ELSEVIER

Using an integrated process of data and modeling in HRA

John Wreathall[a],[*], Dennis Bley[b], Emilie Roth[c], Jordan Multer[d], Thomas Raslear[e]

[a]The WreathWood Group, 4157 MacDuff Way, Dublin, OH 43016, USA
[b]The WreathWood Group, 11738 English Mill Court, Oakton, VA 22124, USA
[c]Roth Cognitive Engineering, 89 Rawson Road, Brookline, MA 02445, USA
[d]John A. Volpe National Transportation Systems Center, 55 Broadway, Cambridge, MA 02142, USA
[e]Federal Railroad Administration, Office of Research and Development, 1120 Vermont Avenue, NW, Washington, DC 20590, USA

Abstract

The paper describes an approach taken to estimate the probabilities of failure associated with various railroad tasks to prevent accidents (principally collisions and derailments). These probabilities were estimated using an expert elicitation process that used partially relevant data available from a variety of databases and that were filtered and scaled to make them more directly relevant to the analyses being performed. Extensive qualitative studies were performed prior to the elicitation process to identify relevant contexts under which the tasks can be performed.
© 2003 Elsevier Ltd. All rights reserved.

Keywords: Railroad; Human reliability analysis; Data; Expert elicitation; Qualitative analysis

1. Introduction

The railroad industry in the United States is in the process of developing positive train control (PTC) systems primarily intended to improve the throughput and safety of railroad operations. See, for example, Ref. [1] for a summary of PTC concepts and systems.

PTC systems address three core safety-related functions:

- Preventing train-to-train collisions
- Enforcing speed restrictions and temporary slow orders
- Providing protection for roadway workers and their equipment

The complexity of these technologies (particularly communication and information technologies) requires additional safety considerations that current FRA regulations do not address. As a consequence the US Federal Railroad Administration (FRA) has developed a proposed rule to assess the safety of these systems using (among other assessment methods) the results of probabilistic risk assessments (PRAs) [2].

The proposed rule adopts a performance-based approach to enable flexibility in the design and implementation of PTC systems while providing a mechanism to achieve

safety goals. The performance standard adopted in the rule requires that the new product or system must not degrade safety below the level of the existing system. To evaluate whether this condition is met requires a risk assessment comparing the new system to the system it will replace.

This proposed rule would require that any railroad wishing to use a processor-based control system (such as positive train control (PTC) systems) to provide more effective or efficient control of train movements must submit a product safety plan (PSP) that includes a quantitative risk assessment that compares the mean time to hazardous events (MTTHE) for related railroad operations with and without use of the processor-based control system to show that there would be no reduction in safety from implementing the system. The proposed rule also requires that MTTHE values must incorporate the impact of all elements of the system, including human factors as well as the hardware and software components. While this rule is not final, it is considered very likely that the final rule will contain the same conceptual requirements for performing a quantitative risk assessment as part of the PSP.

This paper focuses on some techniques for estimating the probabilities of human errors associated with train operations where partially relevant data are available but do not directly correspond with the events being modeled. It is noted that such conditions often apply in modeling human errors in other technical settings, including nuclear power

* Corresponding author.
 E-mail address: jwreatha@rrohio.com (J. Wreathall).

applications. A full description of the work is presented in Ref. [3].

2. Train control, unsafe actions and safety in rail operations

The purpose of the human reliability analyses in this study was to estimate the likelihood of specific unsafe human actions. Unsafe actions include actions not being taken that could prevent hazardous events, as well as actions being taken that may cause hazardous events (by themselves or in combination with other conditions). The particular system being modeled is a new computer-based system for train monitoring and control, TMC, that is being developed by a large North American freight railroad. The TMC system is a form of train control that provides a warning to the locomotive crew when the train is predicted to exceed the limits of its authority[1] and stops the train if the engineer fails to act in time. The system is intended to operate in 'dark territory' where authority for train movements and track occupancy is accomplished by verbal exchanges between the dispatcher and train crew (i.e. no signal displays are used—hence the term dark territory—and, for the most part, a single track with sidings is used for traffic in both directions). Railroad-specific operating rules govern the exchange of information between the dispatcher and the train crew.

Should errors occur, for example misunderstandings about the limits of authority during radio communications between the dispatcher and the train crew, TMC provides an additional layer of defense. TMC receives information regarding the authorized train movements from a computer-aided dispatch system used by the dispatcher to indicate valid track occupancy and compares this information with the current train position (using global positioning systems (GPS) data) to determine whether the train is operating within its authority. This system is 'overlaid' over the existing manual train control system; should it fail, protection is still provided by the existing rules of operation.

One of the uses of human reliability analysis (HRA) proposed by FRA is in support of risk-based evaluation of new technology. When a new technology is introduced that requires human interaction, one cannot evaluate the performance of the new technology in isolation. Rather, the role that the human may play needs to be considered in either enhancing the overall performance or degrading it. This requires performing analyses to quantify risk in the base case (with existing technology) and then comparing this risk to the estimated risk once the new technology is introduced.

The analyses described in this paper were those performed for the existing system for which operating experience and some data exist. Four kinds of conditions involving unsafe actions were selected by FRA and the analysts performing a broader scope probabilistic risk assessment (PRA) to be the subject of the HRA—these are conditions that the TMC has the potential to prevent. These are:

1. Train enters a block without authorization. Entering a block without authorization creates the potential for a collision with a train already occupying the block.
2. Train exceeds the track speed limit. This creates the potential for derailing the train and damaging the track, and for creating a release of hazardous materials to the environment, the workforce, and the public (possibly requiring an evacuation).
3. Train enters a preplanned work zone without authorization. This has the potential for maintenance workers or equipment being struck, or the train entering track that is not intact.
4. Train crosses a misaligned switch. This has the potential for derailing the train and damaging the track, plus the risk of releases of hazardous materials.

Each of these conditions can result from unsafe actions. For example, the train could enter a block without authorization because the train crew were not paying attention to their location and inadvertently passed the end of their authorization. Alternatively the crew may believe they have authorization (for example, because of poor radio communications with the dispatcher) for the block being entered. The dispatcher may provide verbal authorization but not enter the data in the computer system correctly. Also the crew may miss the block boundary location because of very poor visibility or because the boundary sign has been vandalized or damaged. As would be expected, system operating rules and design features reduce the potential for many of these conditions. For example, operation of the dispatch system and the operating rules for communications between the dispatcher and the train crew require double checking of the authorization, to ensure the crew read back correctly the authorization as entered in computer by the dispatcher. Crews are very familiar with the territory and know the block boundary locations even if the signs are missing or damaged.

3. Human reliability analysis

This analysis involved performing both qualitative and quantitative studies to define and evaluate the relevant unsafe actions, the contexts in which the unsafe actions could occur, and the probabilities of the unsafe actions in

---

[1] A train's authority is the permission, generally given by a central dispatcher under the operating rules of the company, to occupy a particular section of track (referred to as a 'block'). At the end of that authority, the train crew must request, and be given, a new authority before they may move into the next block.

these contexts. The quantitative analysis of these unsafe actions was performed primarily by an expert elicitation process using a workshop involving representatives from all the participating technical and management organizations.

For each of the conditions, a detailed qualitative analysis was performed that involved two aspects: (1) an analysis of the current work environment to understand the types of errors that can arise and the factors that contribute to those errors; and (2) an examination of the proposed TMC system, its user interface and proposed human-system interaction, to assess its potential impact on human performance and human reliability. This analysis involved structured interviews and visits to observe both the current operations and field testing of a prototype of the TMC system.

The focus of the interviews and observations included:

- What are the most likely forms of human error in the current railroad operations in the segment of railroad territory where tests of the TMC were being conducted (i.e. the base case)?
- What are the factors that are most likely to contribute to those errors?
- What recovery mechanisms do humans provide that contribute to a robust, high-reliability system?
- What impact would TMC likely have on human reliability and overall safety?
- Could TMC prevent and/or catch and recover from the types of human errors that are known to occur in the base case?
- Would TMC change how the people in the system perform (i.e. locomotive engineers, conductors, dispatchers)?
- Could TMC introduce any new sources of risk? If so, are there mechanisms available to enable the people in the system (e.g. the locomotive engineer, the conductor, the dispatcher) to catch and recover from the TMC-induced 'errors'?

The primary tasks in the quantitative analysis of the HRA were the identification of relevant sources of data, their limitations and gaps, and application of the expert elicitation process used to compensate for these limitations and gaps in the final quantification steps.

3.1. Overall analytical process for quantification

The overall analytical process for quantifying the probabilities of each of the human error events was to answer the following five questions. The first three were answered in large part before the quantification process was started, by defining the scope of the analysis and in the discussions undertaken as part of the qualitative analyses.

- *Identify major events to quantify*. For example, the train exceeds its limits of authority. This could be the result of

two different general human errors—errors by the train crew and errors by the dispatcher.
- *What is the scope?* For each major event identified in the previous step, what kinds of errors does that event include? In the case of the train crew, the scope would simply be the crew fails to stop the train at its limit of authority because of (for example) failing to notice they have reached the end of the last block or erroneously recalling what is the limit of their authority.
- *What kinds of things could cause the errors listed in the previous step?* A partial list of what could cause failure of the train crew to stop at the limit of the authority includes:
  ○ Inattention or failure to recognize their location
  ○ Erroneous recall of authority limits
  ○ Distraction (within the cab or outside the cab)
  ○ Over-reliance on another crew member
- Misjudged braking performance
- *What data exist?* To what degree do the available databases relate to the events being modeled? Do they include all or most of the identified significant causes?
- *What judgments are needed?* Are there additional causes not included or under-reported in the databases that are relevant to the analysis? For example, are all the causes listed above (inattention, etc.) included in the train crew disciplinary database or the FRA incident database? Are there additional causes in these databases that should be excluded because they do not related to causes being modeled? On what basis can these data be filtered and scaled?

Based on the combination of the databases and the judgments, the final probability parameters (usually in the form of probability distributions) for the human error events are estimated. This overall process is shown in Fig. 1.

3.2. Sources of data

Two kinds of data are typically required in HRA studies: information about the numbers of events that have occurred that are similar to those being modeled and information about the number of opportunities for such events (operational experience base), such that a probability of the events can be estimated. Two major sources of data were identified in this study: the databases maintained by the FRA, and databases maintained by the freight railroad. These sources of data contain information about both the frequencies of events and the opportunities for such events.

3.2.1. FRA Databases
FRA maintains a database [4] that contains coded summaries of incidents in railroads that meet the FRA reporting requirements. These summaries identify, among
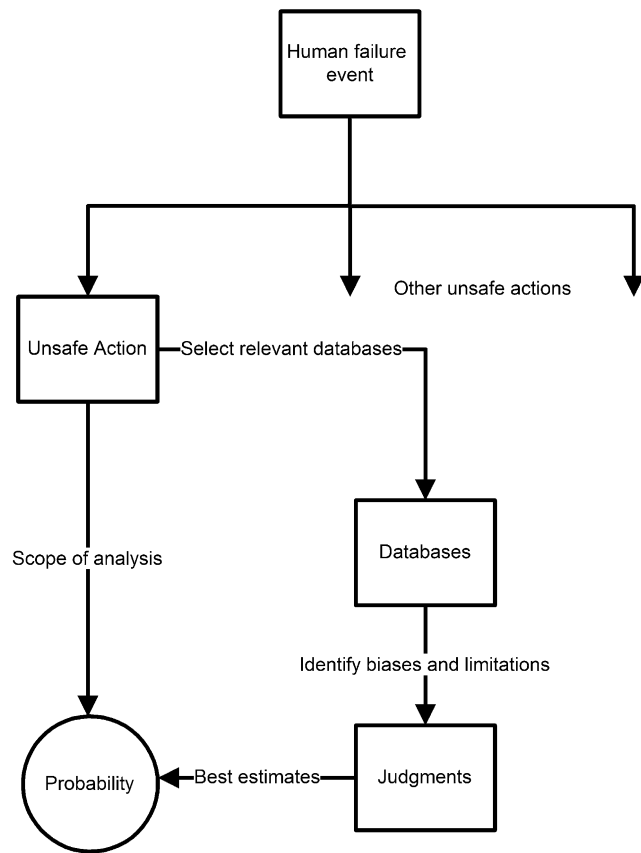
Fig. 1. Overall analytical process.

other things, the railroad(s) and locomotive identifiers involved, the date and location, the type of traffic (passenger, freight, etc.) and a set of cause codes for the event, based on the investigation made by the railroad.

The FRA also maintains an operational experience database, available at [4]. This database summarizes the amounts of train movements (expressed in train-miles) for each railroad, separated by train-miles in yards versus track, passenger versus freight, etc. Totals per calendar month are provided.

3.2.2. Freight railroad databases

The freight railroad identified several incident databases suitable for this analysis. The first was a summary of events that occurred on the test territory in the nine years preceding the study. Each event record included a summary of the type of event, and whether it was hardware-, human-, roadway-or other-related. In total, there were 89 events, of which the largest contribution was from roadway equipment problems. A total of 24 events involved a human-related cause.

The second set of incident databases was associated with disciplinary actions taken by the freight railroad. Two were provided: one associated with train crew disciplinary actions and one with dispatcher actions. No

individuals were identified in either database. Both databases provided a brief summary of the event (either the rule violated (dispatcher data) or the type of event (train crew data)) and the date. From these data, it was possible to identify for example, if the action was a track segment violation (that is, exceeding the limits of their authority) for a train crew, or if the action was a violation of a rule concerning issuing a block authority inappropriately by a dispatcher.

In the case of operational experience that can be used as a basis for estimating the number of opportunities for the unsafe action of interest, the freight railroad provided a set of 'raw' data for the test territory: a set of all movement authorities for the territory covering a 2-week period. This 2-week period was considered to typify operations for the test territory. For this 2-week period, the data allowed the following items to be identified:

- The number of trains traveling the territory
- The number of authorities that were issued
- The time and the number of blocks issued or released for each train
- The number and duration of temporary work zone restrictions in place
- The number of track inspections occurring at any time

Table 1 illustrates data sources and their relevance in estimating probability of unsafe human actions for select examples; a complete set of similar tables was generated for all relevant actions in the actual studies.

3.3. Limitations and gaps in the databases

Each database has certain limitations and gaps with regard to the events being analyzed. These are briefly summarized below. Those limitations were resolved through interactive sessions with experts in railroad operations.

3.3.1. FRA databases

Incident database. There are two primary limitations in this database: events to be recorded must meet certain damage or injury criteria as set forth by FRA,[2] and only a limited set of causal information is required to be provided by the reporting railroad. As a result of the first, there is a potentially significant gap of event information for which no accident occurred-there is no 'near miss' reporting for events involving unsafe actions but no (or only minimal) consequence.

Because almost any accident is the result of multiple causes, the ways in which an event is reported can be somewhat subjective as to what is given a primary focus—equipment or human. Therefore relying only on

---

[2] See FRA Instructions for preparing FRA Form F 6180.54, Rail Equipment Accident/Incident Form—see [4].

Table 1
Example data sources for unsafe actions

Event	Event data	Unsafe human actions	Unsafe action data	Operating experience
Train fails to stop at boundary of authority	Rail company incident data	Train crew fails to stop at block boundary at end of authority	Employee disciplinary actions database	Rail company total and yard miles (FRA)
	FRA incident data	Block sign present		
		Block sign missing Dispatcher fails to protect train authority in CADS	Dispatcher disciplinary actions database	Rail company total and yard miles (FRA)
		Train crew mishears dispatcher as to limit of authority (location, train ID)	Estimates (e.g., estimates obtained from train crews or dispatchers	Number of train movements in territory
		Dispatcher mishears train crew request for authority (location, train ID)	Estimates (e.g. estimates obtained from train crews or dispatchers)	Number of train movements in territory

the cause codes of the events does not provide a sufficient basis for identifying events relevant to this study. The reports do provide the opportunity for presenting a narrative for additional information but there can be quite significant latitude in the way events are reported. However, the combination of types of events and the narratives seems to provide at least a basic starting point for identifying relevant events.

Operational experience database. The FRA operational database provides a basis for estimating total train movements within a given railroad, but the categories only describe whether the movements were in yards or out of yards. There is no distinction between the types of train control system in use (e.g. whether it is dark territory), or any information about traffic within specific territories.

3.3.2. Freight railroad databases

Incident databases. The small number of events limited the freight railroad database associated with incidents in the test territory. The databases associated with disciplinary actions were limited largely by the fact that that an unknown number of similar events could occur, but without any mechanism to detect and report the event outside of the crew involved, and the absence of any incentive to self-report such events, this number cannot be known for certain. On the other hand, some disciplinary events (particularly for train crews) could be the result of performance testing that is more rigorous than (and not representative of) normal operations. Therefore, the potential exists for both underestimating (from unseen events) and overestimating (from the inclusion of non-representative testing) from these databases.

Operational experience database. The details of the traffic were sufficient to identify the total numbers of trains, the numbers of blocks issued, the amount of

maintenance work, etc., in the test territory for the 2-week period. The only question is the extent to which the 2 week period was representative of traffic overall in the test location.

3.4. Expert elicitation process: Quantification workshop

In order to compensate for the limitations and gaps in the above databases, the data need to be filtered and scaled. To perform these adjustments, a 2-day elicitation workshop was held to obtain adjustments to the data represented in the databases available for quantification. A total of 30 attendees participated in the 2-day workshop, including representatives of the railroads, engineers, conductors and dispatchers.

A facilitator who was familiar with the event being analyzed, the results of the qualitative analysis, the types of databases available, and the expertise in the group led the elicitation process for each human error event. Preliminary analyses were performed to identify talking points in the facilitated groups, such as possible causes of human error events (based on the results of the qualitative analyses described earlier), examples of the databases and their possible limitations, and questions to help develop distributions associated with the probabilities being estimated (such as 'How high and how low could the end points of the distributions be, and why?'). The facilitator led the discussion through the items as the list of causes for the events and the available databases to ensure that no significant contributors or sources of information had been overlooked, and that everyone in the group had a common understanding of the events and factors being analyzed. In all cases, an extensive discussion ensued that often clarified the details of the actions necessary for the event to occur, anecdotal examples of near-misses that participants had

witnessed or participated in, and issues associated with the content of the databases.

Following this discussion, the facilitator led the discussion to actually estimate the parameters of the model, including the uncertainty. Two approaches were needed, one for cases where data were available, but not quite appropriate and a second for cases where no data could be found. In the first type of discussion for estimating what adjustments were necessary to make the database most relevant, the group considered such issues as which events needed to be excluded, and where under-or over-reporting could occur. Based on these discussions, the facilitator led those group members who had a working knowledge of the situation to estimate adjustments to the parameters developed from the database. First, any events in the database that were not relevant to the scenario were removed (i.e. *filtering*). An example was to filter out events that could only occur in a switching yard where the analysis was only for events occurring on main tracks. Second, adjustments were made, for example, for the potential for under-reporting, as with disciplinary data where people would be unlikely to self-report their own errors (i.e. *scaling*).

Following these adjustments, probability distributions were estimated using a combination of data and elicited judgment. In almost all cases, these values were obtained by eliciting the endpoints of the event occurrence distribution (from using the 'How high could this event frequency be?', and 'How low…?' questions, as well as the experts' bases for their opinions) and then estimating the shape and the resulting mean of the distribution.

In the second case, where no database applied, the people with hands-on experience were asked to think through how relevant situations could occur, what factors would be most important, and then to focus on the extreme values—what is the most often it could occur, and the least often? In some cases the facilitator had to stimulate the discussion, saying, 'From what you have discussed, it appears that the high and low values must be….' and obtaining consensus through discussion.

4. Example analysis

This example illustrates just one calculation to show how the process described above was used to estimate one parameter: the rate at which locomotive crews exceed their block authority in dark territory operations (an unsafe action that can result in a collision if another train or maintenance equipment is already in the new block). This analysis uses two sources of data to provide this estimate: the freight company's employee disciplinary database, and the train operations data reported to FRA. The analysis process for this example is shown in Fig. 2.

The rate of train crews exceeding their authorized block boundary in dark territory is therefore calculated basically by dividing the number of disciplinary actions related to
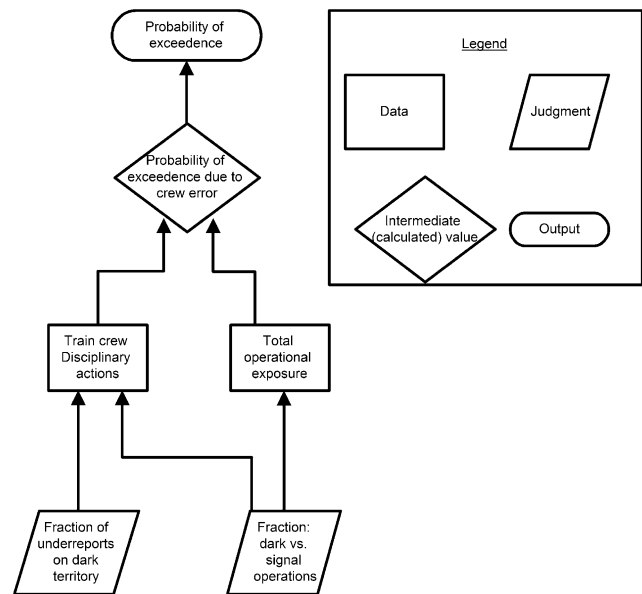

Fig. 2. Example analysis process for train crew exceeds block authority.

dark territory operations by the number of train-miles of dark territory exposure, as shown in Fig. 2. However, the disciplinary events are reported for all track miles. Therefore these events must be scaled for just dark operations. It was the opinion of the attendees that the total event rate could best be proportioned in the ratio of the dark to signal miles. Similarly the total operational exposure (train miles) needs to be proportioned in the ratio of train control modes.

With regard to the crew disciplinary actions, the relevant category in the database is the number of 'track segments' violations, of which 91 occurred in the four year period of the database (1997–2000). Using only these actions represents an example of the filtering described above. The workshop participants agreed that the last 4 years was largely representative of the current operations. It was felt that conditions prior to this period were less likely to be representative of current conditions (for example, because of updates in the rule book, company mergers, etc.) and therefore using the data for just this period is appropriate.

One simple adjustment was to assess whether the 91 track segment violations in the disciplinary database represented an under- or over-reporting of the events in practice. The focus of this analysis was a discussion with the group members. It was generally agreed that the events in the database under-represented the events in practice. The experts in the group—the train crew members and those familiar with the railroad's disciplinary process—were asked how low and how high could the under-estimation be, using the elicitation process described above. The lowest under-reporting was estimated to be 5%, and the under-reporting could be as high as 20%. All intermediate points were judged equally likely, thus creating a flat distribution between the limits of 1.05 and 1.20 for the scaling the number of events in the database.

The 91 track segment violations could have occurred on signal territory, dark territory or in yards—no information is provided in the database to distinguish between these locations. Based on discussions within the group, it was agreed that these 91 events could be distributed on the basis of the train mileage associated with each of the territories and yard traffic.

The workshop participants agreed that the basis for these different train miles could be estimated from a combination of two sources of data: the FRA operating experience database (which provides train mileage data associated with track separately from yards), and the relative track lengths associated with signal, dark territory, and other operations. The data from the yard versus track traffic over the 4 years was analyzed using the FRA data, with the results being shown in the following Table 2.

Based on data provided by FRA, the following were found to be the track lengths associated with different modes of operation (Table 3).

While the 'unknown' category in Table 3 only represents less than 10% of the total track length, it would represent just over 25% of the dark territory if all the unknown operations were, in fact, dark. The reason for it being unknown is largely because of it being made up of short lengths of track, often in connection with industrial (private) tracks, and for which centralized records were not readily available. If the uncertainty with the operating modes in this track was very important, more detailed surveys would need to be sought, which could take significant resources and time. In this analysis, it was concluded that treating the operating modes of the lengths of unknown track as a source of uncertainty would be sufficient for the purposes of this study.

In order to model this as a source of uncertainty, the extremes, of the unknown track being either all signal or all dark, were taken as the endpoints of a distribution. However, this was considered unlikely by the workshop participants. The most likely condition was that the unknown miles would be in proportion to the known miles identified as signal or dark. Automatic train control (ATC) is used very rarely and its locations are well known. Yard miles are also well known and the unknown miles can be excluded from the yards. Therefore the operating modes for the unknown miles are only either signal or dark.

Thus, the probability density function shown in Fig. 3 was created to represent the fraction of the total system

Table 3
Track lengths for different operating modes of freight company (FRA data)

Operating mode	Length (miles)	Fraction (%)
Yard	2963	13.6
ATC	75	0.3
Signal	10,560	48.4
Dark	6,072	27.8
Unknown	2,164	9.9
Total	21,834	

miles that represent dark operations in the unknown territory. That is, if the unknown miles are all signal, then the dark territory would represent 27.8% of the total miles; if all the unknown miles are dark, then the total dark miles would be (27.8 + 9.9), or 37.6% of the total miles. The 'most likely' value for the allocation of the unknown track lengths was discussed by the attendees, and it was agreed that that it would be best represented in the proportion of the known track lengths for the signal and dark territories. That is, the peak of the distribution is located between the lower and upper limits. Given the above upper and lower limits, the peak lies at 31.3%, as shown in Fig. 3.

Using the above data and judgments represented in the distributions, the mean rate of exceedance is 3.1×10^{-7}/train mile, with a range of 2.9–3.6×10^{-7}/train-mile. This analysis was performed using a commercial software add-in to Microsoft® Excel that manipulates discrete data distributions in spread sheets, rather than the single point values normally used in Excel. For this analysis, the distribution is well-represented by a uniform (flat) distribution, with the above limits. The shape of this distribution is strongly influenced by the shape of the distribution used to characterize the range of under-reporting (also uniform). While expressed in units of 'exceedances per train-mile' (since the source of the exposure rate is train-miles), the risk analysis typically uses 'per block boundary' for the hazard rate. The average block length is 6.3 miles, so the corresponding mean exceedance rate is 2.0×10^{-6} per block boundary, with a range of 1.8–2.3×10^{-6} per block boundary.

Table 2
Freight company train miles (from FRA database)

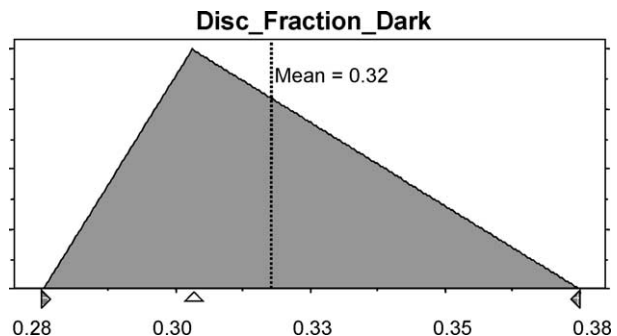| Year	Total train miles	Yard train miles	Non-yard train miles
1997	83,733,024	13,324,933	70,408,901
1998	83,447,524	13,367,246	70,080,278
1999	105,277,723	16,075,426	89,202,297
2000	114,426,120	17,874,254	96,585,866

Fig. 3. Probability distribution of fraction of track miles representing dark operations in the unknown tracks.

5. Findings

The approach taken for HRA was able to generate reasonable results (i.e. acceptable to the workshop participants) despite the fact that there was no directly applicable database. The workshop format permitted experts from many different organizations and backgrounds to work together and reach consensus. Uncertainty was expressed through probability distributions that were accepted by the group. The HRA and PRA teams reached agreement that the HRA results were appropriate for use in the PRA.

The approach taken in this study provides one viable way for others to perform HRA studies in support of the FRA's proposed Standards for Development and Use of Processor-Based Signal and Train Control Systems. The lessons learned from performing this example analysis of the system were documented and provide guidance on avoiding potential pitfalls in future human reliability analyses studies.

Two observations are relevant from our study to other analyses of human reliability, both related to rail and to other industries:

- First, that the analytical situation in this study—having some relevant data but with a variety of limitations (not a perfect match for what we want to estimate, with sources that may lead to both under- and over-estimates of frequency) are far from unique to our case. They happen often and must be addressed explicitly
- Second, the approach we took for combining 'hard data' with expert judgment is a good approach that could be used in other applications. It uses hard data to ground the experts judgments, while using expert judgment to compensate for the known limitations of the existing data.

Recommendations for future analyses of rail HRA studies:

1. Use an HRA team that includes members experienced in performing human factors studies, human reliability analyses, PRAs, and group facilitation.

2. Model human errors at compatible levels in the PRA and HRA tasks, preferably at the level of available data and experience.
3. Verify that the data sources (databases, expert judgment or a combination) are suitable for the tasks and associated errors being analyzed. Identify gaps or mismatches and utilize expert judgment to leverage the available data while compensating for the known limitations of the data.
4. Conduct qualitative task analyses with people experienced in using the existing systems. Activities should include interviews with workers using the existing systems or the target users of the system (in the case of technologies under development), their trainers and supervisors, so that all levels of experience are included.
5. Expert elicitation methods should take into account known biases and other limitations of expert judgment. Experts should express their opinions in terms of ranges rather than single point values.
6. Solicit input from as broad a range of stakeholders as possible so that the analysis takes into account a wide range of perspectives. Accept quantitative inputs only during the elicitation process, from people with relevant operating experience.
7. Ask the broadest range of stakeholders possible to review to the results of the analyses to foster support for the results.

References

[1] Ullman KB, Bing AJ. High speed passenger trains in freight railroad corridors: operations and safety considerations. Washington, DC: US Department of Transportation, Federal Railroad Administration; 1994.
[2] Department of Transportation, 49 CFR Part 209, et al. Standards for development and use of processor-based signal and train control systems; proposed rule (Federal Railroad Administration). Federal Register; 2001. p. 66 (155, Friday, August 10, 2001).
[3] Wreathall J, Roth E, Bley DC, Multer J. Human reliability analysis in support of risk assessment for positive train control. Cambridge, MA: U.S. Department of Transportation, John A. Volpe National Transportation Systems Center; 2002.
[4] Federal Railroad Administration, Office of Safety Analysis Web Site. http://safetydata.fra.dot.gov/OfficeofSafety/