin lines have

been done to assess the risks associated with travelling these lines. These analyses include the most important items that trains encounter when travelling along the line in both directions and permit us to identify the most critical items from the safety point of view. After the analysis, the following conclusions can be drawn:

1. The line Dublin-Cork presents three critical locations, which correspond to a permanent speed limit sign, to a temporary speed limit sign and to a light signal, respectively, where the first is especially critical due to its very large expected number of severe incidents (ENSI) value. As located, the speed limit signs do not permit satisfaction of the safety requirements due to an insufficient length between signs, which do not allow reducing the train speed to the required limit. The three cases can be easily corrected by replacing signs or signals,

2. The line Cork-Dublin presents two critical locations, which correspond to a permanent speed limit sign and a buffer stop at the end of the line. As indicated, these cases can be corrected by installing an ERTMS system to protect against an speed excess at the final station,

3. With the exception of the applied solution for one of the critical points in the Cork-Dublin direction, the designed improvements are based on simple works, such as sign location changes or placements of new signs. Consequently, the costs associated with a significant improvement of the safety of the line are very small. The rest of problems can be solved by installing an ERTMS system.

4. The order in which actions should be taken is indicated by the sorted list where items implying higher risks are located first.

To illustrate the process we present some details on the PSA of the Cork-Dublin line in the Southwest-Northeast direction. This line has *264.323* km and is modeled with *541* items, which imply *6045* variables and a cumulated ENSI of *4.48e-06*.

Table 1 shows the sorted list of the most critical items with the corresponding locations (PK), nodes and local probabilities.

Table 1 Sorted list of critical items for the Dublin-Cork line with the corresponding numbers, names, locations (KP), nodes and local probabilities.

Rank	Item	Item name	KP/mile	Node	Probability
1	532	Permanent Limt	262.720/1.001	A5945-DBl	1.45e-06
2	541	Stop station	264.323/0.000	A6045-Sta	6.92e-07
3	540	Permanent Limt	263.920/0.252	A6036-DBl	5.19e-07
4	196	Permanent Limt	91.120/108.251	A2146-DBl	4.97e-07
5	316	Light sign	157.902/66.513	A3499-E	1.66e-07
6	538	Permanent Limt	263.520/0.502	A6014-DBl	1.62e-07

The first conclusion that can be extracted after consulting the list of items sorted by their

risks is that this line provides a reasonable safety in this direction because the final ENSI is $4.48*10^{-6}$. The critical points in this case are located at the end of the line (see Table 1), i.e. at the Heuston Station environment.

Since the most hazardous items are relatively close to one another, this analysis will consider the two items together, because the possible improvements will affect both items. These hazardous situations are related to the driver's cumulative tiredness at the end of the trip, to the different permanent speed limit signs, which can lead to errors, and to the lack of a continuous protection system against speed limit violations at the station proximity causing failure to stop before the buffer stop. Figure 5 illustrates the main errors and the suggested improvements to reduce these risks.

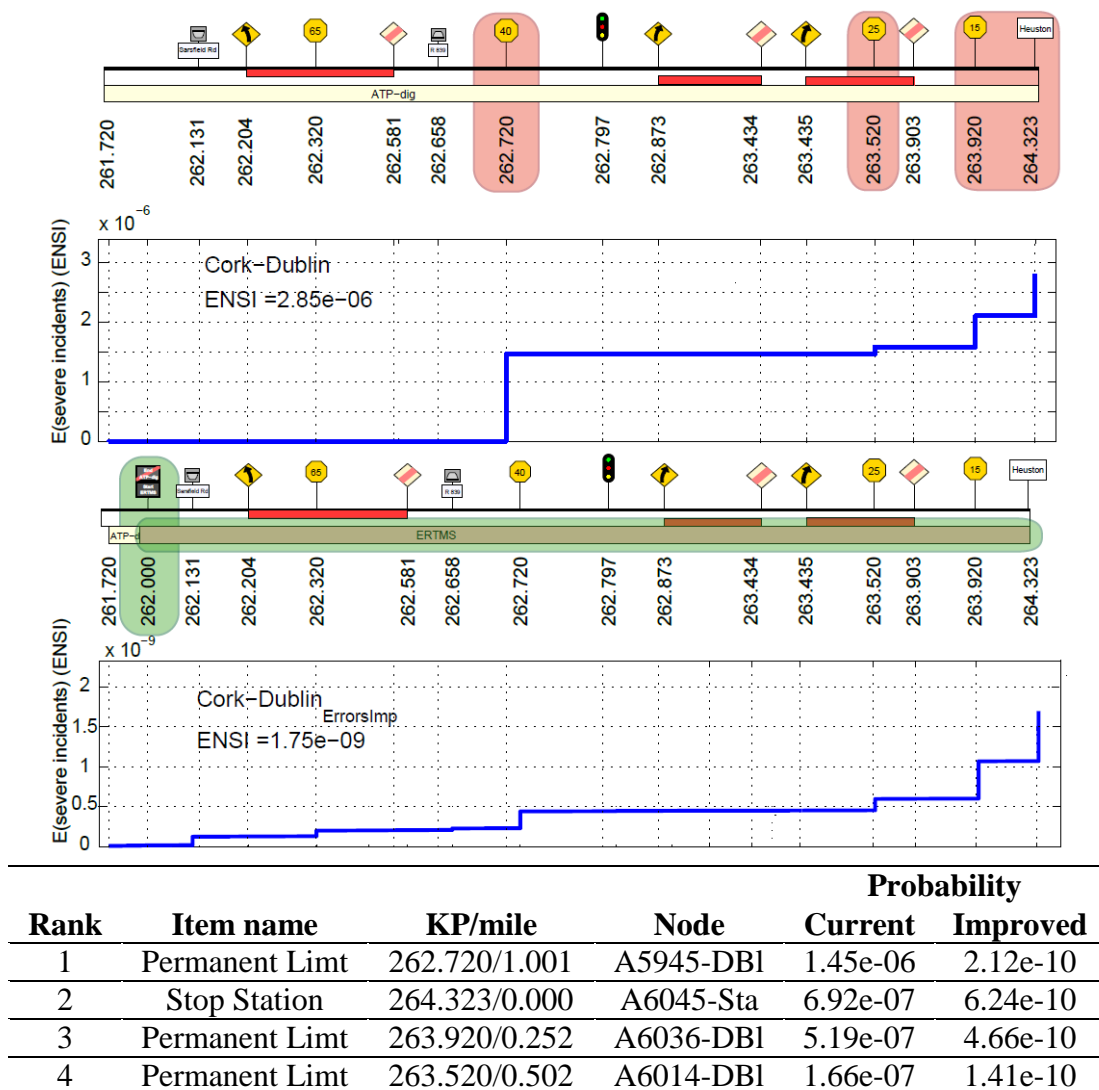| Rank	Item name	KP/mile	Node	Probability	
				Current	Improved
1	Permanent Limt	262.720/1.001	A5945-DBl	1.45e-06	2.12e-10
2	Stop Station	264.323/0.000	A6045-Sta	6.92e-07	6.24e-10
3	Permanent Limt	263.920/0.252	A6036-DBl	5.19e-07	4.66e-10
4	Permanent Limt	263.520/0.502	A6014-DBl	1.66e-07	1.41e-10

Figure 5 Most hazardous point along the Cork-Dublin line before (upper plot) and after (lower plot) the improvement, and their associated probabilities.

Due to the fact that four different items (the three speed limit signs and the buffer stop sign) are involved in the safety of the current location, a joint action has been proposed, which is based on installing the ATP system to protect the last 1.5 miles of the line. This safety

improvement would control the train speed along the line in order to guarantee a safety improvement against permanent speed limit sign violations. It should be noted that if in a short-medium term the ERTMS system implantation were not possible, at least some balises protecting the permanent speed limit signs should be arranged.

The principal result of the proposed corrections is the reduction of the overall risk probability. Since the initial cumulated probability was $4.48*10^{-6}$ and after the correction of the three major critical points, at mileposts 1.001 (KP 262.720), 0.252 (KP 263.920) and 0.000 (KP 264.323), respectively, the final probability becomes $1.29*10^{-9}$.

5. CONCLUSIONS.

The most important conclusions derived from this work are:

1. Bayesian networks are a powerful tool to perform a PSA of a railway line in which all items encountered by trains are reproduced in a natural way to build the network. This tool permits reproducing any probabilistic improved structure and competes with advantage with fault trees.
2. The cases of the Palencia-Santander and Dublin-Cork lines show that they provide a relevant information about the critical points in the line and suggest how the maintenance program must be organized, starting with the most dangerous items in order to save time and money.
3. Work is needed to provide general rules to estimate the parameters of the model and to validate the associated results.

REFERENCES

ANDERSSON, A., O'CONNOR, A., AND KAROUMI, R. (2015). Passive and adaptive damping systems for vibration mitigation and increased fatigue service life of a tied arch railway bridge. *Computer Aided Civil And Infrastructure Engineering*, DOI: 10.1111/mice.12116.

BEARFIELD, G. AND MARSH, W. (2005). Generalising event trees using Bayesian networks with a case study of train derailment. Lecture Notes in Computer Sciences, 3688:52–66.

BJORKMAN, P. (2011). Probabilistic Safety Assessment using Quantitative Analysis Techniques – Application in the Heavy Automotive Industry. *PhD thesis, Uppsala University*, Sweden.

CASTILLO, E., CALVIÑO, A., GRANDE, Z., SÁNCHEZ-CAMBRONERO, S., GALLEGO, I., RIVAS, A., AND MENÉNDEZ, J. M. (2016a). A Markovian-Bayesian network for risk analysis of high speed and conventional railway lines integrating human errors. *Computer Aided Civil and Infrastructure Engineering*, 31:193–218.

CASTILLO, E., GALLEGO, I., SÁNCHEZ-CAMBRONERO, S., MENÉNDEZ, J. M., RIVAS, A., NOGAL, M., AND 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., GRANDE, Z., AND CALVIÑO, A. (2016b). Bayesian networks based probabilistic safety analysis for railway lines. *Computer Aided Civil and Infrastructure Engineering*, in press.

CASTILLO, E., GRANDE, Z., MORAGA, P., AND SANCHEZ-VIZCAÍNO, J. (2016c). A time partitioning technique for railway line design and timetable optimization. *Computer Aided Civil and Infrastructure Engineering*, DOI: 10.1111/mice.12194.

CASTILLO, E., GUTIÉRREZ, J., AND HADI, A. (1997). Expert Systems and Probabilistic Network Models. *Springer Verlag*, New York.

DADASHI, N., SCOTT, A., WILSON, J. R., AND MILLS, A. (2013). Rail Human Factors: Supporting reliability, safety and cost reduction. CRC Press, Taylor and Francis, London.

EVANS, A. W. (2011). Fatal train accidents on Europes railways: 1980-2009. *Journal of Accident Analysis and Prevention*, 43(1):391–401.

FUKUYAMA, H., INUTSUKA, F., TACHI, M., AND ISHIGE, T. (2008). Application of risk assessment method in railway. Sociotechnica, 1(5):163–171.

FUQUA, N. B. (2007). The applicability of Markov analysis methods to reliability, maintainability, and safety, volume 10. *START*

GRANDE, Z. (2015). Mathematical Models for Traffic Observation and Prediction and Design and Probabilistic Safety Analysis of Railway Lines. *PhD thesis, School of Civil Engineers. University of Cantabria*, Santander, Spain.

HENLEY, E. J. AND KUMAMOTO, H. (1992). Probabilistic Risk Assessment; Reliability Engineering, Design, and Analysis. *IEEE Press*, New York.

KOKKINGS, S. J. AND 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.

MIYASHITA, N. (2010). 2013 Safety vision. *JR EAST Technical Review*, 15(1):163–171.

MUTTRAM, R. I. (2002). Railway safety's safety risk model. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transi*t, 216(2):71–79.

PEN, F. AND OUYANG, Y. (2014). Optimal clustering of railroad track maintenance jobs. *Computer Aided Civil and Infrastructure Engineering*, 29(4):235–247.

PETERMAN, D. R., FRITTELLI, J., AND MALLET, W. (2009). High speed rail (hsr) in the United States. *CSR Report for Congress R40973, Congress of the US*.

SLOVAK, R., KASSEV, K., STOYTCHEVA, N., IVANOV, E., AND SCHNIEDER, E. (2007). General stochasti modelling for quantitative safety analysis using markov chains and Petri nets. *Information Technologies and Control*, 2:17–30.

TODOROVICH, H. P. AND HAGLER, Y. (2011). High speed rail in america. *Technical report, America 2050*.