Causal Modelling of Lower Consequence Rail Safety Incidents

Dr George Bearfield,

Safety Risk Assessment Manager, RSSB, Block 2 Angel Square, 1 Torrens Street, London EC1V 1NY. Tel: +44 (0)20 3142 5464, Fax: +44 (0)20 3142 5668 Email: george.bearfield@rssb.co.uk

Dr William Marsh

Lecturer in Computer Science, Risk Assessment and Decision Analysis Research (RADAR) Group, School of Electronic Engineering and Computer Science, Queen Mary, University of London, Mile End Road, London E1 4NS. Tel: +44 (0)20 7882 5254. Email: william@dcs.qmul.ac.uk

ABSTRACT: The Safety Risk Model (SRM) is a key source of information for the GB rail industry. It is a structured representation of the 120 hazardous events that can lead to injury or death during the operation of the railway and is used to estimate the risk to passengers, workers and third parties. The SRM includes both rare but high consequence events such as train collisions and more frequent but lower consequence events such as passenger accidents at stations. In aggregate, these lower consequence events make an important contribution to the overall risk, which is measured by a weighted sum of injuries of different severity.

Where possible, the SRM is derived from historical incident data, but the derivation of the model parameters still present challenges, which differ for different subsets of events. High consequence events occur rarely so it is necessary to use expert judgement in detailed models of these incidents. In comparison, the low consequence events occur more frequently, but both records of incidents and the models in the SRM are less detailed. The frequency of these low consequence events is sufficient to allow both the absolute risk and trends in the overall risk to be monitored directly. However, without explicit causal factors in the data or the model, the models are less able to support risk management directly, since this requires estimates of the risk reduction possible from particular interventions and control measures. Moreover, such estimates must be made locally, taking account of the local conditions, and at each location even the low consequence events are infrequent.

In this paper we describe an approach to modelling the causes of low consequence events in a way that supports the management of risk. We show both how to extract more information from the available data and how to make use of expert judgement about contributory factors. Our approach uses Bayesian networks: we argue their advantages over fault and event trees for modelling incidents that have many contributory causes. Finally, we show how the new approach improves safety management, both by estimating the contribution of the underlying causes to this risk and by predicting how possible management interventions and control measures would reduce this risk.

1 INTRODUCTION

There are approximately 2,500 rail stations on the UK rail infrastructure controlled by Network Rail. Health and Safety legislation requires the organisation managing the station premises (either the Train Operating Company or Network Rail) to ensure that they are safe. Incidents are recorded using a national reporting system and guidance is provided to station staff on the measures (e.g. station cleaning regimes) likely to maintain station safety.

One category of station safety incidents is those relating to passengers boarding or alighting from trains; approximately 1000 of these incidents, of varying severity, are recorded every year. This data is used to track the overall level of risk across the whole rail network. At this scale, the large number of (low severity) incidents permits a reliable, trended estimate of the risk (6.35 FWI/year as estimated in 2009) and of the year on year trend of this risk. This analysis is used to assess progress against national targets and can inform national initiatives.

1.1 *Aims*

It would also be desirable to use the data to guide local decision-making, by estimating:
– the risk at specific stations and the profile between stations, and
– the effectiveness of possible risk reductions interventions and controls.

The difficulty of this is immediately apparent: since the number of stations greatly exceeds the number of incidents per year nationally, the number of events per station is low, with many stations where there are no recorded incidents.

1.2 *Key Features of Our Approach*

The key features of our approach to resolving this problem are:
1 Use the records of incidents to estimate the impact of different causal factors on the probability of each type of incident.
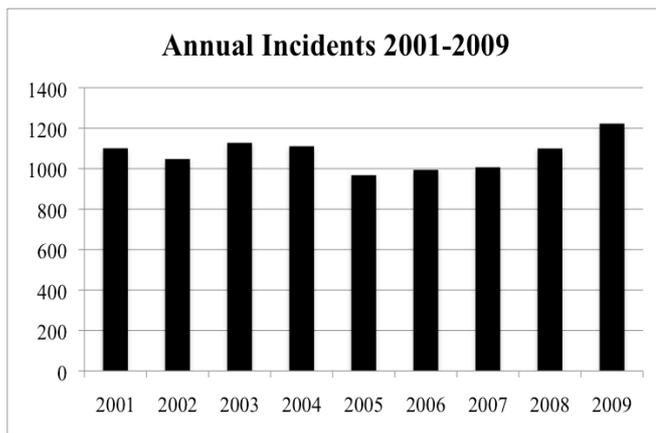
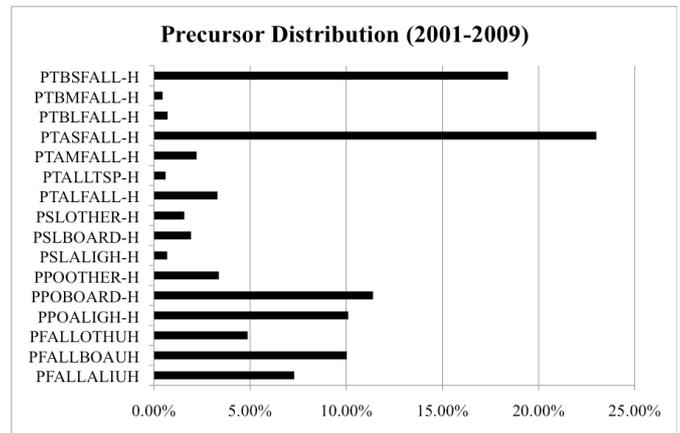Figure 1 Boarding and Alighting Incidents Per Year

Figure 2 Relative Probabilities of Different Precursors

2 Use data on the characteristics and usage of stations on the network to determine the presence of causal factors at each station and the exposure to the risk at each station.

Overall, the model is developed at two levels. The first level uses aggregated incident data to understand the causes of incidents. The second level uses data about stations to complete the model of risk at each station. Although not all the data we would like is available it is clear that much of the data needed has already been collected for other purposes.

1.3 Outline

The remainder of the paper is organised as follows: in section 2 we describe the incident data and the way it is used in the existing Safety Risk Model (SRM). Section 3 outlines the proposed causal model for train boarding and alighting events, explaining how the different data sources are used to build the model. This is followed, in section 4, by a description of the way that the causal model can be used as part of a toolkit for local decision-making. Conclusions are in the final section.

2 EXISTING ANALYSIS OF LOW CONSEQUENCE EVENTS IN THE SRM

2.1 Data Reporting using SMIS

SMIS (Safety Management Information System) is a national reporting system for the GB mainline rail network. SMIS is used by all Railway Group members (including Network Rail, the Train Operating Companies, Freight Operating Companies and Infrastructure Contractors) to record all safety related events that occur on Network Rail controlled infrastructure. The primary requirement for safety incident reporting arises from the Reporting of Injuries, Diseases and Dangerous Occurrences Regulations 1995 (RIDDOR). Since 2003, SMIS has been a web-based system. Data about station incidents is entered locally before being coded by a dedicated SMIS team.

Figure 1 shows the number of reported incidents per year, for the years 2001 to 2009, covering falls while boarding and alighting from trains and incidents of passengers being trapped in train doors.

The data for boarding and alighting incidents includes:

– Time and date
– Location
– Nature of injury
– Narrative description of the event

Table 1 shows the distribution of severity of events over the years 2001-2009. Severity is measured as FWI (Fatalities and Weighted Injuries) with an FWI = one fatality = 10 major injuries = 200 RIDDOR reportable minor injuries or class 1 shock/traumas = 1,000 non-reportable minor injuries or class 2 shock/traumas.

Table 1: Distribution of Event Severities

Injury Category	FWI	% of Total
Fatal	1.0	0.06%
Major	0.1	3.42%
Shock trauma	0.005	4.55%
Minor Reportable	0.005	17.79%
Minor	0.001	74.18%

2.2 The Safety Risk Model (SRM)

The SRM consists of a series of risk models of 120 hazardous events that, collectively, define the overall level of risk on the mainline railway in Great Britain. The reported risk estimates relate to the network-wide risk situation, and they indicate the current level of 'residual risk' (i.e. the level of risk remaining with the current risk control measures in place and with their current degree of effectiveness). The risk associated with a particular hazardous event is calculated as: Risk = Frequency x Consequence

In this calculation 'frequency' is the estimated number of events occurring across the network per year and 'consequence' is the average consequences
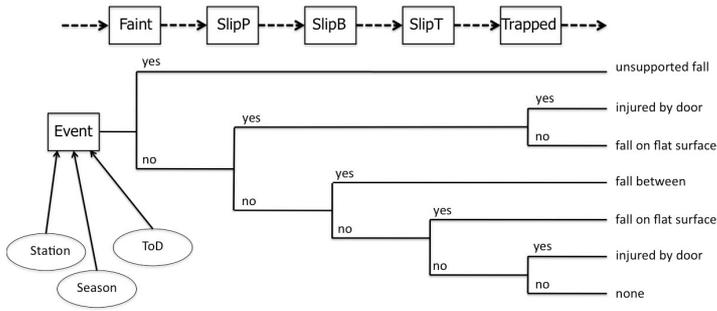
Figure 3 The Events and Top-Level Factors

Table 2: Events in the model

Event	Description
Faint	The passenger faints
SlipP	The passenger slips/falls on the platform
SlipB	The passenger falls between the train and the platform
SlipT	The passenger slips/falls on the train
Trapped	The passenger becomes trapped in the train door

(in FWI) that are judged to occur should the hazard-ous event occur, so that the risk is therefore the FWI expected per year.

Low severity events, where the number of re-ported incidents is relatively high, are analysed as follows:

1 The incidents are first assigned an appropriate precursor, combining the person type, the primary cause of the incident and/or the location in which an incident occurs.

2 A current incident rate for each precursor is judged by smoothing and where appropriate ex-trapolating the year-on-year trends.

3 Precursors with similar accident consequences are grouped together and severity estimates for these groups are calculated by averaging the conse-quences over a number of events, over a number (typically 5) of years.

Figure 2 shows the distribution of the precursors used to categorise incidents of boarding and alight-ing from trains. The most frequent precursors are:

– PTBSFALL-H: Passenger injury while boarding stationary train
– PTASFALL-H: Passenger injury while alighting stationary train
– PPOBORAD-H: Train door closes on passenger (non-slam door stock) whilst boarding

3 A CAUSAL MODEL OF TRAIN BOARDING AND ALIGHTING

In this section we outline a causal model which can help manage (rather than just measure) the risk from train boarding and alighting incidents. The causal model is formed from three layers:

1 Events with distinct causes
2 Causal factors, influencing events
3 Top-level factors, determining usage and influen-cing causal factors.

We describe each of the layers in turn.

3.1 Events

We classify the incidents into 'events', which have distinct causes. The events are shown in Table 2.

Figure 3 shows the events; we use the technique of Marsh & Bearfield (2008) to present the model as a generalised Event Tree, representing a Bayesian Network.

3.2 Top-level factors

The top-level factors are also shown in Figure 3. The factors jointly partition the event that exposes the passenger to risk: 'boarding or alighting from a train'.

Top-Level Factor	Values	Description
Station	There are 2,518 stations on the GB rail network	
Season	Winter / Spring / Summer / Autumn	The season influences the weather.
Time of Day	Rush hour / Nor-mal	Train usage is higher at some stations during the rush hour, making crowding more likely.

We need to determine the proportion of the events occurring at each combination of station, season and time of day. The Office of Rail Regula-tion (ORR) (2010) provides station usage data that can be used for this purpose.

3.3 Causal factors, influencing events

The causal factors are shown in Figure 4. For exam-ple, the causes of falling between the train and plat-form (event SlipB) are judged to be:

– Vertical gap: the difference in height of the plat-form surface and the train floor.
– Curvature: the platform curvature, which creates a larger horizontal gap from the edge of the plat-form to the train.
– Behaviour: passengers who are intoxicated or rushing are more likely to fall.

3.4 Determining probabilities.

The probabilities in the model are completed by ana-lysing the data in the following steps:

1 **The events**. We estimate p(Faint), p(SlipP), p(SlipB), p(SlipT) and p(Trapped | SlipP) from the aggregated data. This requires each incident to be allocated to one of these categories: inci-dents in each category are counted and the num-ber of events divided by the number of boarding (alighting) events. We are exploring the use of search techniques to look for keywords in the un-structured part of the SMIS data record to partly automate this classification.
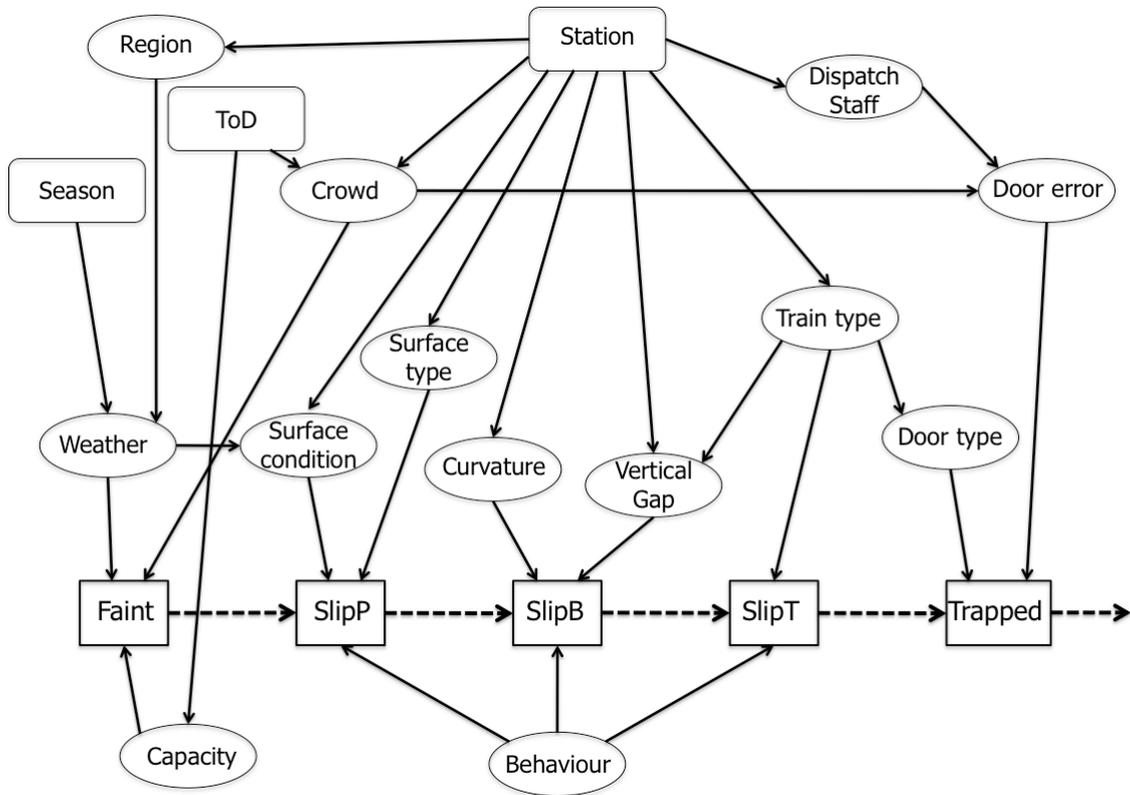
Figure 4 The Complete Bayesian Network Model

2. **The causes of events**. We need to go from p(Faint) to p(Faint | Weather, Capacity). Suppose there are 10 Faint incidents in 10,000 boarding events. We allocate these 10 events to both the Weather and the Capacity categories (there are 4 of each). We then evaluate:

$$p(Fa\,\text{int} \mid Weather, Capacity)$$

$$\propto p(Weather \mid Fa\,\text{int}).p(Capacity \mid Fa\,\text{int}).p(Fa\,\text{int})$$

Suppose Weather is 'normal' or 'hot' or 'icy'. The value p(Weather=hot | Faint) is a table:

Weather	Faint	No faint
normal	proportion of faints in normal weather	proportion of rail journeys in normal weather
hot	proportion of faints in hot weather	proportion of rail journeys in hot weather
icy	proportion of faints in icy weather	proportion of rail journeys in icy weather
	total =1.0	total = 1.0

3. **The station properties**. We need to determine the following for each station: Region, Crowd, Surface condition, Surface type, Curvature, Vertical gap, Train type and Dispatch staff. Some of these are deterministic (e.g. Region). Train type will often be deterministic, except for stations used by a variety of train types. We are examining a variety of data sources for this data, but expect to supplement it with some expert elicitation.

4. **The weather and season**. Data is needed for weather and season, by region. We will connect this to the incidents using dates and time of day.

4 EXTENDING THE MODEL TO A TOOLKIT

4.1 *Requirements*

Legal requirements placed on organisations operating in the GB railway industry mandate some degree of risk assessment. Sections 2, 3 and 4 of Health and Safety at Work Act 1974 (HSWA) (HMSO 1974) require all employers, including railway companies, to manage safety 'so far as is reasonably practicable'. The duty is not just to identify the risks inherent to the company's work activity, but also minimise them as far as reasonably practicable. Determining whether an action is reasonably practicable involves balancing its risks, costs and benefits. In order to determine what measures might be required under the law, there is therefore a requirement to be able to estimate the likely risk reduction achieved by a particular measure to combat the risk from boarding and alighting incidents, in order to weigh this against the cost (Bearfield 2009). In order to make a comparison of risks with costs, the risk needs to be translated into a financial value. This is done using the industry 'Value of Preventing a Fatality' (VPF), a figure endorsed for use by the Department for Transport, which is currently £1.6 million per FWI averted.

Companies might also consider a broader requirement for applying measures, above and beyond any actual legal requirement, taking into account for example the possible impact of incidents on the rail company's reputation and revenue, or the possibility

of civil lawsuits from injured passengers. To support decisions about how to intervene to manage this risk a toolkit is therefore needed which:

– Allows the risk reduction achieved by a range of measures to be estimated.
– Supports the comparison of the costs of the measure with the risk reduction it achieves.

4.2 *Use of the Causal Model*

The proposed causal model can be used in two ways:

1 Profiling the risk at different locations to set priorities for interventions to reduce risk.
2 Estimating the benefit of possible risk reduction measures.

A range of measures could be considered, such as:

– An appropriate cleaning regime: cleaning away contamination improves the slip resistance of a floor, but cleaning itself creates hazards, for example by leaving excess water on a floor surface.
– Winter precautions, such as the use of de-icing chemicals or grit to combat icy surfaces.
– Prevention of tripping hazards: procedures for ensuring that any trip hazards are prevented from occurring or removed in a timely fashion.
– Application of robust train dispatch procedures to minimise the risk of passengers falling while boarding or being trapped in closing doors.
– Civil engineering work to raise the height of platforms.

The potential cost of each of these measures varies significantly, as does their likely impact on train boarding and alighting risk. These various measures target different causes of boarding incidents. For the causal model to be used to analyse such interventions, it is necessary for it to be sufficiently detailed that the precise causes targeted by each measure are specifically modelled allowing 'before' and 'after' estimates of risk to be calculated.

There are precedents for such types of models in the GB railway industry. For example the 'All Level Crossing Risk Model' (ALCRM) (RSSB 2007) allows risks to be estimated for the full range of different types of level crossing and incorporates cost-benefit analysis and 'what-if' functionality. The model outlined here does not include the consequence estimates, but it can be extended following the approach described in Marsh & Bearfield (2009).

5 CONCLUSIONS

We have outlined a way to enhance the modelling of low consequence incidents, intended to support the management of these risks, using the example of train boarding and alighting. The approach has the potential to be applied to a wide range of risks where data is available, but the number of incidents and the detail of the records does not allow direct estimation of the frequency and consequence of events at the level at which safety management interventions are made, conditioned on the factors that can be varied.

5.1 *Related Work*

The problems caused by excessive crowding on trains have been investigated (RSSB 2008). The issues ranged from how crowding is defined by the industry, to developing a consistent approach across the railway network to the controls that might be put in place by station and train operators.

Although the problem of falling from a platform when boarding or alighting a train are particular to the railway industry the problem of managing the risk from slips, trips and falls more generally is well known. The UK Health and Safety Executive (HSE) provides a slip assessment tool which can be used to gather relevant information concerning floor surface properties, contamination, cleaning regimes, footwear to calculate friction coefficients for different surfaces. This approach is useful in known problem areas, but is less useful for identifying areas of key risk and benchmarking performance.

Statistical analysis of the causes of slips, trips and falls has also been undertaken for postal workers (Bentley, 1998).

REFERENCES

Bearfield, G. J. (2009) 'Achieving clarity in the requirements and practice for taking safe decisions in the railway industry in Great Britain' Journal of Risk Research, Volume 12, Issue 3 pages 443-453 Routledge.

Bentley, T.A. (1998) 'Slip, trip and fall accidents occurring during the delivery of mail' Ergonomics, Volume 41, Issue 12 December 1998, pages 1859 - 1872

D. W. R. Marsh & G. Bearfield (2008), 'Generalizing event trees using Bayesian networks' in Proc. IMechE, PartO: J. Risk and Reliability vol. 222, no. O2, pp. 105-114

D.W.R Marsh & G.J. Bearfield (2009), 'Why risk models should be parameterised', Modelling And Analysis Of Safety And Risk In Complex Systems (MASR 2009), St Petersburg, Russia.

HMSO (1974). Health and Safety at Work etc Act 1974. (Elizabeth II 1974. Chapter 37), Her Majesty's Stationary Office.

RSSB (2007) T737: Development of the All Level Crossing Risk Model – A History, 1993-2007. Available from http://www.rssb.co.uk/

RSSB (2008) T605: Management of on train crowding. Available from http://www.rssb.co.uk/

Safety Management Information System (SMIS) – Safety Performance and Risk – RSSB http://www.rssb.co.uk/safety/spr/smis.asp (retrieved April 2010)

ORR (2010) Station Usage: Rail Statistics: Publications and statistics: Office of Rail Regulation. http://www.rail-reg.gov.uk/server/show/nav.1529 (retrieved April 2010)