10-2013

Identification of Causal Paths and Prediction of Runway Incursion Risk using Bayesian Belief Networks

Benjamin Jeffry Goodheart
Embry-Riddle Aeronautical University - Worldwide

IDENTIFICATION OF CAUSAL PATHS AND PREDICTION OF RUNWAY INCURSION RISK USING BAYESIAN BELIEF NETWORKS

by

Benjamin Jeffry Goodheart

A Dissertation Submitted to the College of Aviation
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Aviation

Embry-Riddle Aeronautical University
Daytona Beach, Florida
October, 2013

IDENTIFICATION OF CAUSAL PATHS AND PREDICTION OF RUNWAY INCURSION RISK USING BAYESIAN BELIEF NETWORKS

by

Benjamin Jeffry Goodheart

This Dissertation was prepared under the direction of the candidate's Dissertation Committee Chair, Dr. Dothang Truong, Professor, Daytona Beach Campus; and Dissertation Committee Members Dr. Alan Stolzer, Professor, Daytona Beach Campus; Dr. Steven Hampton, Professor, Daytona Beach Campus, and Dr. James Luxhøj, Professor, Rutgers University; and has been approved by the Dissertation Committee. It was submitted to the College of Aviation in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Aviation

Dissertation Committee:

Dothang Truong, Ph.D.
Committee Chair

Alan J. Stolzer, Ph.D.
Committee Member

Steven Hampton, Ed.D.
Committee Member

James Luxhøj, Ph.D.
Committee Member - External

Alan J. Stolzer, Ph.D.
Department Chair, Doctoral Studies

Tim Brady, Ph.D.
Dean, College of Aviation

Michael P. Hickey, Ph.D.
Dean of Research and Graduate Studies

10/31/13
Date

ABSTRACT

Researcher: Benjamin Jeffry Goodheart

Title: IDENTIFICATION OF CAUSAL PATHS AND PREDICTION OF RUNWAY INCURSION RISK USING BAYESIAN BELIEF NETWORKS

Institution: Embry-Riddle Aeronautical University

Degree: Doctor of Philosophy in Aviation

Year: 2013

In the U.S. and worldwide, runway incursions are widely acknowledged as a critical concern for aviation safety. However, despite widespread attempts to reduce the frequency of runway incursions, the rate at which these events occur in the U.S. has steadily risen over the past several years. Attempts to analyze runway incursion causation have been made, but these methods are often limited to investigations of discrete events and do not address the dynamic interactions that lead to breaches of runway safety. While the generally static nature of runway incursion research is understandable given that data are often sparsely available, the unmitigated rate at which runway incursions take place indicates a need for more comprehensive risk models that extend currently available research.

This dissertation summarizes the existing literature, emphasizing the need for cross-domain methods of causation analysis applied to runway incursions in the U.S. and reviewing probabilistic methodologies for reasoning under uncertainty. A holistic modeling technique using Bayesian Belief Networks as a means of interpreting causation even in the presence of sparse data is outlined in three phases: causal factor identification, model development, and expert elicitation, with intended application at the systems or regulatory agency level. Further, the importance of investigating runway incursions probabilistically and incorporating information from human factors, technological, and organizational perspectives is supported. A method for structuring a Bayesian network using quantitative and qualitative event analysis in conjunction with structured expert probability estimation is outlined and results are presented for propagation of evidence through the model as well as for causal analysis.

In this research, advances in the aggregation of runway incursion data are outlined, and a means of combining quantitative and qualitative information is developed. Building upon these data, a method for developing and validating a Bayesian network while maintaining operational transferability is also presented. Further, the body of knowledge is extended with respect to structured expert judgment, as operationalization is combined with elicitation of expert data to create a technique for gathering expert assessments of probability in a computationally compact manner while preserving mathematical accuracy in rank correlation and dependence structure.

The model developed in this study is shown to produce accurate results within the U.S. aviation system, and to provide a dynamic, inferential platform for future evaluation of runway incursion causation. These results in part confirm what is known about runway incursion causation, but more importantly they shed more light on multifaceted causal interactions and do so in a modeling space that allows for causal inference and evaluation of changes to the system in a dynamic setting. Suggestions for future research are also discussed, most prominent of which is that this model allows for robust and flexible assessment of mitigation strategies within a holistic model of runway safety.

ACKNOWLEDGEMENTS

To James B. Crain, my grandfather, and the man to whom I owe my enduring sense of discovery, unrelenting curiosity, and more than occasional stubbornness. Thank you for always pushing me to ask why and teaching me to never give up.

To my wife Cathy, and to my children Isabel and Mateo, thank you for being my source of calm in what has often been a challenging journey for all of us. Without their collective patience, love, and support, this dissertation would never have been. If I have succeeded to any degree, it is because of them. If I have failed, I owe it only to my own shortcomings.

Thank you to Dr. Roger Cooke and to Daniel Ababei and Dr. Anca Hanea for so graciously answering what must have seemed like an infinite barrage of questions.

To my advisor, Dr. Dothang Truong, and to my mentor, Dr. MaryJo Smith, who push me ever forward in search of answers. I am forever grateful for your humor and wisdom through this process, and I am happy to count you as dear friends. Thank you also to my entire committee for their invaluable advice, guidance, friendship, and tolerance throughout the process of creating this research.

My sincere thanks are due the Transportation Research Board and the National Academies of Science. Without their support, this research would have been impossible.

Finally, to Sohail Khan, whose very visceral introduction to physics showed me a world made much more interesting by science, and whose dedication to teaching young people is an inspiring example of the power of a great educator.

To those who I have not mentioned by name, please know that you have my thanks, appreciation, and gratitude for your patience with me as I navigated this research. I could not have done it without you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

A runway incursion (RI) is defined by the International Civil Aviation Organization (ICAO) as, "any occurrence at an aerodrome involving the incorrect presence of an aircraft, vehicle or person on the protected area of a surface designated for the landing and takeoff of aircraft" (EASA, 2011, p. v; ICAO, 2007, p. vii). Effective October 1, 2007, the Federal Aviation Administration (FAA), also adopted the current ICAO definition of a runway incursion in an effort to harmonize global efforts to identify and reduce RI incidents. In the U.S., RIs are classified by four severity categories as shown in Table 1 and by type, as shown in Table 2 (FAA, 2009).

Table 1. *FAA Runway Incursion Severity Classification.*

Category D	Category C	Category B	Category A	Accident
Increasing Severity →				
Incident that meets the definition of runway incursion such as incorrect presence of a single vehicle/person/aircraft on the protected area of a surface designated for the landing and take-off of aircraft but with no immediate safety consequences.	An incident characterized by ample time and/or distance to avoid a collision.	An incident in which separation decreases and there is a significant potential for collision, which may result in a time critical corrective/evasive response to avoid a collision.	A serious incident in which a collision was narrowly avoided.	An incursion that resulted in a collision

Note. Adapted from http://www.faa.gov/airports/runway_safety/news/runway_incursions.

Runway incursions were recently identified as one of aviation's most critical challenges in the 2011 FAA NextGen Implementation Plan, the FAA National Runway Safety Plan for 2012-2014, and in the 2011 National Transportation Safety Board (NTSB) *Most Wanted List of Transportation Safety Improvements* (FAA, 2011a, 2011c;

Table 2. *FAA Runway Incursion Factors.*

Operational Errors (OE)	Pilot Deviations (PD)	Vehicle/Pedestrian Deviation (VPD)
Action of an Air Traffic Controller that results in: Less than required minimum separation between 2 or more aircraft, or between an aircraft and obstacles, (vehicles, equipment, personnel on runways) or Clearing an aircraft to take off or land on a closed runway	Action of a pilot that violates any Federal Aviation Regulation Example: a pilot crosses a runway without a clearance while enroute to an airport gate	Pedestrians or vehicles entering any portion of the airport movement areas (runways/taxiways) without authorization from air traffic control

Note. Adapted from http://www.faa.gov/airports/runway_safety/

NTSB, n.d.). Despite their explicit identification as a target for mitigation strategies, the rate at which RIs occur continues to escalate as shown in Figure 1 (note that while count decreased in FY 2009, rate increased as a function of reduced traffic volume). Data from the FAA Runway Safety website (http://www.faa.gov/airports/runway_safety/) and from FAA Annual Runway Safety Reports describe an increase in RI rate from 12.3 to 18.9 occurrences per million surface operations over the past six years of available data (FAA, 2010a).

Although the increasing rate of RI occurrence in the U.S. is disconcerting in and of itself, its continued escalation in combination with the substantial growth of air traffic indicates an urgent need to address the rise in RI events. In their Aerospace Forecast for FY2012-2032, the FAA projects annual growth of the domestic aviation sector at between two and three percent per year, with passenger numbers expected to rise from 731 million in 2011 to 1.2 billion in 2032. Extrapolation of present air operation totals over the next ten years results in a calculated estimate of 1,125 RI annual events in 2020, ten years from the last publicly available totals in 2010 (FAA, 2010a). This assumes that RI rate remains static, when in fact it has increased steadily in the past several years, and

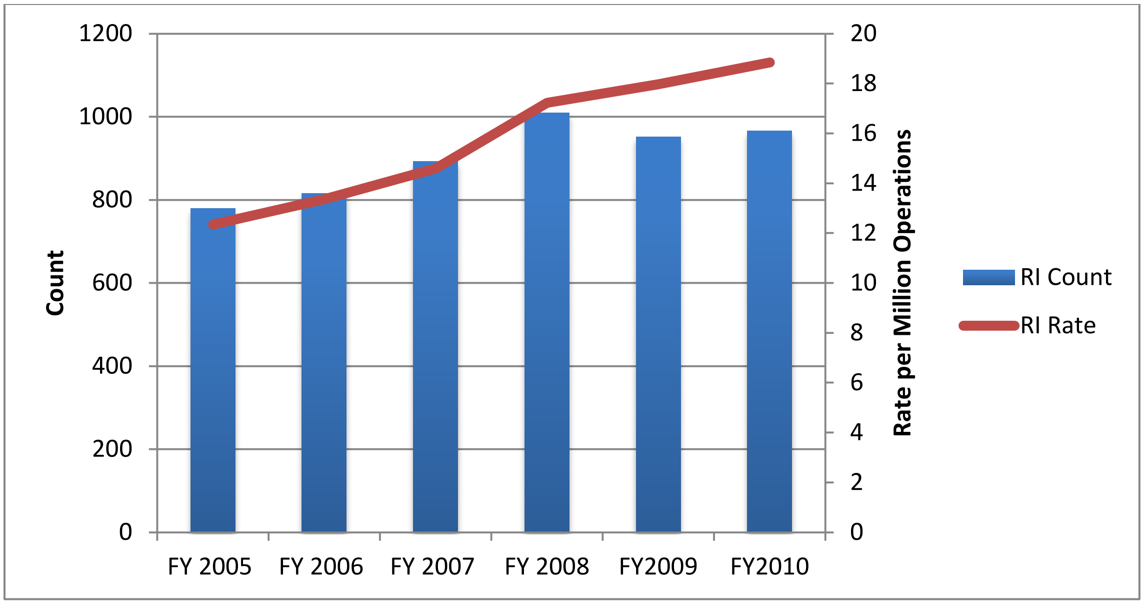


Figure 1. U.S. runway incursion count and rate FY2005 to FY2010 (http://www.faa.gov/airports/ runway_safety/).

that air traffic will grow only modestly as it recovers from the current economic downturn. Although these assumptions render the example given here difficult to generalize, the likely result of increasing or even static RI rate alongside increasing air traffic volume is more frequent RI occurrence. Cognizant of this, the FAA issued a *Call to Action* in late 2007, prompting not only an internal FAA challenge to address runway incursions, but a focused industry response as well (FAA, 2011b). Though the Call to Action prompted an initial decrease in pilot deviation type incursions, the trend reversed and the rate of RI events has continued its upward trend (FAA, 2011c).

In search of meaningful reduction of the RI rate, many strategies for RI mitigation have been presented, tested, and implemented. While some of these solutions have been met with success, the persistently increasing rate tempers premature declarations of successful widespread reduction of RI events by regulators and others. Rankin II (2008) reviewed a number of these initiatives and objectives with particular focus on those implemented by or at the direction of the FAA. On the whole, and as illustrated in the analysis by Rankin II (2008), RI reduction plans generally fall into three categories: those that address RI events from a human factors perspective, those that assume an organizational perspective, and those that identify technological or engineering solutions to the problem. Within these broad categories, individual research is often limited to a particular domain, examples of which are embodied by organizational, psychological, physiological, and technological or engineering-based theories and models (FAA, 2010a, 2011c; McLean and Monro, 2004). Schönefeld and Möller (2012) suggest that effectively addressing RI occurrences relies upon removing the human from the system to the greatest extent possible. Evidence shows that this approach can be effective (Dabipi, Burrows-McElwain, & Hartman, 2010; McLean and Monro, 2004; Torres, Metscher, & Smith, 2011); however, technological solutions must also be developed in conjunction with the fullest possible understanding of the nature of the problem (Rankin II, 2007). Given that the rate of RIs has not decreased in spite of these efforts is indicative that this level of understanding has yet to be achieved.

The existing research investigating RI data is important and meaningful, but it falls short in some respects beyond a failure to address RIs through a holistic, cross-domain approach. In addition, many studies of the modes by which RIs occur do not account for the substantial uncertainty involved in the investigation of rare events such as runway incursions, which is operationally evident in the lack of high-resolution RI data. In other areas of safety research, probabilistic risk assessment (PRA) has been successfully utilized in numerous applications characterized by uncertainty, such as that found in investigating RIs. PRA, pioneered in large part by the National Aeronautics and Space Administration (NASA) in the 1960s, seeks to provide answers to three basic questions: What can happen? How likely is it to occur? If it is to occur, what are the consequences? (Bedford & Cooke, 2001; Stolzer, Halford, & Goglia, 2008). PRA has seen widespread use in the nuclear, chemical, energy, aerospace, and financial industries, all of which share the common trait of high consequences of failure despite relative rarity of events (Stolzer, Halford, & Goglia, 2008). After the *Challenger* accident in 1986, NASA once again became a strong proponent of PRA, strengthening its position as a powerful tool for the prediction of risk where a system or systems are highly variable (NASA, 2002). As previously discussed, PRA often involves the study of rare events for which data are sparsely available, and while it provides a probabilistic alternative to deterministic point estimation of risk, PRA also has shortcomings in the context of complex, rare events such as aviation accidents. Zio (2009) argues that the complexity of systems such as those in which RIs occur renders event sequence-based techniques such as found in traditional PRA of limited utility. To this end, complex, multidisciplinary systems require safety risk assessment approaches that can dynamically model the complex interactions of actors and events (Stroeve, Blom, & Bakker, 2013). Even more advanced PRA methods such as Monte Carlo simulation in its naïve form – wherein simplified sampling methods do not support higher-order uncertainties – cannot capture the conditional or state-dependent nature of an event sequence leading to an accident or incident. The distinction between causal factors and what are often referred to in this study as causal paths (also causal sequences) is more than a semantic argument. Whereas the methods principally used to investigate RIs to date are focused on individual contributors to RIs, or static sequences, this research proposes a method of evaluating and understanding the interactions of causal factors that result in dynamic causal interactions that form paths or sequences.

As noted previously, PRA is used to answer three basic questions with respect to the nature of risk. Issues arise, however, in adequately doing so in the presence of both epistemic and aleatory uncertainty while relying solely on a frequentist view of probability estimation. Markov chain Monte Carlo (MCMC) is a means of describing the successive probabilities of events in relation to the immediately preceding occurrence. As an extension to the classical PRA methods described previously, it offers insight into the conditionality inherent to a complex sequence of events (Gamerman & Lopes, 2006). From a practical standpoint considering the variability that almost always accompanies complexity, Bayesian inference exhibits more widespread utility as a risk assessment tool given its treatment of probability as a measure of degrees of belief and inherent assignment of epistemic distributions to model parameters. This subjective view reflects on partial belief as a function of behavior choice and consequence. Only the subjective interpretation of probability allows for the integration of epistemic uncertainty in its analysis. Bayesian parameter estimation also accommodates a variety of data types, most notably expert elicitation in addition to classical statistical information (Siu & Kelly, 1998).

The proposed method in this study capitalizes on the ability of Bayesian methods, specifically probabilistic graphical models, and advanced sampling techniques to allow evaluation of causal factors to RI events across domains. Steadily increasing rates of RIs highlight the need for a more complete understanding of the complex interaction of factors and systems that contribute to RIs, and Bayesian belief network models allow for a novel means of examining the problem. This proposal outlines the use of Bayesian Belief Networks and stochastic sampling to effectively capture the interdependencies that characterize the chain of events that lead to an undesired state – "the incorrect presence of an aircraft, vehicle or person on the protected area of a surface designated for the landing and takeoff of aircraft" (EASA, 2011, p. v; ICAO, 2007, p. vii). In conjunction with a holistic theoretical basis for discovery, this method will allow for prediction of RI events, evaluation of mitigation strategies, and identification of key causal paths in the face of substantial uncertainty and across areas of knowledge. In the discussion to follow, these elements are synthesized into a method by which RI events may be dynamically modeled such that causal paths – the interaction of those components that lead to an undesired state – can be stochastically modeled and evaluated for more complete understanding of the problem and of how mitigation efforts are best applied to address it.

Statement of the Problem

The catastrophic collision of two Boeing 747 aircraft on the runway at Tenerife, Spain in 1977, which resulted in the deaths of 583 passengers and crew, elevated runway incursions in the public psyche (Tarrel, 1985). As a matter of public interest in aviation safety, this concern remains pervasive even today. In the U.S., runway incursions (RI) have been a topic of intense scrutiny by the Federal Aviation Administration (FAA) and National Transportation Safety Board (NTSB) for at least the past three decades. During this period, each agency has addressed RIs in its strategic planning as well as through a variety of initiatives designed to meaningfully reduce the frequency with which such incidents occur.

At present, strategies aimed at reducing RI threats have primarily been those that independently implement training and engineering protection by way of proposals to modify airport lighting, surface markings, signage, ground-based monitoring displays, and cockpit display devices (FAA, 2007, 2010a, 2011b; Moertl & McGarry, 2011). In contrast to the concrete engineering solutions identified by Rankin II (2008), comparatively little research has been conducted with a focus on understanding the covert errors that inform, or should at least be considered, in the design of effective mitigation strategies (Hendrickson, 2009). Even fewer studies appear in the literature addressing the dynamic interaction of causal factors as they combine resulting in undesired events, especially when those factors exist across domains of knowledge such as human factors, mechanical systems, or organizational dynamics (Luxhøj, 2003). This paucity of research is conceivably a function of the uncertainty that results from the small number of data points and relative infrequency of RI events acting in combination with a lack of cross-disciplinary research, presumably because of the complexity that often accompanies it. The research that does exist is not without merit; however, its narrow scope often fails to capture the dynamic conditionality of the sequence of events and states that lead to incidents such as RI events. In light of these apparent gaps in research and understanding, it is proposed here that the key to effective reduction of RI events is not in the application of independent solutions, but in a holistic understanding of the causal structure of RIs and the identification of pivotal interactions where mitigation strategies will be most effective.

Purpose Statement

The purpose of this research is to evaluate the feasibility and effectiveness of Bayesian belief network models, supported by structured expert elicitation, as a tool to examine causal factors and dynamic causal paths to RI events (regardless of Federal Aviation Regulations under which an aircraft is operated) in the U.S. This study employs techniques from many domains, and it is guided by the proposition that Bayesian inference and associated modeling techniques offer a robust and natural inferential platform for understanding RI events and the complex dynamics that influence them under the uncertainty of sparse data. The elicitation protocol and resulting BBN are tools intended for implementation at the regulatory level to assist in design and evaluation of RI mitigation and causation. Although probable end users are agencies such as the FAA or Department of Transportation, this study also provides tools for operationalizing BBNs at the airport level where sufficient sophistication in data availability and analysis is accessible.

Research Questions

The present research addresses two principal questions, the first of which informs the latter. First, what are the interacting causal factors that lead to RIs in the U.S.? Second, can runway incursions in the U.S. and their dynamic causal factors and interactions be modeled through the use of a Bayesian belief network supported by expert-elicited data?

Significance of the Study

This study is significant in its use of a novel means to investigate dynamic causal interactions across many domains. Modeling RI events in this way offers the potential for insight into why RIs occur, and what events or event interactions present the most promising prospect for substantial reduction of incursions. To date, the increasing rate and growing numbers of RIs in the US over the past several years reveals that this has not yet been accomplished. The combination of methods investigated in this research provides a unique opportunity to eliminate this considerable knowledge gap.

Delimitations

This study does not attempt to address RI incidents around the globe. Rather, it focuses only on RIs within the U.S. because of the infeasibility of obtaining homogeneous data across many countries. Although other countries are excluded, U.S. data are likely representative of RIs experienced worldwide, given that ICAO standards are nearly universally applied.

RI data for this study are limited to those years where the definition of RI is consistent with ICAO and with the definition used today. This purposively limits data collection to a five-year period (2008-2012) inclusive only of the data collected under the presently-used RI definition and severity categorization scheme, which substantively changed in 2007 to align with the ICAO definition of runway incursion. This change is addressed in more detail in the review of literature to follow. Although data from years prior to the definition alignment may be revised to the new standard by estimation, as has been done in some FAA reports (FAA, 2010a), the benefit to the additional data points is unlikely to be so great as to outweigh the liability of approximated figures given that variation can be observed even in the FAA figures.

While the model discussed and developed in the present study will have the *capacity* to support sensitivity analysis and evaluation of mitigation strategies via inference algorithms, this is beyond the scope of the research. These methods, which include influence diagrams, are extensions of Bayesian networks that can be used to evaluate strategies to achieve optimal utility and support decision making (Darwiche, 2010). The results of the current study will allow future researchers to extend the network to achieve just such results.

Limitations and Assumptions

As mentioned briefly in the preceding section, it is assumed, based on commonality of air and ground navigation procedures as well as aerodrome design in part because of general adherence to ICAO standards, that the findings of this study can be reasonably generalized to other populations. Nevertheless, the purposely constrained scope of this research is a potential limitation to the application of results found therein.

Although RIs have the capacity for catastrophic results, they do not occur in that mode with frequency such that large numbers of data points are available. Small numbers of some data may affect the power of the study, which may also affect broad application of results. This scarcity of data is frequently a limitation to studies of RI phenomena, but in the present research, sparse research is supplemented with expert elicitation. Given the state of knowledge of RIs, and the spread of that knowledge across domains, expert elicitation is well-supported in its role in this study (Mosleh, Bier, & Apostolakis, 1988).

Definitions of Terms

Aleatory Uncertainty: Aleatory uncertainty is due to the natural, unpredictable variability of a system or a process. With respect to Bayes' theorem, aleatory uncertainty is that about which one cannot or chooses not to learn. Though it cannot be resolved through expert knowledge or judgment, it may be quantified (Bedford & Cooke, 2001).

Bayesian: Referring to the methods of inference based on Bayes' Theorem and made in terms of probability statements that are updates as additional evidence is made available.

Causality: The relationship between states or events such that one is understood as the consequence of the other. One implication of causality is that a model must demonstrate more than correlation in the classical statistics sense.

Conditional Probability Table: A tabular representation of the conditional probability distributions for variable relationships.

Decision Maker (DM): The weighted combination of expert judgments under Cooke's Classical Model. In the Classical Model, the DM is used in place of individual assessments, and is based on the aggregation scheme outlined by Cooke (1991).

Deterministic: Pertaining to exactly predictable (or precise) processes, the outcome of which is known with certainty if the inputs are known with certainty. This type of model is the antithesis of aleatory (Kelly & Smith, 2011).

Directed Acyclic Graph: A structured flowchart with parametric relationships connected by line segments (edges) in order to map out the paths of priors used to illustrate causal structure in a Bayesian network (Gill, 2008).

Edge: The connecting lines between nodes in a graph, which may be either directed or undirected.

Emic: Assuming the viewpoint or perspective of a cultural insider.

Epistemic: Pertaining to the degree of knowledge about models and their parameters.

Epistemic Uncertainty: The uncertainty in the model of a system or a process that arises through a lack of knowledge of the system. Epistemic uncertainty relates to those things about a system that can be learned. In this way, epistemic uncertainty may be resolved, at least conceptually, via sufficient study such as through the use of expert elicitation (Bedford & Cooke, 2001).

Etic: Assuming the viewpoint or perspective of a cultural outsider.

Frequentist: With respect to probabilistic reasoning, the long-run expected frequency with which a phenomenon will occur. *Frequentist* refers to the inferential framework within which common statistical methodologies such as hypothesis testing and confidence intervals function.

Informative Prior: A prior distribution function that expresses some positive information or knowledge, by way of the selected distribution, about an unknown parameter.

Joint Distribution: A probability density function that involves more than one random variable (Lynch, 2012, p. 19).

Markov Chain: A stochastic process that deals with the characterization of random variable sequences where given the present state, past and future states are independent (Gamerman & Lopes, 2006).

Markovian: A Markov process is a stochastic process that is considered to be memory-less in that the future states of such a process depend only on the present state.

Monte Carlo: A statistical approximation technique that uses computer algorithms and random number generation to produce probabilistic inputs in a pre-specified manner to solve problems and gain insight to phenomena that have some random component (Shonkwiler & Mendivil, 2009).

Moral Graph: The undirected graph equivalent of a directed acyclic graph (Koller & Friedman, 2009).

Naïve: As used to describe probabilistic reasoning, naïve refers to methods that do not account for interdependence or uncertainty. Naïve (or standard, or crude) Monte Carlo methods consider a sample of *n* independent copies of a random variable, and estimate event probability based on the proportion of a rare event occurrence over the sample (Rubino & Tuffin, 2009).

NP Hard: Referring to the complexity of an algorithm, NP-Hard indicates a problem at least as hard as the hardest problems in NP (non-deterministic, polynomial time).

Node: Values and variables in the model as specified by the model builder (Gill, 2008).

Posterior: When data are combined with the prior, an updated probability distribution is mathematically computed and is called the *posterior distribution* or posterior.

Prior: What is currently known about parameters within the model is expressed as a probability distribution on those parameters, called the *prior distribution* or simply the prior.

Probabilistic Graphical Model: A graphical declarative representation of the conditional dependence or independence between model variables (Koller & Friedman, 2009).

Probabilistic Risk Assessment: a comprehensive, structured, and logical analysis method aimed at identifying and assessing risks in complex technological systems for the purpose of cost-effectively improving their safety and performance (NASA, 2002).

Runway Incursion: Any occurrence at an aerodrome involving the incorrect presence of an aircraft, vehicle or person on the protected area of a surface designated for the landing and take-off of aircraft (FAA, 2007, para. 2).

Stochastic: A reference to the randomness of a system. The opposite of deterministic.

Surface Incident (Deviation): Any event where unauthorized or unapproved movement occurs within the movement area, or an occurrence in the movement area associated with the operation of an aircraft that affects or could affect the safety of flight (FAA, 2009a).

List of Acronyms

AMASS	Airport Movement Area Safety System
ASDE-3	Airport Surface Detection Equipment Model Three
ASDE-X	Airport Surface Detection Equipment Model X
ASIAS	Aviation Safety Information Analysis and Sharing
ASRS	Aviation Safety Reporting System
BBN	Bayesian Belief Network
DAG	Directed acyclic graph

EASA	European Aviation Safety Agency
ET	Event tree
FAA	Federal Aviation Administration
FTA	Fault tree analysis
ICAO	International Civil Aviation Organization
MITRE	The MITRE Corporation
MTBF	Mean time between failures
NTSB	National Transportation Safety Board
OE	Operational error
PD	Pilot deviation
RI	Runway incursion
RITA	Research and Innovative Technology Administration
RSO	Federal Aviation Administration Runway Safety Office
RWSL	Runway Status Lights
VPD	Vehicle/pedestrian deviation

CHAPTER II

REVIEW OF THE RELEVANT LITERATURE

This chapter examines the literature that informs this research, discussing not only the current state of RI research and mitigations in the U.S., but also the relevant practical and theoretical considerations of investigating RI causation through BBNs and structured expert judgment. The complex systems found in aviation organizations are dynamic in their behavior, and the response of these systems to perturbation from a desired or normative state is an intricate and varied series of interactions with the environment as well as between the components, including humans, of the system itself. Generally accepted guidance within the sphere of safety management points to a requirement for probabilistic risk assessment as a set of methods for predictive analysis in systems characterized by uncertainty and high reliability (Stolzer et al., 2008). However, such methods are frequently limited to discrete event simulation and do not often behave dynamically. This limitation is apparent when considering the increasing rate of runway incursion incidents despite concerted efforts to reverse this trend. Bayesian Belief Networks (BBNs) enable probabilistic estimation that more accurately reflects the interaction of each state of an event sequence and the end state frequencies associated with the simulation of events while retaining the ability to classify risk importance as in classical probabilistic risk assessment (PRA) (Darwiche, 2009; Gamerman & Lopes, 2006; Kelly & Smith, 2009; Koller & Friedman, 2009). Bayesian Belief Networks (BBNs) can seamlessly incorporate expert opinion in the face of uncertainty, and they are flexible to systems that change over time. In the context of runway incursions, the incorporation of BBNs allows a more complete inferential and predictive analysis of risk and better informs choices between alternative safety interventions in a causal framework (Darwiche, 2009; Darwiche, 2010; Napoles, 2010; Pearl & Russell, 2003).

Runway Incursions

As was previously discussed, RI events remain one of aviation's most critical challenges, and remain a prominent fixture in the annual plans and reports issued by organizations including the NTSB and FAA (FAA, 2010a, 2011c; NTSB, n.d.). Despite the efforts of these groups and others, the rate at which RIs occur in the U.S. continues to escalate. Records from the FAA Runway Safety website (http://www/faa.gov/airports/runway_safety/), illustrated previously in Figure 1, point to an increase in RI rate from 12.3 to 18.9 occurrences per million aircraft operations over the past six years of publicly available data (FAA, 2010a). In its most recent Annual Runway Safety Report, the FAA cites a "drop by 50 percent over the previous year" (2010a, p. 1), a position repeated in a 2010 press release announcing "terrific progress in the area of runway safety" (FAA, 2010b, para. 2). While these proclamations are encouraging, a closer look at the data indicates that RIs are not declining on the whole in rate or in number. To some extent, this discrepancy may be attributed to constraints on the data analyzed in reports as well as to the evolving definition of RIs.

Defining runway incursions. Through its publicly available reports (FAA, 2010a; 2010b), the FAA has in some ways defined the RI problem into success by constraining data to include only those incidents that fall into the most serious RI categories, but the risk posed by RIs has continuously trended upward over several years. Changing definitions of RIs affect more than the way they are represented in news releases. Prior to October 1, 2007, RIs were defined differently in the U.S. than by ICAO:

> Any occurrence in the airport runway environment involving an aircraft, vehicle, person, or object on the ground that creates a collision hazard or results in a loss of required separation with an aircraft taking off, intending to take off, landing, or intending to land (FAA, 2004, p. 9).

With the 2007 change, the FAA aligned its definition to that adopted by ICAO in an effort to standardize RIs, which were previously described by "at least 20 definitions" (FAA, 2007, para. 7) in countries around the world. The definition currently in use by the FAA and ICAO is:

> Any occurrence at an aerodrome involving the incorrect presence of an aircraft, vehicle or person on the protected area of a surface designated for the landing and take-off of aircraft (FAA, 2007, para. 2).

This change to the definition of an RI not only aligned the U.S. to the international standard, it meant that events previously classified separately as surface deviations would be categorized as a Category C or D (refer to Table 1) runway incursion. This change is evident in the FAA's reported numbers of RIs, which increase dramatically after the definition alignment as illustrated in Figure 2. Also illustrated in Figure 2 is that runway incursion figures were retroactively estimated in concordance with the updated definition for FY 2006 through 2007.

Runway incursion data. In spite of FAA reports (FAA, 2010a; 2010b) that indicate runway incursions are decreasing, Figure 3 illustrates that the opposite is true for all but the most serious of RIs. Returning to Figure 2, it can be observed that the reported

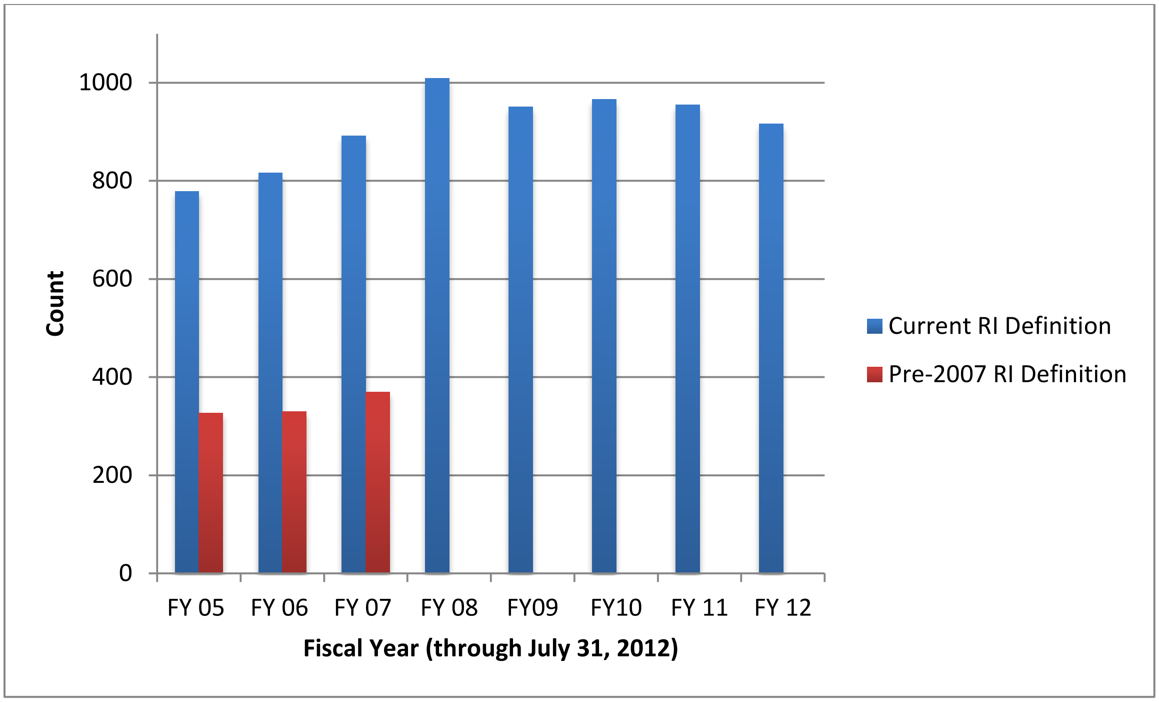


Figure 2. Comparison of U.S. runway incursion count FY2005 to FY2012 (Oct. 1 – Sept. 30) by definition (data from http://www/faa.gov/airports/runway_safety/).

count of RIs in FY 2012 as of July 31 are nearly equal the total of RIs in the entirety of FY 2011 despite there being three months of unaccounted data remaining in FY 2012. Also shown by Figure 3 is that even those severe RI events categorized as A or B have begun to show a rise in rate once more. Examined in greater detail, as in Figure 4, it is apparent that the trend for each category of severe runway incursions is increasing, as count for severe RIs has nearly doubled over the two previous years through only the third quarter of 2012.

The focus on severity rating in FAA Annual Runway Safety Reports (FAA, 2009b, 2010b) as a measure of success in mitigation of RIs is somewhat misleading as an indicator of the magnitude of the problem associated with RIs. Recalling the categorical definitions outlined in Table 1, it becomes evident that the severity ranking schema in use by the FAA and by ICAO is less a function of the manner in which an RI occurs as it is a

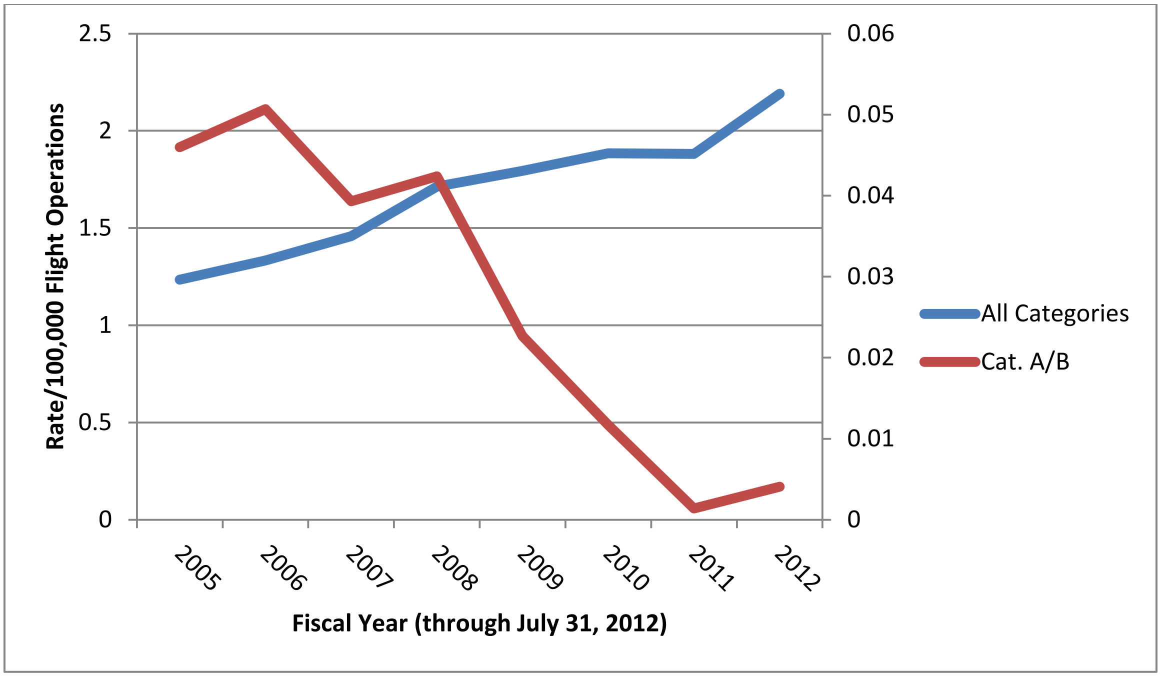


Figure 3. Trend of RI rate, 2005-2007 data estimated to reflect definition change definition (data from http://www/faa.gov/airports/runway_safety/).

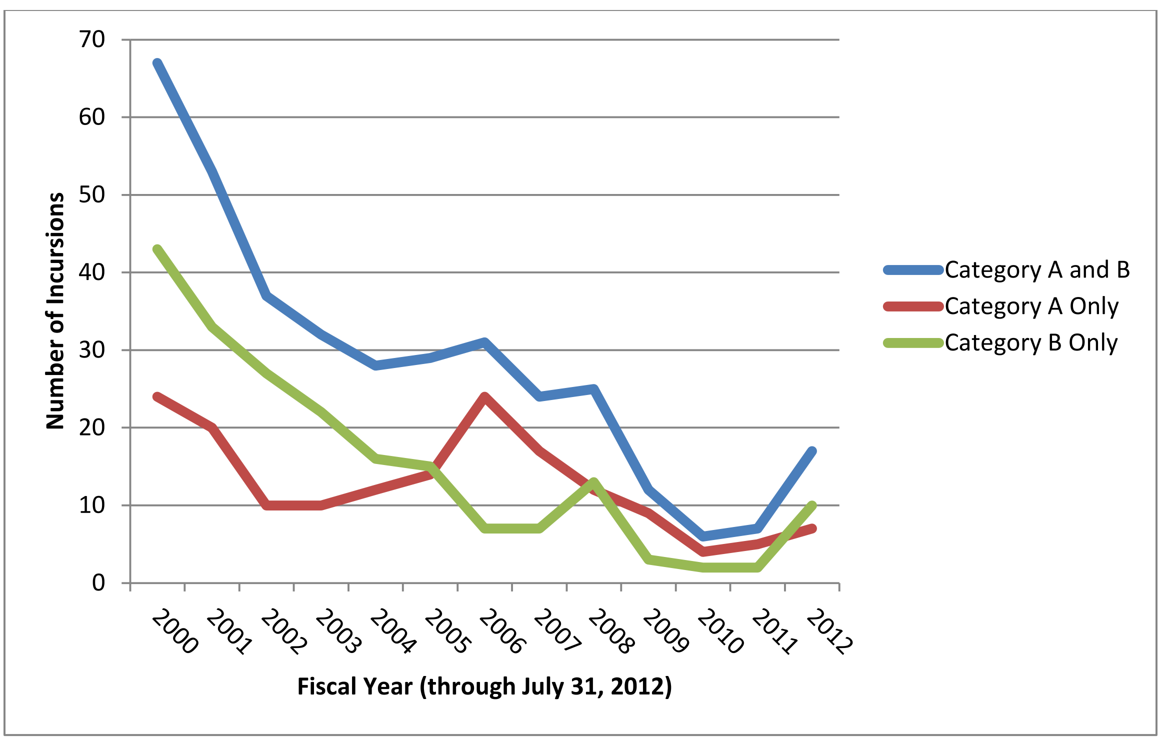


Figure 4. Category A and B runway incursion trend by count, 2000-2007 data estimated to reflect definition change (data from http://www/faa.gov/airports/runway_safety/).

product of the time separation between aircraft, vehicles, or pedestrians occupying the same protected surface area.

Considering this, gleaning causal information based solely on RI severity strata may be difficult if not altogether misrepresentative of the underlying structure that contributes to RI causation because of the artificiality of the separation-based severity ranking (Dr. K. Cardosi, personal communication, November 9, 2012). Nevertheless, increasing rates of RIs are clear, and point to a growing need to address the problem utilizing dynamic analytical methods in order to account for the full depth and breadth of existing knowledge with respect to the safety of aircraft surface operations. A number of studies have been conducted in attempts to shed light on the growing issue of RIs.

Review of runway incursion research and study of causal factors. The figures in the preceding discussion use data from the FAA Runway Safety Office (RSO), which is available via the Aviation Safety Information Analysis and Sharing database, the FAA Runway Safety Program website (http://www.faa.gov/airports/runway_safety/), or via the RSO directly. As the regulatory body charged with oversight of air traffic operations, including those on the surface, the FAA has a keen interest in researching the causes of and potential solutions to RIs. FAA research on RIs began in earnest in the late 1980s, when the FAA Assistant Administrator was "directed to identify the causes of runway incursions and formulate measures for alleviating this problem" (NTSB, 1990, p. 44). The NTSB issued a report in 1986 that encouraged this action by the FAA and included 33 recommendations concerning RI prevention, many of which had previously been issued by the Safety Board as early as 1973 (NTSB, 1986). Table 3 summarizes the relevant literature addressing RI causation.

Table 3. *Summary of Reviewed Runway Incursion Causal Factor Literature.*

Year	Author(s)	Summary of Factors	Domain(s)
1981	Bellantoni & Kodis	No causal patterns aside from human errors	Human Factors
1985	Tarrel	Communication breakdown and the influence of taxiing aircraft highlighted	Human Factors
1986	NTSB	Disorientation and communication breakdown noted as well as ATC training	Organizational, Human Factors
1989	Bales, Gillian, & King	An analysis of ATC-related runway incursions, with some potential technological solutions	Technology, Human Factors
1991	Steinbacher	An analysis of ATC-related runway incursions in the national airspace	Human Factors, Organizational
1995	Koenig	Cockpit procedures and unfamiliarity with airport layout compound pilot error	Human Factors
1994	Adam & Kelly	Catalogued RI contributing factors based on pilot surveys with focus on communication and navigation errors	Technological
1996	Adam & Kelly	Catalogued RI contributing factors based on pilot surveys with focus on procedures and factors affecting pilot performance	Human Factors, Organizational
2000	Knott, Gannon, & Rench	Literature review highlighted numerous human factors-related contributors to RI causation	Human Factors
2001	Cardosi & Yost; Cardosi	Reviewed previous data with analysis focused on aids to memory and situational awareness for pilots and air traffic controllers	Human Factors
2005	Cardosi	Human factors effects on ATC	Human Factors
2008	Rankin II	Review of causation and FAA mitigation efforts noting that ground vehicle training was a common causal factor to RIs	Human Factors, Organizational
2011	Torres, Metscher, & Smith	A study of the relationship between human factor errors and RI occurrence identified loss of flight crew situational awareness and miscommunication as common causes	Human Factors
2012	Chang & Wong	Summary of human factors research and factors associated with RI events	Human Factors

In its role as the agency responsible for runway incursion study and mitigation in the U.S., the FAA is a primary contributor to research that seeks to understand RI phenomena. On the basis of data analysis conducted by the FAA and through its partner agencies (such as MITRE, RITA, NTSB, and others), the 2002-2004 FAA Runway Safety Blueprint identified five points that defined the agency's view of RI events:

- Operational performance in the airport movement area must be further improved to reduce runway incursions.
- Runway incursions are systemic, recurring events that are unintentional by-products of NAS operations.
- Operations must be standardized to reduce risk at a time when growth is challenging runway and infrastructure expansion.
- Collision-avoidance safeguards need to be developed for the high-energy segment of runways, where aircraft are accelerating for take-off or decelerating after landing.
- Human factors are the common denominator in every runway incursion (FAA, 2001).

From a research perspective, most notable of these points is that human factors are identified as pervasive to RIs. As early as 1981, Bellantoni and Kodis (1981) began investigating RIs, noting, "there does not appear to be any pattern to the causes of runway/taxiway transgressions other than human errors on the part of both air traffic controllers and pilots" (p. v). Some of the underlying factors identified by Bellantoni and Kodis (1981) as contributing to system error were: deficiencies in attention, judgment, and phraseology (p. 11). Within these categories, a number of specific human errors were identified for pilots and controllers based on analysis of ASRS, NTSB, and ATC report data (Table 4). The researchers found that within the runway and surface incidents

Table 4. *Errors in Factors in RI Causation.*

Pilot	Controller
Proceeded without clearance	Directly conflicting clearances
Failed to see and avoid	Insufficient separation
Failed to display proper lights	Cleared to obstructed runway
Lost/disoriented	Provided inadequate information
Failed to understand message	Erroneous instruction
Failed to follow instructions	Faulty GC/LC coordination
	Failed to track aircraft
	Poor supervision

Note. Adapted from Bellantoni and Kodis (1981).

studied roughly 95 percent were due to an element of human error, with the error lying approximately evenly-distributed between pilots and controllers. Though it was a first step toward understanding RI data, Bellantoni and Kodis' (1981) research could not explore the underlying nature of the errors it identified because of a lack of detailed data.

Tarrel (1985), in one of the earliest studies of the causes of what are referred to in the subject study as runway transgressions, sought to "uncover patterns of behavior that lead to these incidents" (p. 2). Citing work by Billings and O'Hara (1978), and utilizing NASA ASRS reports to obtain information through a bi-directional analysis procedure, Tarrel noted two critical characteristics of RI events: that the breakdown in information transfer between parties was an important factor, and that taxiing aircraft were a major contributor to RIs. On the flight deck, this analysis showed that forgetfulness, distraction, disorientation, and misunderstanding of an ATC clearance were most commonly observed. Tarrel noted factors quite different from those in pilot-enabled RIs when assessing ATC-enabled incidents. Misjudgment of aircraft spacing, coordination between ground and local controllers, non-standard phraseology, and high-workload conditions were often indicated in instances of ATC error. Tarrel concluded that both pilot and ATC-enabled RIs carried unique risk profiles, and that the task of reducing them would fall to the proper identification of those operational areas where the greatest effects might be achieved. This proper identification is the impetus for the use of more dynamic means of investigating RI events, such as through BBNs.

The NTSB, in a 1986 Special Investigation Report, investigated 26 runway incursions at control tower-equipped U.S. airports in an attempt to discern their "underlying causes and to recommend appropriate remedial actions" (p. 1). Failure to identify traffic, inconsistent supervision, memory failure, boredom, and coordination between controller positions were all cited as factors from an ATC standpoint. While each of these factors is well-supported, the Special Investigation does not address the means by which these causal elements interact and manifest as an RI event.

Koenig (1995) reviewed the results of MITRE Corporation survey studies sponsored by the FAA in 1993 and 1994 that gathered survey data from U.S.-based airline pilots. Though the questionnaire was not structured such that it could be used for formal statistical analysis, a number of factors were identified as potential precursors to RIs, most of which fall within the definition of human error. MITRE's reports (Kelley & Steinbacher, 1993) emphasized the insight that could be gained through plain-text operational reports of RI events.

An examination of human error was central to the work completed by Cardosi and Yost (2001) and Cardosi (2001), which noted a number of areas for improvement in communication and memory aids for ATC controllers. Their research also reviewed previous work by MITRE analyzing operational errors in the ATC system (Bales, Gilligan, & King, 1989; Steinbacher, 1991) as well as the surface incident study that was the subject of Kelley and Steinbacher (1993), Adam and Kelley (1994), and Adam and Kelley (1996). As in a number of other studies, including those conducted by MITRE, NASA ASRS reports were evaluated. In particular, Cardosi and Yost's research evaluated nearly 250 reports submitted to the ASRS system by ATC controllers to analyze causal factors affecting controller error. In the same study, over 75 reports submitted by pilots were also evaluated. While the analysis of ASRS reports was undoubtedly enlightening, the reports selected for study were done so on the basis of recency, which may have affected the representativeness of the sample.

Rankin II (2008) approached RI causation from the perspective of personnel and airport vehicle transgressions into the runway environment. Rankin II focused on ranking effectiveness of FAA initiatives aimed at RI reduction using a survey instrument to collect operational data. Rankin II's survey achieved only modest responses, with a reported response rate of 35 percent. As such, the responses that indicated driver training was a common cause and that FAA efforts should refocus on such programs may have been limited in their generalizability.

Torres, Metscher, and Smith (2011) also attempted to identify common human factors causes of RIs with a focus similar to the present research: that a better understanding of causation would allow the FAA and others to gain greater success in mitigating RI events. The researchers noted that in reviewing nearly 300 ASRS reports, the most common causal attribution in RI events was to a loss of situational awareness on the part of the flight crew, followed closely by miscommunication. However, no further analysis of the underlying causes of these two contributing factors was conducted, with recommendations limited to a broad suggestion for increased focus on the human in the loop (Torres, Metscher, & Smith, 2011). The authors acknowledge this limitation, noting that further research is necessary so that a "positive impact on the reduction of the loss of situational awareness that leads to runway incursions" may be realized (Torres, Metscher, & Smith, 2011, p. 24).

This closing statement points to an apparent shortcoming in the understanding of what were referred to previously as covert errors in the context of RI causation. It is an increased focus on these errors and their complex, diverse interactions with "interrelated system components" that the present research suggests is critical to a fuller understanding of RI causation and mitigation (Luxhøj, 2003, p. 17). Despite this gap in understanding on which interventive measures are based, a number of strategies directed at reduction of RI incidents have been developed and implemented over the past several years.

Review of runway incursion mitigation strategies. Runway incursion mitigation efforts generally fall into categories that can be classified by the domain on which they focus. For the purposes of discussion here, those domains are: infrastructure/organizational factors, human factors, and technological/engineering factors. Certainly, some research incorporates elements of more than one of these domains, and those that do are discussed herein. Because of the regulated nature of air transportation, a number of mitigation strategies aimed at the reduction of RIs have been implemented or directed by the FAA. Table 5 provides a brief overview of the relevant literature addressing RI mitigations.

As Table 5 indicates, a substantial amount of research to date has focused largely, and in many cases wholly, on technological improvements to reduce the contribution of human-in-the-loop systems to RI causation and avoidance. Systems such as ASDE-3,

Table 5. *Summary of Reviewed Runway Incursion Mitigation Literature.*

Year	Author(s)	Summary of Mitigations	Domain(s)
1998	Kelly & Jacobs	Recommended changes to memory tools for ATC as well as improvements to ASDE-3 and stop bar lights	Human Factors, Technological
2000	Knott, Gannon, & Rench	Review of technological barriers to human error, including IMAGE and TARMAC systems	Organizational, Technological, Human Factors
2004	McLean & Monro	Runway ASDE-3, ASDE-X, and AMASS	Technological
2002,2005, and 2006	Jones, Jones, and Jones & Prinzel	RIPS system and HUD symbology for situational awareness	Technological
2007	Vernaleken, Urvoy, & Klingauf	SMASS for wrong-runway departure avoidance	Technological
2007	CAST	Numerous recommended safety enhancements across domains	Human Factors, Technological
2010	Dabipi et al.	Low cost alternative to FAROS and RWSL systems	Technological
2011	Moertl and McGarry	Cockpit display of traffic information with indicators and alerts	Technological
2006,2011b	FAA	Runway Status Light technical requirements and use	Technological
2012b, 2012c	FAA	Parts 91, 121, 125, and 135 flight crew procedures during taxi operations / Parts 91 and 135 single pilot, flight school procedures during taxi operations	Human Factors, Organizational

ASDE-X, AMASS, RWSL, and others have undoubtedly been met with success (Dabipi et al., 2010; FAA, 2006, 2011b, 2012b, 2012c; Kelley & Jacobs, 1998; Moertl & McGarry,2011); however, that success is tempered by the knowledge that little improvement in RI rate has been realized as demonstrated in the foregoing discussion. In much of the reviewed literature, mitigation strategies are approached as domain-specific, rather than as a cross-domain effort to combat the dynamic causation of RIs. The research proposed here is intended to address this shortcoming by providing a modeling environment that supports cross-domain mitigation investigation and implementation, which has to date been generally lacking. Perhaps the most structured example in aviation using probabilistic, interdisciplinary mitigation of RI events has come from the Commercial Aviation Safety Team (CAST).

CAST, which was formed in 1998 to bring together industry and FAA stakeholders in an attempt to reduce the air travel accident rate, has also addressed the issue of runway safety. CAST is primarily responsible for identification and prioritization of mitigation strategies based on its consensus review of event sequences that lead to accidents. Although the CAST methodology is based principally in a panel review of accidents, and is therefore limited by the relatively small number of RI occurrences, the process does include some elements of probabilistic reasoning. Of note in discussing the methods used in the CAST process is that they are inherently deterministic, utilizing single-point estimates in all, or nearly all cases. However, the methods employed in the CAST process produce several inputs critical to effective probabilistic risk modeling: some estimation of the severity or probability of an occurrence, an estimated probability of severity or probability, estimated effectiveness of proposed controls, and some knowledge of the cost of such controls (Stolzer, Halford, & Goglia, 2008). These products of the CAST methodology make it ideally suited for application of more robust probabilistic and decision-making modeling than is presently employed, specifically through the applied use of more dynamic and flexible probabilistic methods as are presented here.

As has been previously discussed, methods such as those that utilize probabilistic reasoning are called for because the causation and associated mitigation strategies discussed here have generally failed to achieve the intended result of substantially reducing the incidence of RIs in the U.S (FAA, 2010a). As the volume of U.S. air traffic is forecast to rise, this gap in mitigation efficacy carries with it a growing potential for disastrous results and the associated need for more robust analysis under uncertainty. Probabilistic reasoning offers the capacity for analysis and decision making under uncertainty, and methods such as Bayesian network modeling do so while remaining flexible to the dynamic interaction of factors across domains.

Probabilistic Risk Assessment

As the aviation industry experiences accident rates that are among the lowest in history, identification of threats or new approaches to mitigate consequences has become an increasingly difficult task (International Air Transport Association, 2011). Risk assessment methods in an aviation safety context are often focused only on reactive and proactive efforts, and when predictive modeling is employed, it is frequently limited in scope by addressing only discrete event probabilities rather than viewing outcomes as a sequence of conditionally-dependent events. Regulators and operators are encouraged by current research to implement infrastructure that fosters not only proactive, but predictive identification of hazards, and guidance provided by the Federal Aviation Administration (FAA) and International Civil Aviation Organization (ICAO) points to the need for robust risk assessment programs (Stolzer, Halford, & Goglia, 2008).

Programs such as the Commercial Aviation Safety Team (CAST) have demonstrated the effectiveness of probabilistic risk assessment (PRA) as a tool for identifying causal structures in aviation accidents and for developing and prioritizing mitigation strategies; however, these methods rely heavily on deterministic point probability estimates gained primarily through subject-matter expert (SME) consensus. Institutional knowledge is often the primary tool for identification of hazards and assignment of probability estimates (as in the CAST model discussed previously). Brooker (2011) points to a potential shortcoming in this regard, asking “what are the mechanisms by which experts ‘know’ such probabilities?” (p. 1154). In a frequentist paradigm, this argument is more concerning; however, in the context of structured Bayesian inference for risk assessment, the model is capable of learning as additional information and expertise is gained, and in any event, probability remains bound by the fundamentals of probability theory.

Predictive safety modeling. In its manual on the subject, the International Civil Aviation Organization (ICAO) offers the following brief explanation of the purpose of predictive analysis, the scope of which includes probabilistic modeling such as discussed herein:

> Predictive safety data collection systems are essentially statistical systems, whereby a considerable volume of operational data, which alone are largely meaningless, are collected and analyzed, and combined with data from reactive and proactive safety data collection systems. The aggregation of data thus leads to the development of a most complete intelligence that allows organizations to navigate around obstacles and currents and position themselves optimally within the drift (2009, p. 3-11).

Modeling is one means of achieving the level of predictive analysis detailed by ICAO's description. Modeling by BBNs works with and within the concept of probabilistic risk assessment (PRA) to provide answers to three basic questions: what can happen? How likely is it to occur? If it is to occur, what are the consequences?" (Bedford & Cooke, 2001; Stolzer et al., 2008). PRA and its associated modeling techniques are not the sole means by which the pivotal task of safety management may be achieved, but as Figure 5 illustrates, they are necessary tools as organizations seek to move beyond proactive management of safety risks. PRA often involves the study of rare events for which data are sparsely available, and while it provides a probabilistic estimation of risk, PRA has shortcomings in the context of complex, comparatively rare events such as aviation accidents. Rubino and Tuffin (2009) cite mean time between failure (MTBF) rates of 10^{-9} as a representative example of aviation system reliability requirements. Given this low probability of occurrence, traditional PRA tools such as Monte Carlo simulation become infeasible due to sample size requirements unless other techniques are introduced. Additionally, PRA in its simpler forms is unable to accommodate evolution of risk estimates as a process changes or as updated information enters the system by investigating events as they occur. In the discussion to follow, an approach is proposed that allows the effective use of probabilistic risk modeling while providing robustness to the dynamic uncertainty inherent to complex systems while also accommodating the totality of knowledge of a domain. More importantly, this approach captures these elements in a truly predictive approach, providing an elegant means of addressing the truism cited by Stolzer et al., that "you can't expect to meet the challenges of today with yesterday's tools and expect to be in business tomorrow" (2008, p. 219).

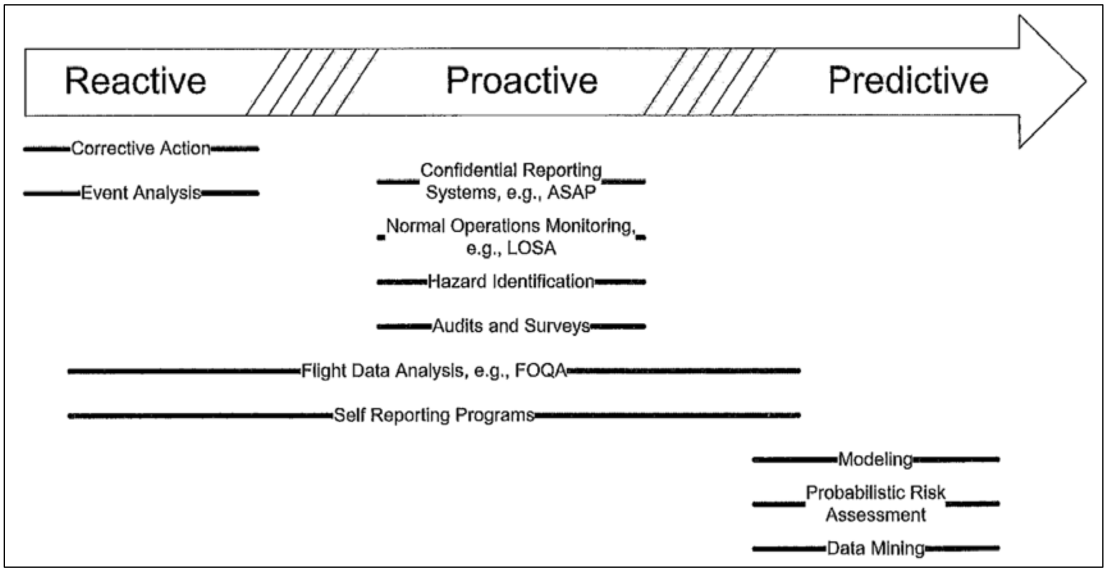


Figure 5. Safety management continuum (Stolzer et al., 2008, used with permission).

Bayesian Reasoning

Bayesian estimation of conditional probabilities is important in a subjective estimation of risk as it allows for the epistemic distributions of the aleatory model parameters in a model to be updated as new knowledge becomes available. In PRA, where uncertainty is often supplanted by expert elicitation, Bayesian inference is especially appropriate given its treatment of expert judgment simply as another source of evidence. Using Bayes' theorem (given in Equation 1), discrete as well as continuous probability distributions are addressed.

$$P(\mathrm{A}|\mathrm{B}) = \frac{P(B|A)P(A)}{P(B)} \tag{1}$$

where:

P(A) is the prior, or *a priori*, probability of *A* in that it does not account for any knowledge about *B*;

P(A|B) is the conditional, or posterior, probability of *A* given that *B* is true;

P(B|A) is the conditional probability of *B* given *A*; and

P(B) is the prior, or marginal, probability of *B*.

Bayesian belief networks. Bayesian Belief Networks (BBNs) illustrate composite, conditional probabilities in the form of directed acyclic graphs very similar to the illustration of a Markov chain. In these graphs, univariate random nodes, representative of variables of interest, are linked by arcs representative of influences between nodes. The acyclic requirement of BBNs simply means that there is no directed path that returns to its own starting point, a logical premise given the present application of the method Bedford and Cooke (2001) describe. Equation 2 specifies the joint distribution, and Figure 6 illustrates an example of the simplest form of a belief net.

$$p(x_1, x_2) = p(x_1)p(x_2|x_1) \tag{2}$$

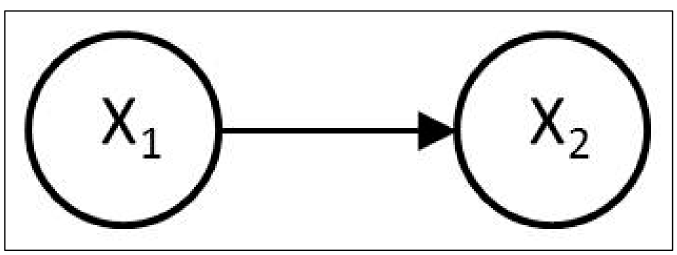


Figure 6. Simple BBN.

In the foregoing figure and equation, the probability specification is the marginal distribution of x_1 and the conditional distribution of x_2 given x_1 for every value of x_1. This expression is in the simplest possible form, and does not account for Markovian properties. However, rather than discuss notation of Markov chains here, their representation by BBNs are addressed in the discussions that follow. Figure 7 illustrates

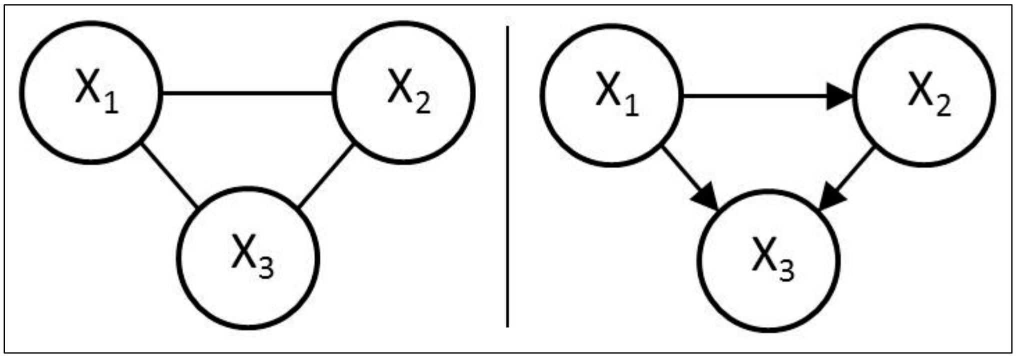


Figure 7. Undirected (left) and directed graphs.

basic undirected and directed graphs, with the latter as an example of the type utilized by BBNs.

The differentiation in BBNs comes from the inclusion of belief as a means of establishing subjective probability distributions in the dependency model. As in a typical Markov Chain, each node in the BBN can assume any one of its possible states, the belief in which is associated directly with the preceding node state probabilities. In a PRA application, BBNs serve to limit variable interaction to those nodes that have direct interaction, simplifying the updating process and contributing to the computational efficiency of the model.

Support for Bayesian belief nets. Though it is not necessarily unique in this regard, aviation, and especially aviation safety, is often regarded as a field comprised of varying degrees of art and science. In this context, it is often the case that *art* is used as a euphemism for uncertainty, implying that practitioners rely on wit and experience to reach successful outcomes. This paper does not fully address the particulars of expert opinion in comparison to strictly stochastic techniques, but worth noting is that this dichotomy of knowledge can lead to substantial gaps in organizational understanding. In pursuit of safety improvements, the unification of knowledge toward process understanding and quality management is a central tenet (see FAA, 2010c; ICAO, 2009; and Stolzer, Halford, & Goglia, 2008); BBNs can leverage organizational knowledge in qualitative and quantitative forms to support more accurate prediction and decision-making through a single form of representation. In representing the multifaceted knowledge that describes a domain, Bayesian networks amalgamate this knowledge consistently and completely given that the constraints of the network are satisfied by only one probabilistic distribution per node (Darwiche, 2010). Because BBNs rely on both probabilistic and causal semantics, they are a natural platform for the representation of this combination of prior knowledge and new data (Nadkarni & Shenoy, 2004). Modern propagation algorithms allow this unification of knowledge through BBNs to occur in a computationally efficient format, and even in topologically complex networks such as may arise in aviation, BBNs can compactly provide a robust inferential tool.

Unifying knowledge is one matter, but communicating it is another. In domains characterized by uncertainty, as in risk assessment in high reliability fields such as aviation, capturing and communicating the complex behaviors of a system to many and varied stakeholders is a daunting process when undertaken in conventional frequentist reasoning. BBNs represent causal connections and dependencies while also capturing uncertainty to intuitively communicate the state of a domain, even in the face of a dynamic operating environment. Conrady and Jouffe (2011) refer to the effectiveness of BBNs in this regard as creating a “portable knowledge format” that succinctly encapsulates the state of a domain of knowledge and the multifaceted interaction of the variables within that domain (p. 3).

As Pearl and Russell (2003) assert, the most remarkable feature of BBNs is that "they are direct representations of the world, not of reasoning processes" (p. 158). It can be argued that human cognition follows a general pattern of solving probabilistic, rather than logical, inferential challenges (Oaksford & Chater 2009). Instead of relying on instinctual, undefined expert opinion to perform risk analysis and subsequent mitigation evaluation and selection, BBNs provide a means of identifying what is already known and what will result from future processes and circumstances. Dynamic networks allow propagation of reasoning processes to flow naturally and in closer harmony with perceptions of reality, as opposed to the rule-based systems one may otherwise encounter (Pearl & Russell, 2003). Naïve PRA methods identify only discrete events, often outside their operational context, and many advanced methods operate in an operational vacuum, far removed from the dynamic world in which they function. In contrast, BBNs can capture not only conditional probability and uncertainty over time, but can also allow for dynamic assessment of interventions, such as safety improvements, within the conditionally dependent model.

Bayesian networks and causality. As a means of understanding RI events, a chief advantage to the use of BBNs is that causal inference is possible. Conrady and Jouffe (2011) note that one reason Bayesian networks have seen a rise in popularity in recent years is the possibility that they may allow discovery of causal structures otherwise hidden within raw statistical data. BBNs construct a causal structure as a function of formalizing causation through identification of direct interactions from a given variable set, something experts and people on the whole are good at (Darwiche, 2010). The nature of the directed graphs central to BBNs is both a feature of and a foundational element of the idea that causal structure can be revealed by these methods. The directed acyclic graphs (DAGs), as in Figure 7, are inherently a system of processes that lead to causal interpretation. When a BBN is constructed such that the directed structure is consistent with causal theory, the network is capable of updating probabilities based on the interaction of an intervention inserted into the model (Darwiche, 2009).

Nadkarni and Shenoy (2004) suggest a causal mapping approach to the construction of Bayesian networks as an intuitive and systematic means of building a BBN. BBN construction is generally a function of either a data-based or a knowledge-based approach. The former is a function of deriving independence relationships from the data; however, in the context of runway incursions, the availability of data is questionable. While a data-based approach alone may produce a defensible network structure, adding elicitation of expert judgment, especially in the present framework of rare events, is critical to maintaining network sensitivity toward a more complete domain understanding and greater effectiveness. Causal maps provide a starting point for the representation of knowledge necessary to create a BBN. As cognitive maps, causal maps capture causal knowledge of experts that is otherwise difficult to ascertain, and they do so more descriptively than regression or structural equation methods (Nadkarni & Shenoy, 2004). Causal maps are made up of three components: a node representing causal concept, a link representing the causal connection among concepts, and strength representing the causal value of a connection Nadkarni and Shenoy (2001). Figure 8 illustrates a simple causal map of expert opinion in the context of prediction of an aircraft runway overrun. In the Figure 8 example, the unidirectional arrows indicate causal connections, with a positive or negative influence. A high level malfunction in an aircraft

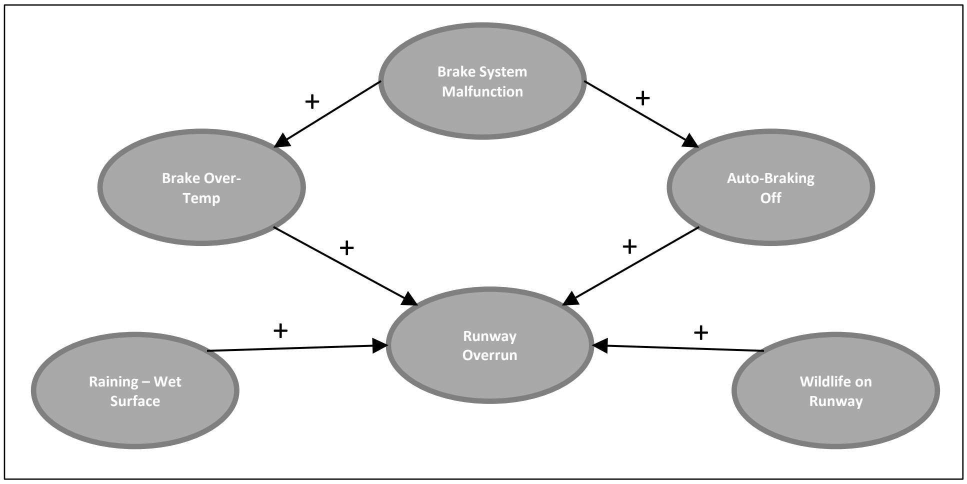


Figure 8. Causal network.

braking system, for instance, increases the likelihood that the auto-braking system will not function (positive relationship). This continues along the causal path to a higher probability of runway overrun. Pearl (2009) refers to the utility of causal BBNs as "oracles for intervention", a reference to the distinct parent-child node relationships and modular characteristics of the network that facilitate the evaluation of relationships with a minimum of changes to the network structure (p. 22). In essence, a causal network is a Bayesian network, with "the added property that the parents of each node are its direct causes" (Conrady & Jouffe, 2011, p. 10).

Practical application and considerations of BBNs. On the surface, employing Bayesian inference in PRA seems logical enough; and when presented in terms of a practical aviation example, this is even more apparent. Before suggesting a review of an exemplar scenario, it is first appropriate to mention that PRA, and by virtue of association, MCMC and BBN, are not always the best tool in every application. This is not to imply that probabilistic modeling is not a powerful means of inference, but rather that the time and resources involved may not be appropriate for every situation. As an

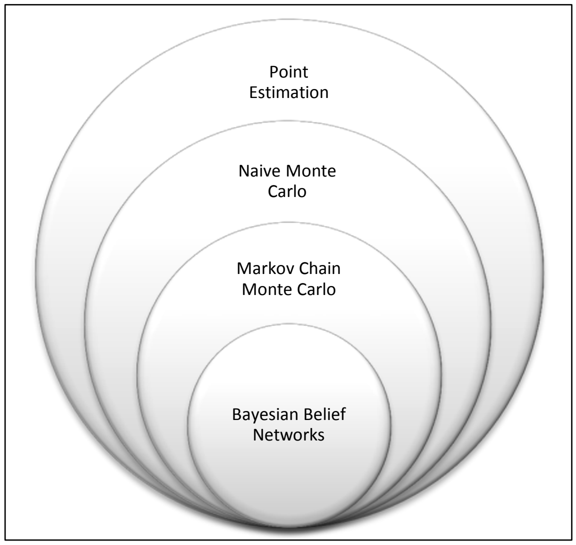


Figure 9. Continuum of PRA tools.

example, and as illustrated in Figures 5 and 9, deterministic point estimation of risk remains a useful tool for relatively simple processes, and even as a first step for scenarios that will eventually utilize more complex methods. In aviation, however, as previously discussed, the complexity of operations and rarity of failure provides ample opportunity for application of PRA in the Bayesian framework.

As has been discussed, the probability estimates in many attempts to assign rare event probability (shown also in Tables 12 and 13 in the example scenario in Appendix A) are simply contrived point estimates, and as such do not account for uncertainty. What this example expresses is how correctly integrated random sampling can in fact allow for the uncertainty that characterizes rare, but high-consequence events such as aviation accidents. By treating the network as a Markov chain, similar distributions and node correlations, either through Monte Carlo sampling or other methods such as the junction tree sampling algorithm (as in Borsotto et al. (2006)) or incorporation of the copula-vine approach (Bedford & Cooke, 2002), can be created from which an expected conditional value may be approximated from the sample mean of a function of simulated random variables for each node.

Whereas sampling in a naïve Monte Carlo method via importance sampling or rejection sampling is possible and will produce results reducing variance, both methods are limited in their applicability for the approximation of joint distribution. In the context of the model described above, and for similar network models where the joint distribution is unknown but marginal probabilities are known with at least some level of certainty, the Gibbs sampling method is a means of sampling the posterior network distribution. The Gibbs sampler is intuitive in its sequential method of sampling from a target distribution within MCMC algorithms (Lynch, 2007). However, the Gibbs sampling algorithm applies primarily to parametric or discrete parametric models. Alternatively, the joint tree or rank correlation methods of sampling are also appropriate to the model proposed here. Given that interval representation is an important element of the expert elicitation process described in the following sections, these algorithms can be used to achieve more informative priors by narrowing intervals that may otherwise widen as a result of propagation (Borsotto et al., 2006). Readers are directed to Borsotto et al. (2006), Bedford and Cooke (2002), Congdon (2003), Hanea, Kurowicka, and Cooke (2006), Koller and Friedman (2009), Gelman et al. (2004), and Gamerman and Lopes (2006) for more detailed descriptions of the sampling methodologies available for use in BBNs.

Examples of BBN use for probabilistic causal modeling in an aviation safety context are few, and the review of literature on the subject revealed only one study with substantial commonality to that proposed here. Lechner and Luxhøj (2005) conducted case studies of three specific RI accidents using the Aviation System Risk Model

(ASRM) described by Luxhøj and Coit (2006) and intended to supplement fault and event tree models by representing interdependency and dynamic, interactive causation using BBN modeling. Lechner and Luxhøj (2005) used the Human Factors Analysis and Classification System to map influence diagrams and the structure of a BBN for the three accident case studies of interest. The authors address the rarity of events by asserting that the case study approach is in fact generalizable to the broader population of RI events. While there is demonstrated merit to the case study method of evaluating causal structure, it can also be construed that reliance on past events is inherently a reactive, forensic process. In the case of Lechner's and Luxhøj's study, case studies are warranted because the research sought also to evaluate the impact of mitigation strategies after the fact. In the present research, where it is suggested that causal structures must first be understood in a dynamic, holistic setting, case studies artificially restrict generalizability and limit the extent to which uncertainty is accounted for in determining causal interactions of future RI events. It is this differentiation that supports the need for methodology described here.

In addition to Lechner and Luxhøj's work, a study by Morales, Cooke, and Kurowicka (2008) stands out among the literature as having particular bearing on the present research. In describing causal modeling methods for air transport, the authors utilize BBNs, building the models, as did Wang (2007), on the basis of more traditional PRA tools such as ET and FTA. Their use of BBNs focused primarily on human error and probabilistic influence on error by a complex system of interdependent factors. This is relevant not only because of the relationship of human error, but because of the successful use of BBNs as a tool for causal inference in an air transportation setting. Also focusing on causal modeling in transportation, Hanninen and Kujala (2010; 2012) sought to evaluate ship-ship collision causation using BBNs. The authors note in their conclusions about the success of the model that "a Bayesian network causation probability model also provides the means to examine how the underlying factors influence the collision probability" (Hanninen & Kujala, 2012, p. 32). Although this study addressed ships, not aircraft, the validity of BBNs as a tool for discovering paths or nodes of influence among a complex web of variables and their underlying parameters is clear.

Theoretical Considerations

Perhaps the element most often overlooked or underestimated in the creation of probabilistic models, including BBNs, is the theoretical architecture underlying the model itself. Theoretical concerns in the present research are in the form of defining the construct of study, identifying and codifying causal theory, and understanding the theoretical applications and limitations of the data generation process. These must each be addressed comprehensively before a model can be developed, tested, or deployed with any measure of success.

Causal theory. Causation is revealed in the conditional interdependencies that characterize an underlying structure of data (Pearl, 2009). In frequentist statistical analysis, covariation, not causation, is the basis upon which inference is made. In the present context though, a more natural inferential model is proposed within the Bayesian framework. Temporal precedence is a reasonable, if not implicit indicator of causation, but it is not required. Intransitive dependencies also exist that reflect on the natural game of induction the researcher is often forced to play (Pearl, 2009). The directed acyclic graphs discussed previously serve as a foundation for discovering causal structure in this way, but the acyclic structure that defines the graph must be derived somehow from observation, however limited. To inform the construction of the acyclic structure of parent and child nodes, it is possible to employ a familiar technique from PRA, the fault tree.

In FTA, an event sequence is represented based on the relationships between pivotal events and consequences. In the context of fault trees (FTs), system perturbation is captured by initiating events (IE), pivotal events (PE), and end states. Fault trees, as shown in Figure 10, can be useful in modeling complex pivotal events, but non-binary event outcomes make it infeasible to rely solely on FTs for modeling of dynamic system behavior (NASA 2002; Roelen et al., 2003). It is not difficult to see the relationship between the structure of the FT and a BBN, and the FT may be used as a basis for node selection in a BBN, though it is by no means the only way to arrive at the structure of the BBN.

Although the literature suggests that FTs may be used as a method of populating the BBN acyclic graph, it is worth noting that FTs quite often are very heavily focused on

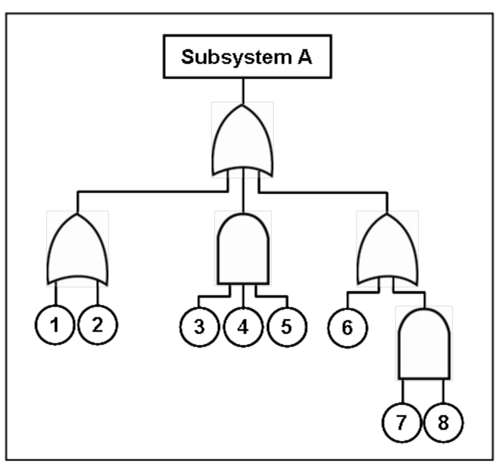


Figure 10. Fault tree.

failure from an engineering perspective (NASA, 2002). In and of itself, this does not dilute the effectiveness of an FT, but from the perspective of creating an inclusive causal model, multiple domains must be considered as part of the construction of causal theory. On the surface, this idea is simple enough, but one must bear in mind that it requires the merging of two models that are philosophically different (Mosleh, Dias, Eghbali, & Fazen, 2004; Roelen et al., 2003). The engineering approach described in the discussion on FTA is an effective structure for the identification of failure paths in a system, but it does not account for what Reason (1997) calls organizational accidents.

To this end, the extension of the research by Joslin, Goodheart, and Tuccio (2011) is appropriate in that it addresses a "holistic understanding of the contributory elements of runway incursion incidents" outside of the ordinary constraints of ET and FTA (p. 2). This *holistic* perspective addresses accident causation not only from the mechanistic framework of discrete pathways of failure, but also from the perspective of management factors, such as illustrated in Rasmussen's (1997) model of functional abstraction (Figure 11). Of course, as Mohaghegh-Ahmadabadi (2007) and Roelen et al. (2003) prominently note, the organizational and engineering approaches, which could arguably be described as being qualitative and quantitative, respectively, must be somehow combined in the model structure if it is to be truly encompassing of the explicit and latent factors that interact to create an accident as is the objective here.

Wang (2007) discussed the importance of considering the total environment within which a system operates in a comprehensive model. If dynamic interactions are evaluated for risk and causal structure separately, how is one to combine disparate

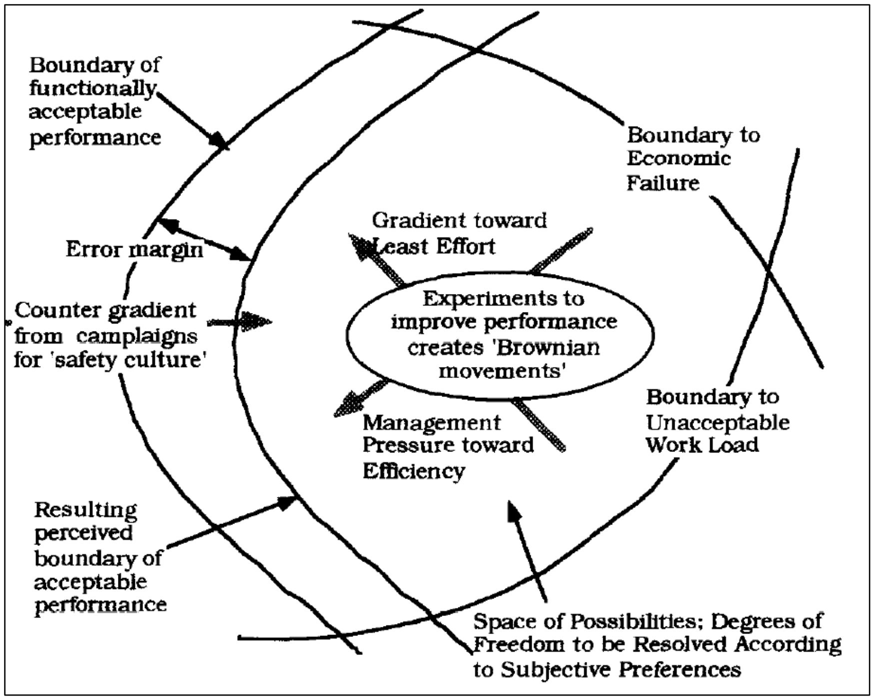


Figure 11. Migration of behavior toward unacceptable performance (Rasmussen, 1997).

methodologies? Wang (2007) proposes a hybrid causal model that captures the complex interactions of a system with an environment comprised of regulatory, physical, and socio-economic components as shown in Figure 12. To create such a model, Wang (2007) began with basic PRA tools such as event sequence diagrams and extended them to also interact with Bayesian networks. Pai and Dugan also proposed a hybrid causal model that captured the complex interactions of a system with an environment comprised of regulatory, physical, and socio-economic components as shown in Figure 12. Fault trees and similar methods are a graphical representation of logical semantics that allow reasoning about causal paths to failure, but as are many traditional PRA methodologies, they are often static and limited to binary probability states (Dugan, Pai, & Xu, 2007).

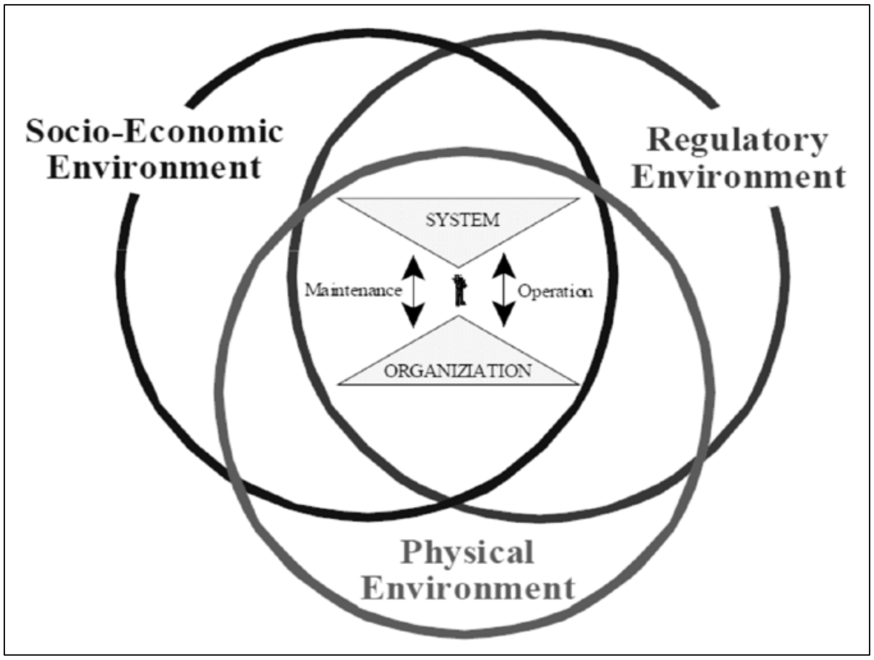


Figure 12. Interaction between systems and environment (Wang, 2007)

Used alone, and even in interaction with BBNs, traditional tools such as event and fault tree analysis are not capable of analyzing causation dynamically and across domains. As discussed previously and as illustrated in the literature, BBNs can accomplish this, and they are flexible to inclusion of traditional PRA methodologies as a means of building the model structure. Once a structure has been created, the matter of populating network nodes with data must be addressed.

Data generation. Because obtaining data collected by authorities such as the FAA is addressed in the Joslin et al. (2011) study design, it is unnecessary to expand upon it in great detail here. The emphasis of the Joslin et al. (2011) methodology is on reviewing data in a mixed-method framework, relying in part on expert raters to evaluate pilot self-reported narratives and code them (but not rank) in terms of causal contributors. This method is appropriate because it prioritizes qualitative information based on the collection method, and uses it to enhance the inferences that would ordinarily be drawn only based on quantitative information. For more information on the data available through the NASA Aviation Safety Reporting System (ASRS) and the methods for extraction of RI data from this information, readers are directed to Joslin, et al. (2011).

While the coding scheme outlined by Joslin et al. (2011) is useful as a starting point for RI study in that it identifies thematic causal elements, the resolution of the data prove insufficient as a sole source for derivation of probability distributions as are required in BBNs. As a means of achieving this level of granularity, the low frequency of these events requires supplementation of data, and in this case, expert elicitation is appropriate. In the Bayesian context, existing data is used to establish the prior belief about the distribution of the unknown, and expert opinion is solicited to inform and update posterior probabilities as the model is deployed. Expert opinion can of course be used to establish a prior in the same sense as the FAA data can, and the specifics of this implementation are beyond the scope of this paper. In either event, the method by which expert information or opinion (purposely identified as distinct concepts, as in Kaplan (1992)) is collected, structured, and aggregated must be carefully considered in order to limit bias and propagation of erroneous data throughout the model.

Expert Elicitation

Insufficient or unobtainable data has led to the development and use of probabilistic risk and safety assessment in a variety of fields, from nuclear power (DeWispelare, Herren, & Clemen, 1995) to public health (Hoelzer et al., 2012; Knol, Slottje, van der Sluijs, & Lebret, 2010), and security (Levine, 2012). Cooke (1991) references Plato's allegory of the cave in his description of expert elicitation as somewhat of a contradiction in the scientific world (p. 3). In his parable of the cave, Plato describes the evolution of knowledge from its lowest orders of imagination and belief to *episteme*, the highest levels of knowledge, separating these strata with his so-called *divided line*, shown in Figure 13 (Heidegger & Sadler, 1988). Although the process of expert elicitation could be supposed to deal predominantly with the visible realm of illusion of belief, Plato himself reasons that in discussing these abstractions often and deeply enough, it is possible to eventually achieve knowledge of what he often refers to as *forms* (Heidegger &Sadler, 1988). With this in mind, this research and the body of knowledge suggest that thorough expert elicitation knowledge can evolve toward new solutions while maintaining scientific rigor.

Knowledge (*Episteme*)		**Opinion (*Doxa*)**	
Intelligible Realm (Forms)		Visible Realm (Substance)	
A Pure thought (*gnosis*)	B Reason (*dianoia*)	C Belief (*pistis*)	D Illusion (*eikasia*)

Figure 13. Plato's *divided line* (adapted from Heidegger and Sadler (1988)).

Support for expert elicitation. The limitations in both resolution and availability of RI data discussed previously establish that in order to perform more detailed analysis of the problem, data sources outside those conventionally used in this regard must be explored. When data are sparse, the literature demonstrates that a scientifically structured and mathematically transparent method is a robust means of assessing uncertainty and cumulative probability, both of which are important elements in the study or RI events and the causal factors and interactions that lead to them. Because expert elicitation is a fundamentally "interdisciplinary" method, it is particularly well suited to the investigation of complex phenomena with causal influences that extend across domains, such as is often the case for RIs (Knol et al., 2010, p. 27).

Its utility and multidisciplinary applicability notwithstanding, the true scientific value of expert elicitation is, as it is with almost all analytical methods, dependent on methodological appropriateness. Some current methodologies, such as the consensus-based methods used in the CAST studies discussed earlier, approach this issue directly but remain open to scrutiny because of the somewhat black box nature by which expert opinion is gathered and combined. Within the existing body of knowledge, alternative means for elicitation and aggregation of expert opinion exist, and the following discussion reviews techniques most applicable to the study of RI events in a Bayesian network context.

The continuously variable structure of airspace, technology, system capacity, and even training paradigms makes it impractical to assess the probability of RI events in the frequentist sense, where probability is treated as the long-run tendencies of events that will eventually converge upon the true proportion of a population. Instead, the Bayesian interpretation of probability as a degree of belief is a uniquely appropriate method, where past experiences and knowledge of the likelihood of events can be expressed in terms of a prior distribution function.

In the context of the research presented here, these prior distributions are established as an expert-elicited informative prior. Though some criticism has been leveled at this technique (e.g. Brooker, 2011) on the basis of what Gill refers to as the "supposedly personal-subjective nature of priors", it is suggested that very few, if any, would comfortably approach any model or its related specifications without at least some cursory prior understanding (Gill, 2008, p. 136). That said, priors vary from flat, non-informative (uniform distribution) to those that are strongly influential. In any case, accounting for the existing scientific knowledge in the field is addressed in the present study as a matter of scientifically structured expert elicitation.

Methodological review. Though Brooker (2011) and others have questioned the mechanisms on which expert elicitation is based, the methodology has been empirically established over a number of independent studies at the Technical University of Delft in The Netherlands. The experiences there have shown the utility of the technique, and have answered some of the questions that plagued the theories early on as a result of the rarity of available studies for review. With the benefit of substantial experience in elicitation and aggregation models, the following observations have been made (Cooke, 2004, p. 317):

1. Experts don't mind performance measurement.
2. Experts are leery of 'non-objective' or psychologically based methods, and are suspicious of the 'academic sandbox'.
3. Experts have no problem understanding (subjective) probability and no problem quantifying degree of belief in terms of quantiles of a subjective probability distribution
4. Experts are *not* uniformly overconfident, though overconfidence certainly does arise.
5. It is *always* possible to find suitable calibration variables.
6. In general, though not always, the performance based combination of expert

> judgments performs better, in terms of calibration and information, than an equal weight combination and also better than the best expert.

A separate problem to the elicitation of information from multiple experts is addressed by Wu, Apostolakis, and Okrent (1990), who note that "In PRA, an important issue related to knowledge representation under uncertainty is the resolution of conflicting information or opinions" (p. 170). Considering that the impetus for using expert judgment is that substantial scientific uncertainty impacts on a modeling process, it is reasonable to assume that the experts themselves are not certain. As such, agreement among experts is almost certainly unattainable. However, some means by which differing expert opinions may be translated into a structured consensus of sorts must be specified if the elicited information is to be treated as data.

The elicitation and aggregation of expert opinion generally falls into two methodologies: those characterized as behavioral approaches and those that rely on mathematical calculus, whether in non-Bayesian or Bayesian models (Clemen & Winkler, 1999). While behavioral approaches are useful from a broad, common-sense perspective, they generally fail to satisfy the conditions of rational consensus that scientific, structured expert judgment requires, more specific elements of which are included in Appendix C and can be found in Cooke and Goossens, 2006. It is generally agreed (Clemen & Winkler, 1999; Lin & Bier, 2008; Mosleh et al., 1988) that mathematical methods for aggregation produce better results because of this adherence to the principles of scientific inquiry.

The natural question that arises in the context of this research is whether or not a Bayesian approach can also be used in the process of expert elicitation. Procedures that approach the problem of elicitation and aggregation from a Bayesian perspective have been presented (Clemen, 1986; Clemen & Winkler, 1999; Ouchi, 2004); however, Bayesian methods can be problematic in practice. As one example, many Bayesian elicitation and aggregation models assume that experts are independent. Furthermore, the use of the Bayesian paradigm in elicitation and aggregation requires that the practitioner formulate joint distribution over the variables of interest, the seed variables used for calibration, and the experts' distributions over those seed variables and the variables of interest (Cooke & Goossens, 2006). This is not an impossible task, but it is mathematically complex, and is subject to the difficulties of overcoming resistance to the group decision problems that arise because a group of rational individuals cannot be treated as a single rational individual (Cooke, 2009). Though the Bayesian approach meshes harmoniously with the present research from a theoretical standpoint, it is impractical from an operational one. Instead, Cooke's (1991) classical model provides a more readily workable procedure for eliciting and combining expert opinion. Cooke's classical model has been applied in practice over dozens of studies conducted in conjunction with the Technical University of Delft, and its value has been established empirically (Cooke & Goossens, 2006).

The classical model. The classical model operationalizes the mathematical principles for combination of probabilities such as those summarized by Genest and Zidek (1986) and is perhaps the most widely used method by which expert judgments can be combined (Clemen, 2008). Cooke (1991) borrows from Savage's theory of Rational Decision as a basis for the probability calculus that describes many of the ideas surrounding uncertainty. The classical model has been demonstrated in a variety of applications including food safety (Brown et al., 1997), finance (Bakker, 2004; Qing, 2002), geologic and infrastructure erosion (Brown & Aspinall, 2004), public health (Tyshenko et al., 2011; Winkler et al., 1995), volcanic eruption assessment (Aspinall & Cooke, 1998), equipment failure (Bedford, Quigley, & Walls, 2006; Akkermans, 1989), and nuclear risk assessment (Goossens & Harper, 1998). In some social science contexts, the fact that Cooke's model purposely does not capture consensus in its aggregation of results may be viewed as a disadvantage (Albert et al., 2012); however, the present approach fits well with the classical pooling approach that has "stood the test of time" (French, 2011, p. 183). The classical model (Cooke, 1991) meets the theoretical and operational requirements of the research at interest in that it uses real data to evaluate experts and assign weights to their assessments (Clemen, 2008). Its unique approach to empirical control and performance-based expert scoring lends an empirical formality to the elicitation process, contributing to the transparency with which a model such as the BBN discussed here may be developed. The algorithms for elicitation and aggregation of opinion within the classical model are presented in detail as Appendix C.

Summary

RI occurrence in the U.S is a problem that has been well established in terms of severity and frequency. While a number of mitigation strategies have been demonstrated with promising results, the rate at which RIs occur in the U.S. airspace system has increased over the past several years and continues to increase in the face of concerted efforts to stem this trend. Investigation of RI causation has typically focused on individual factors within isolated domains of knowledge, despite a general recognition that these factors do not operate separately, instead interacting dynamically in the sequence of events leading to an undesired event (Vernaleken, Urvoy, & Klingauf, 2007). As the literature demonstrates, existing studies are limited by the following gaps:

- They generally address only one method of determination of causal factors rather than examining causal interaction.
- Research of RI causation is typified by a focus on one domain, rather than by a holistic approach that addresses the needs of multiple stakeholders.
- A reliance on case studies or deterministic estimates of frequency are common in the RI literature, and most studies lack a probabilistic approach that effectively captures uncertainty.
- Many efforts at understanding RIs look at individual causal factors and map intervention strategies that are equally limited to discrete events rather than designed to work within a complex sequence of interacting factors.

To address the dynamic nature of RI causation despite limited data, stochastic processes such as Monte Carlo simulation are useful. However, such methods do not capture conditionality, nor do they incorporate uncertainty in the context or rare events or sparse data. To this end, Bayesian belief networks are an appropriate tool for investigating causal pathways to RI events. Such methods also support more robust decision-making through sensitivity analysis and evaluation of intervention strategies.

Because RIs are relatively rare events from a statistical analysis standpoint, a structured means of data generation is necessary so that a network model can function. To accomplish this, expert elicitation is an appropriate methodology. Expert elicitation can be used to express uncertainty in the language of probability, and it can be conducted such that scientific scrutability is maintained. In combination, these methods provide a means of investigation into the causal interactions that lead to RIs that is integrative, flexible, and dynamic, and that will allow more focused strategies to combat them through a fuller understanding of their structure.

CHAPTER III

METHODOLOGY

A review of the relevant literature supports the use of modeling the complex and dynamic interactions that lead to RI events. However, as has also been demonstrated in the literature, many investigations into the causation of RIs are one-dimensional. As a means of examining causal patterns across theoretical domains and in the face of uncertainty and sparse data, predictive, simulation-based modeling through BBNs is a more holistic method than what has been used to date, and it is flexible to system variance and technological change. The purpose of this study is to evaluate the feasibility and effectiveness of Bayesian belief network models, supported by structured expert elicitation, as a tool to examine causal factors and dynamic causal paths to RI events in the U.S.

As discussed previously, this study addresses two questions: What are the interacting causal factors that lead to RIs in the U.S.? And, can runway incursions in the U.S. and their dynamic causal factors and interactions be modeled through the use of a Bayesian belief network supported by expert-elicited data? The methodology outlined herein is directed at answering these questions as fully as possible.

Research Approach

The methods utilized here are more appropriately described as an algorithm, wherein three structured phases of research are undertaken toward the eventual objective of developing a functional predictive model. The phased approach described here is iterative and additive in large part, though some elements of the research approach may be accomplished concurrently. While the following discussion addresses each phase independently, these phases were naturally subject to some temporal overlap.

Phase 1: Runway Incursion Data and Causal Factors

Although quantitative data describing occurrences of RI exist, the resolution with which these events are described is limited, and what information is available fails to capture the emic, or insider, perspective. Whereas the more common etic perspective looks at RI events in the more general sense, investigation of RIs from the emic viewpoint provides a more specific understanding of the conditions and impact of RI events and the causal interactions that lead to them (Ng & Earley, 2006). To facilitate this emic view and to more holistically examine the causal structure of RI events, textual data from open-ended narratives (ASRS reports) were evaluated by a panel of raters to determine causal factors for each RI within the sample. This merging of emic and etic perspectives toward a fuller contextual understanding of RI events was accomplished as a mixed methods process, and in the context given here, was approached as an explanatory, sequential element of the research as a whole. The qualitative rater data were combined with quantitative FAA RSO data and the literature review and examined in a multi-modal, quan→QUAL analytical process.

Data collection and generation. The data collection sites for this study were from two sources of publically available U.S. historical data: RI pilot deviation (PD) type incidents from the FAA RSO database, and pilot-reported pilot accounts of RI incidents from the NASA ASRS database. Consistent with the sampling constraints described herein (to limit the population to the period to which the current RI definition applies), the data were considered only if they stemmed from an occurrence within the period inclusive of calendar year 2008 through calendar year 2012. Using basic database query strings, the data were subjected to a specific, targeted search string in order to limit the sample to PD-type RI incidents. Using a technique similar to Joslin, Goodheart, and Tuccio's (2011) expert rating of RI event causation and severity, the data were collected via a spreadsheet template and drop-down menus were utilized to facilitate selection accuracy. As in Joslin, Goodheart, and Tuccio (2011), the following procedure was followed:

a) Unique text for identification was recorded for each ASRS case. This helped ensure data entry accuracy by guarding against a rater recording data in the wrong row.
b) Causal Factor Taxonomy of ICAO 9870, Appendix D*, Manual on the Prevention of Runway Incursions* were used as the set of available causal factors for rater selection.
c) ASRS cases were sorted using random number generation.
d) A limited number of ASRS cases were evaluated per day to prevent expert rater fatigue.
e) Raters conducted their evaluation in a "single pass" for each case without going back to change or re-evaluate causal factor assignment.
f) Raters were limited to a maximum of five causal factors per report as a means of constraining the volume of data collected, and consistent with the pre-test results of the Joslin et al. study.
g) Cases were marked as an "exception" if they did not meet the criteria (FAA/ICAO definition) for a runway incursion or if insufficient data exist.

Population/Sample. The data were drawn from the population of U.S. RI events (regardless of Federal Aviation Regulations under which an aircraft was operated) as recorded by the FAA RSO and in ASRS reports at airports within the U.S. These data captured all reported RI events in the U.S. for the timeframe of interest and supplemented them with all available narrative reports over the same frame, respectively. Operational segment or regulatory part was not differentiated on the basis that a chief objective of this research is to maintain a holistic perspective on RI causation, and purposive narrowing of operational scope could have unintentionally limited capture of important data points. The sampling frame, consistent with the definition of an RI at the time of this study, was limited to data from January 2008 up to but not including January 2013. Within this frame, quantitative RSO data included 6,185 RI cases for analysis, including primarily ATC-reported RI events and quantitative descriptors. ASRS data included 87 cases reported in qualitative, narrative form by flight crews (sufficient according to Cantor (1996) for evaluating rater agreement using Cohen's kappa) describing RI events to be evaluated by subject matter experts for causation.

Sources of the data/Rater selection. Data used in this phase of the study existed in three forms: RSO RI data and ASRS RI reports, rater-identified causal factors, and expert-elicited judgments. The first data set is publicly available via ASIAS and the RSO directly. Derivation of the expert opinion data is an element of this study, though in this first phase it is not intended that expert data will be collected using the same structured elicitation process outlined in the following phases. For this phase only, raters were selected using the following criteria as guidance:

- independence (from one another and the research),

- diversity of operational experience,
- interest in and availability for the project,
- flight qualifications and experience, and
- familiarity with RI causation and mitigation strategies in place or in development.

To reduce bias in the selection of the raters, scoring criteria shown in Appendix M was used to evaluate potential SMEs and make rater selections based on the prospective raters' submitted biographical information, curricula vitae, and peer nomination. The raters, as described previously, were asked to assign causal factors based on the ICAO Taxonomy (ICAO, 2007) to narrative ASRS reports. In this phase of the research, three raters were used as a measure against threats to reliability as experienced by Zuschlag (2005) in his review of ASRS reports for a similar purpose.

Descriptive statistics. In the first phase, which involved identification of causal contributions to RI events based on RSO data in combination with expert opinion, descriptive statistical analysis was utilized in the form of a frequency count of each ICAO causal category used by each rater as well as basic analysis of data from the FAA RSO on RI events. A sort operation of these frequency counts will be used to gain insight into the most common causal factors among ASRS reports examined. Descriptive frequency counts also allowed for identification of causal codes not used by any rater in the ASRS review process.

Interrater reliability. In the context of this study, reliability was applicable insofar as it applied to the expert rating of causal contributors in the first phase of the research. To establish interrater reliability, the expert-assigned causal factors were recoded to a numeric value, assessed for normality, and if necessary, transformed prior to analysis for reliability using Cohen's Kappa (Leech, Barrett, & Morgan, 2008). As in Joslin, Goodheart, and Tuccio (2011) SMEs were used to examine the plain-language narrative ASRS data and re-code the reporter's comments into nominal, quantitative data. Causal codes and the reliability with which they are assigned by expert reviewers were assessed using three methods: union of causal factors, intersection of causal factors, and Cohen's Kappa as demonstrated in and for the reasons described in Joslin et al. (2011). Cantor (1996) suggests that assuming an *a priori* rater agreement of roughly 50 percent, and an error margin of 30 percent, 44 cases should be evaluated. Given that 87 RI cases were initially considered by raters, that minimum sample size was met.

Cohen's Kappa interrater reliability. Interrater reliability, the statistical measure of agreement, or consistency, between the raters on the same variables, was evaluated using a Kappa statistic. Cohen's kappa provides a measure of interrater reliability of two raters assigning one nominal code to a list of items and ranges between 0 and 1.00 with a value of 0.70 generally considered satisfactory (Leech, Barrett & Morgan, 2008). Cohen's kappa is defined by Equation 3:

$$\kappa = \frac{p - p_e}{1 - p_e} \tag{3}$$

where:

p is the proportion of units where agreement exists; and

p_e is the proportion of units that would be expected to agree by chance alone.

Thus, Cohen's kappa is the agreement between observers (SMEs) adjusted for that proportion of agreement that would ordinarily be expected to occur by chance. Though percent absolute agreement is also commonly used as a measure of agreement among raters where multiple levels exist, no provision for level of disagreement or correction for chance is available. Because three raters were presented with 47 possible codes to fill a variable response space between 0 and 5 codes per case, the use of a traditional kappa measure, such as Fleiss' kappa or Cohen's kappa, becomes computationally untenable in its ordinary application. To overcome this issue, a Structured Query Language (SQL) query was used to compare raters pairwise for each potential level of match (i.e. 1, 2, 3, 4, or 5 matching codes across the rater pair). The actual SQL code used to perform this function is presented in Appendix G. Given the possibility that each rater selected a unique set of causal codes, dummy variables were automatically inserted to meet the requirement of the statistic that the number of codes being compared be the same. Using the general procedure from Joslin, Goodheart, and Tuccio (2011) for reach rater pair, the SQL operation was used to sequentially evaluate matching codes as independent operations, with each generating a unique but matching character string between the raters. When matches were not present, a unique string per rater was inserted as a placeholder. At the end of the procedure, dummy ASRS records were inserted to comply with Cohen's procedural requirement of both raters using all possible codes. While this method overcame some basic mathematical limitations of kappa in this application, it did not capture the total domain of possible codes. It does maintain the basic tenet of kappa in that chance is accounted for, however limited by the inability to account for all possible causal factors.

To assist in describing the interaction and agreement (or disagreement) between raters more accurately, SQL was also used to perform union and intersection functions from the rater responses. In this instance, union is intended as the total number of codes that make up a set across all raters for one case – the number of unique codes. Intersection describes the number of codes used by all raters across a single case.

Union of causal factors. Because each expert rater was asked to assign up to five ICAO causal ratings per ASRS report (see Joslin, Goodheart, & Tuccio, 2011), it was possible that many codes could be assigned – up to five codes per rater – if the codes were considered as the union of the set. The union technique provided a distinct list of total causal codes per ASRS report as well as a count of how many union codes were used per ASRS report across the SME raters.

Intersection of causal factors. In contrast to the union operation per ASRS report, an intersection of ratings was also assessed for each ASRS report. The intersection identified only those codes used by all raters per ASRS report. The intersection technique generated a list of distinct causal codes per ASRS report used by all raters as well as a count of these codes.

Merging of data streams. To enable development of a model representative of RI causation in the US, the data from the RSO and ASRS reports were combined with results of the literature review in conjunction with SME input resulting in a pool of potential causal factors. This set of causal factors of RI events is included in this study as Appendix F. Descriptive quantitative statistics were weighed against review of ASRS reports highlighted by the SMEs and with consideration of the existing body of research.

Phase 2: Belief Network Model Creation

The initial belief network model structure was generated based on the causal contributors identified in the review of RI literature in combination with the causal factors and interactions identified within ASRS reports in the preceding phase. Once a list of causal factors was identified from these sources, an influence diagram as discussed in the literature review was developed in preparation for building the BBN. A panel of three experts was utilized to iteratively validate an initial model structure by evaluating the model for completeness and accurate causal interaction and direction. Because of their familiarity with the model and purpose of the research, the SMEs used for ASRS case review in phase one of the study were also used for review of the BBN model to ensure it accurately reflected the combined knowledge of the relevant domains.

Constructing the network model. Building a BBN is a process of structuring the graphical model and defining the causal dependencies within the graph itself, and this process can be generally described as a qualitative one. As discussed, the model structure began with an influence diagram. This process was a means of identifying the “causal, functional, or information relations among the variables” (Kjaerulff & Madsen, 2008, p. 117). The variables of interest in the present research were identified from the literature in conjunction with the first phase of the methodology outlined here. In the case of each variable set to represent a node in the BBN model, a unique set of mutually exclusive events was described. Examples of such events or states are found in the causal literature and include items such as intersecting taxiways (yes or no), restrictions to visibility (yes or no), AMASS system in place (yes or no), etc. Variables within the model were identified as problem variables of interest, information variables for which data may be known (background or outcome variables), and mediating or intervening variables, all of which are connected by edges that represent notions of causality, or as previously discussed, “the way the world works” (Koller & Friedman, 2009, pp. 52-53). Figure 16 illustrates a simple, exemplar causal network with each of these variable types represented in the general directionality of causation (Kjaerulff & Madsen, 2008; Pearl, 2009). Parenthetically noted in each node within Figure 14 are exemplar variables that apply to RI causation. Special attention during the model building phase was paid to inclusion of variables that may not be of direct interest, but that serve an important mediating function and whose exclusion would adversely affect the accuracy of model results (Bedford & Cooke, 2001; Kjaerulff & Madsen, 2008). Figure 14 contains an example of this type of variable by including *visibility restriction* as a mediating variable between *rain* and *crew distraction* and the outcome variable, the inadvertent *crossing of the hold short line*. Where data were known for network model nodes, probability distribution functions were fitted to the data and integrated into the model building process (Bedford & Cooke, 2001; Darwiche, 2009; and Luxhøj, 2003). The

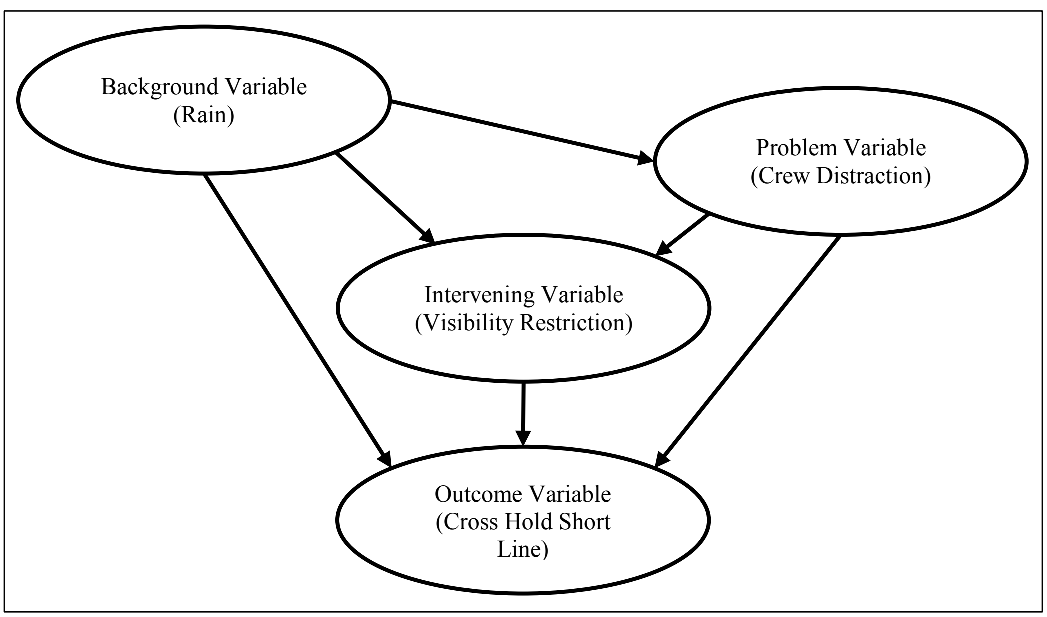


Figure 14. Basic BBN causal interaction.

conditional probabilities that correspond to the random, unknown variables of interest in the network model were elicited as described in phase three; however, if the model structure informing the elicitation protocol is incorrect, the risk that elicited information is inaccurate increases. This being the case, the model structure was finalized through the iterative, expert review-driven process described here prior to elicitation of expert opinion for model quantification (Kjaerulff & Madsen, 2008). This process followed the sequential alpha-level, beta-level, and gamma-level model process described by Marcot, Steventon, Sutherland, and McCann (2006) and the specification of dependence structure described by Hanea, Kurowicka, and Cooke (2006). As suggested by Marcot et al., the model was sequentially revised according to SME comments generated through formal review based on the principles of the Delphi Method (Landeta, 2006). Like the Delphi method process, SMEs confidentially reviewed the model structure and answered a set of structured questions. They were provided feedback based on the combined responses of other SMEs, and were given the opportunity to refine their opinions on the structure and content of the model. Diverging somewhat from Kjaerulff and Madsen (2008) and Marcot et al. (2006), model iterations were made successively more compact to allow for more efficient verification and validation, as well as elicitation of probability and rank correlation functions, with a parsimonious variable set and structure. Appendix H gives the protocol utilized during the model review sessions. For the scope and purpose of this study, and to minimize measurement errors and error associated with validation traceability through the dependence structure, the model was condensed to its most basic structure. While still maintaining adequate resolution, parsimony was achieved using further input from SMEs coupled with the previously discussed RSO and ASRS data.

This additional input was again drawn from a structured SME protocol as in the construction of the more granular model.

Verification and validation of the BBN model. Once the BBN structure was identified as in the preceding section, the variables were checked for conditionality and directionality of the edges that connect them (Bedford & Cooke, 2001). In the present study, the model structure as shown in Figure 14 was initially created by the researcher as discussed previously, and the structure as well as the directionality of connecting edges (dependence and independence relationships) was evaluated by domain experts to ensure that posterior probabilities are correctly indicated. Construction of the model was approached as a continuous process, wherein the structure underwent revision throughout review of the literature and the data collection processes outlined in phase one. Structural issues such as unintended directed cycles were evaluated within the UNINET software package, though this was also a function of the progressive review process. In the preceding discussion, moralization of the DAG was briefly discussed in the context of examining conditional independence. In the present case, the inferential engine for analysis of the network model is native to the software package and graphically manipulating the model is unnecessary (Cowell, 1999). The structure of the network model was such that domain-specific nodes in the model could be separated within UNINET, and the networks evaluated separately against available data and SME judgment. This feature allowed for simplified error tracking across the BBN dependence structure. Once the model was refined through SME input, it was tested using simple probability distribution functions, dummy data ranges, and consistent sampling constraints (in this case, 30,000) to examine interacting nodes as well as the model output under controlled, artificial conditions.

The BBN model was created in the UNINET software package, a screen shot of which is given in Figure 15. UNINET has been used in a variety of modeling studies that involve the creation of BBNs for simulation, including aviation safety applications (Ale et al., 2009; Napoles, 2010). Other software packages for BBNs, HUGIN (Handling Uncertainty in General Inference Network) Expert A/S and BayesiaLab, were evaluated and salient features are summarized in Table 6. While Figure 15 shows the basic workspace available in UNINET, Figure 16 is an example of the complexity of a model with a large number of variables as well as fault and event trees inserted as part of the BBN developed within the UNINET platform.

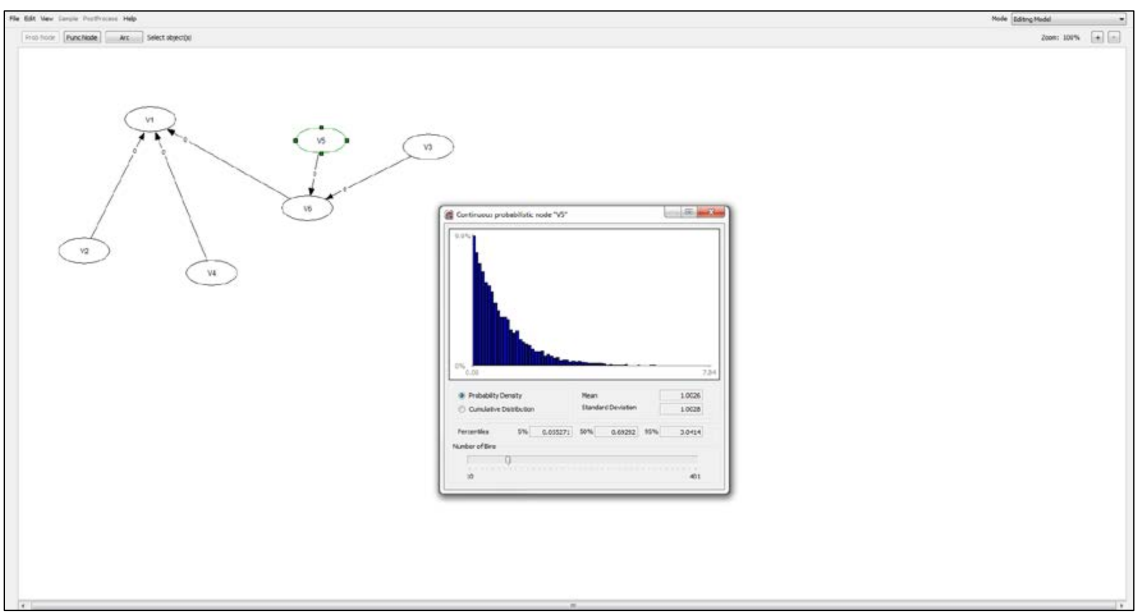


Figure 15. UNINET software for Bayesian belief networks.

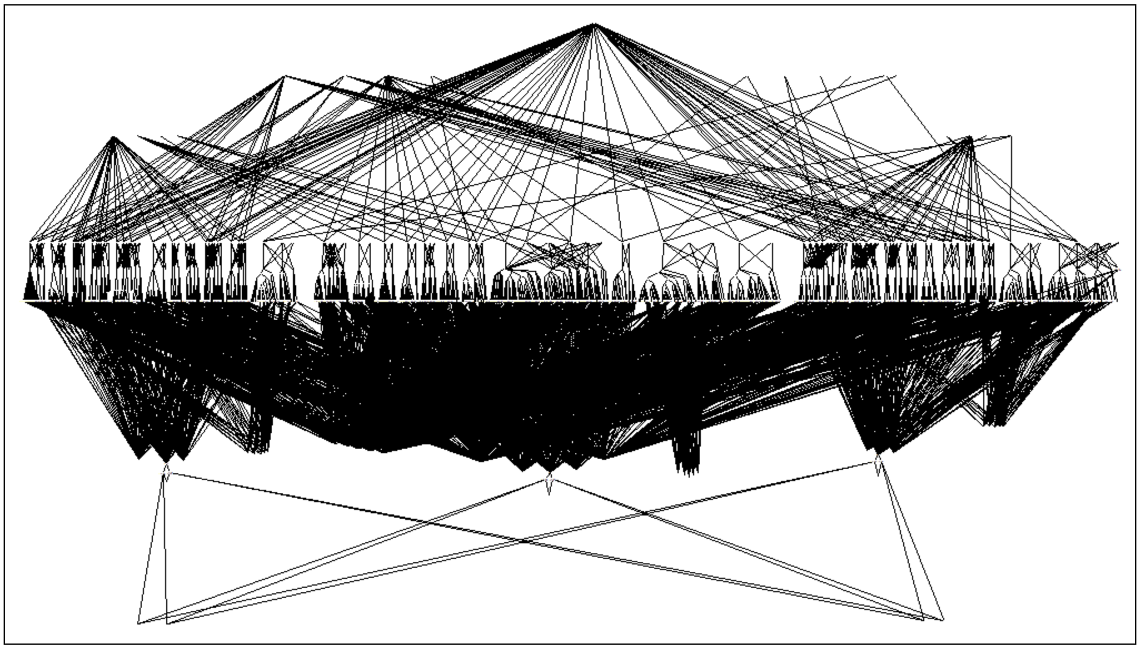


Figure 16. Sample BBN with many nodes and edges.

Table 6. *Software Packages Evaluated for BBN Development.*

Name	Graph Types	Inference	Continuous Nodes
BayesiaLab	Undir/Directed	Joint Tree, Gibbs	Yes
HUGIN Expert A/S	Undir/Directed	Joint Tree	Yes
SamIam	Directed	Recursive Conditioning	Some
UNINET	Undir/Directed	Vine-Copula/Gaussian	Some

The UNINET package used to facilitate this phase of the proposed research utilized the copula-vine approach (Hanea, Kurowicka, & Cooke, 2006; Napoles, 2010) to reduce the computational assessment burden ordinarily associated with large, complex network models and to allow for sampling and analysis through Monte Carlo algorithms, drawing 32,000 samples by default or as specified by the user (Lighttwist Software, n.d.). The copula-vine approach was appropriate in this context because it generalizes the Markov chains often used in high-dimensional problems, and it relies on rank correlation (as discussed in phase three) as a dependence measure of the copula between two variables (Bedford & Cooke, 2002). In addition to simplifying the sampling process, this removes assumptions of normality often associated with discrete BBNs, and more closely follows the distributions found in factors associated with RI causation. The sampling process generates an output from the network model, and though this falls within the broad scope of the phase of research methodology described here, the sampling and analysis of the BBN was not possible until the elicitation process was completed and expert data was processed and populated in UNINET to quantify the model.

Phase 3: Expert Elicitation and Aggregation

Representing the third phase of the research is the selection and elicitation of data from a group of experts. To enable this approach, Cooke's classical model guided the process of expert elicitation (Cooke, 1991). This model structures the elicitation process such that expert judgment is treated as scientific data and is thus formalized in the decision-making process through rational consensus. The Classical model is "essentially a formal method for deriving the requisite weights for a linear pool in which…these weights are expressed as the product of an individual's calibration and information scores" (Aspinall, 2011, p. 3). The aggregation process rewards performance, rather than consensus, and through the application of a strictly proper scoring rule avoids the pitfalls of experts who may intentionally or unintentionally attempt to game the system. This formalization of the elicitation process is advantageous in its transparency, and Appendix C details the mathematics that underlie the algorithm for combining the sometimes dissimilar opinions of experts to derive a rationally-derived *decision maker*.

Data generation and sources. The first step in Cooke's classical algorithm is the selection of experts who, through the structured elicitation process of the classical model, provide probability-based estimates of uncertainty to quantify the BBN. For the purposes of this research and as suggested by Cooke and Goossens (1999) and Renooij (2001), experts were solicited for the study on the basis of the literature review and based on cumulative experience in the field of aviation as well as on operational and theoretical knowledge of the relevant domains that affect RIs. This initial identification of an expert pool was supplemented as outlined in the following process description. Experts were also expected to be conversant in basic probability calculus such that they are comfortable expressing their opinions in terms of probability distributions by way of quantile assessments.

Expert selection. While the meaning of the term expert may be subject to interpretation, in the present context, Wood and Ford's (1993) observations about the differentiation in problem solving between so-called experts and amateurs offered some guidance in this regard. In particular, Wood and Ford's assertion that experts base their opinions less on declarative knowledge than on perceived relationships was important given that elicitation is presented here as an alternative or supplement to scarce data. Cooke and Goossens (1999) suggested that an expert is "a person whose present or past field contains the subject in question, and who is regarded by others as being one of the more knowledgeable about the subject" (pp. 29-30). Expert selection was based on the following general criteria suggested in the EU *Procedures Guide for Structured Expert Judgment*:

- reputation in the field of interest,
- experimental experience in the field of interest,
- number and quality of publications in the field of interest,

- diversity in background,
- awards received,
- balance of views, and
- interest in and availability for the project (Cooke & Goossens, 1999, p. 30).

Though it is mentioned here, the nature of the expertise required for this study did not lend itself to experts having a number of publications or experimental experience, and thus these criteria were approached as guidelines, but not requirements. Once potential experts were identified they were contacted by electronic mail and phone to gauge interest in participation in the elicitation process. Experts for consideration in this research consisted of both general practitioners of aviation and aviation safety and specialists in RI causation and mitigation. During the recruitment process, expert candidates were interviewed with respect to their own area of expertise as well as for recommendations of other potential experts (James et al., 2010; O'Hagan et al., 2006). Where it was not possible to achieve complete impartiality among experts, all practical efforts were made to clarify any potential conflicts of interest. Cooke and Goossens (1999) suggest that a minimum of four experts be chosen for any subject area, and that a good rule of thumb for expert panel size is at least eight members with a representative diversity among participants. Three domains were identified previously as primary to the subject of RI research: infrastructure/organizational factors, human factors, and technological or engineering factors. Consistent with Cooke and Goossens (1999) recommendations, at least two experts with specific knowledge in each domain were engaged for the expert judgment process. That three general domains emerged in the literature review is misleading to an extent, at least as it effects selection of SMEs.

Appendix D provides general biographical information for the experts used in this phase. Review of this background information reveals that the panel of experts had extensive and varied experience from the flight crew, aviation safety, airport operations, and air traffic perspective. Though Cooke and Goossens (1999) suggest that a larger number of experts is desirable, O'Hagan and Oakley (2010) caution against groups larger than five because of the unnecessarily lengthy discussions that may result. In any event, the minimum size of the expert panel identified herein was bound by the need for appropriate diversity across relevant domains, and was the subject of evaluation by the researcher in conjunction with other experts.

Once a pool of potential experts was identified, each was informed of the general processes, procedures, and expected outcomes of the study, and a curriculum vitae (CV) for each SME was obtained and retained by the expert as confidential, but available in de-identified format for review as necessary. Expert CVs were reviewed based on the aforementioned criteria to determine a final list of potential panel members. Because human subjects were involved in the elicitation process, Institutional Review Board (IRB) approval was sought concurrently, and each potential expert was advised of the conditions and risks of participation, including:

- the subject areas for elicitation,
- compensation structure,
- confidentiality,
- intended distribution of study results, and
- feedback of elicitation results.

IRB board approval was obtained in accordance with the most recently adopted guidelines for Embry-Riddle Aeronautical University and in accordance with the IRB Human Subject Protocol Application Form included as Appendix A to this dissertation. Appendix B is an unsigned copy of the informed consent form used in the recruitment of raters and experts.

Upon agreement to participate, each expert was asked to execute a consent form in accordance with IRB guidelines, and executed forms were retained by the researcher. Related to the issue of confidentiality is the use of experts' names within the study results and associated material. Protection of experts' reputations is a legitimate concern, and the use of names and affiliations in the present research must be carefully weighed as a balance between protection of identity and transparency of the elicitation process. To this end, Cooke and Goossens (1999) propose the following in the EU *Procedures Guide for Structured Expert Judgment*, which were, with the exception of publication of names and affiliations due to privacy concerns, adopted in this study and communicated to SMEs:

- Expert names and affiliations are published in the study.
- All information, including expert names and assessments, is available for competent peer review, but is not available for unrestricted distribution.
- Individual assessments are available for unrestricted distribution, assessments are not associated with names but identified as "expert A, B, C,…" etc.
- Expert rationales are available for unrestricted distribution.
- Each individual expert receives feedback on his/her own performance assessment.

- Any further published use of the expert's name requires the expert's approval (p. 31).

The EXCALIBUR software program, which was developed in conjunction with and with support from the European Union, was used to facilitate the classical model of elicitation and aggregation (Cooke & Probst, 2006; Ouchi, 2004). EXCALIBUR processes parametric and quantile uncertainty estimates and calculates expert weights via the classical elicitation technique. The software has been used extensively for high-level elicitation studies, including such critical risk assessment exercises such as the evaluation of eruption risk at Mount Vesuvius, a volcano with the potential to impact millions of people (Neri et al., 2008). Tyshenko et al. (2011) also successfully used the EXCALIBUR package in their investigation of the risk of iatrogenic prion transmission, as did Dawotola, van Gelder, and Vrijling (2011) with respect to risk assessment for crude oil pipelines. Other elicitation packages exist and were evaluated, including SHELF (O'Hagan & Oakley, 2010), ELICITOR (Kynn, 2006), and Elicitator (James, Low Choy, & Mengersen, 2010); however, EXCALIBUR provides the most direct interface with the classical elicitation method and has seen long-term use in large, complex, and varied risk assessment studies (Goossens et al., 2008).

Structured expert judgment. Next, each expert was elicited independently to express their knowledge and degree of uncertainty regarding potential observations. It is this process, structured as discussed in the preceding section, which allows a joint probability distribution to be formulated on the basis of a person's knowledge and beliefs. After selection, experts were trained on basic probability calculus and logic so that they were able to express opinions relative to the RI model via probabilistic estimation.

Training familiarized experts with the expected format of elicited responses. Pre-elicitation training of experts also included expert completion of estimation training questions, in which panel members were asked to provide probability estimates in a variety of formats to prepare for the formal elicitation process. This training was crucial to the effectiveness and reliability of the elicitation process, and was conducted as a formal meeting between the researcher and selected experts either in person or through the combined use of telephone and computer-based video conference (DeWispelare et al., 1995; Mosleh et al., 1988).

Following training in the basic process of elicitation and in expressing uncertainty in the form of probabilistic assessments, a facilitator, in this case the author, ascertained explicit distributions for the selected elements of the model based on answers to specific questions in the elicitation protocol attached here as Appendix E. In the classical model, questions are based upon target variables, query variables, and seed variables. In this study, target variables and query variables generally coincided, and they represented the variables whose values were elicited for inclusion in the network model. In developing these variables for elicitation, dependencies were evaluated and prepared for further assessment by domain experts so that probabilistic dependence between variables could be identified and accounted for. Seed variables were those for which values are known to the researcher but not the expert, and that were used in the calibration and aggregation of expert opinion (Cooke, 1991), and are included in the elicitation protocol in Appendix E.

Bias is a topic of considerable weight in the expert elicitation literature, beginning largely with the seminal work of (Tversky & Kahneman, 1974), which investigated the implications of common heuristics employed in the judgment of probability and the biases to which they lead. Bias is certainly a topic of concern, but (Kynn, 2008) argues that much of the statistical research on elicitation since Tversky and Kahneman published their research has failed to address new knowledge in psychology that has refuted many of the original heuristic-based concerns. This lapse in collaboration between the statistical and psychological communities creates a form of bias in itself wherein elicitation researchers may not adequately address the importance of framing elicitation questions such that bias is systematically avoided (Clemen & Lichtendahl, 2002; Garthwaite et al., 2005). After providing quantile estimates for each question, experts were allowed to evaluate the implied frequencies and distributions of their responses. No attempt was made to encourage experts to change their estimates, but in some cases, experts did alter their responses after they were presented with a distribution or after reviewing frequency of events on various scales. Aside from providing process feedback during elicitation to reduce bias, elicitation questions were framed to:

- encourage rule-based approaches,
- focus on an expert’s specific domain,
- incorporate assigned confidence,
- avoid extreme probabilities,
- require repetitive sampling of knowledge, and
- allow for deliberate practice (Kynn, 2008; Martin et al., 2012; O'Hagan et al., 2006; Renooij, 2001; Renooij and Witteman, 1999; Speirs-Bridge et al.,2010)

Since probability distributions are the desired outcome of the elicitation process, and direct elicitation of rare event probabilities has been shown to increase bias through overconfidence (Tversky & Kahneman, 1974; Kynn, 2008), questioning focused instead on interval elicitation. These intervals were elicited as a four-step procedure as outlined by Speirs-Bridge et al. (2010), the defining feature of which is that experts were asked to estimate their level of confidence that the interval captures the true value along with assessing the lower bound, upper bound, and most likely values. This questioning format was merged with the basic procedures suggested by Cooke (1991) and Cooke & Goossens (1999), to maintain the integrity of the classical model.

Unique to the study here was the method by which rank correlation and dependence information was elicited from SMEs. Ordinarily, the Classical model specifies that rank and dependence information is gathered through questions that closely follow the probabilistic calculus used to interpret it. While this has been demonstrated in the literature (Cooke & Goossens, 2006; Morales, Kurowicka, & Roelen 2008), it is also the case that experts in more operationally-focused research are uncomfortable answering such questions or find the process cumbersome, potentially affecting the quality of results. In this study, experts were first asked to address questions that conformed to the typical, Classical format, as in Morales, Kurowicka, and Roelen (2008). These questions took the general form of: *What is the probability that w is above its* q_w*th quantile given that x is above its* q_x*th quantile, y is above its* q_y*th quantile and z is above its* q_z*th quantile?* Their responses suggested that a more intuitive elicitation format might improve expert understanding of the dependence structure under question, and the protocol was modified to ask experts questions based to some extent on the format suggested by Roelen, van Baren, Morales, and Krugla (2008) in their development of a model for aviation maintenance behavior. Merging SME feedback with ideas from the Roelen et al study, a new questioning format was developed and tested for elicitation of rank correlation information. The validity of the questioning protocol was assured through the literature review, by consultation with researchers in the field (Dr. R. Cooke, Dr. M. Wittmann, personal communication, 2013), and through comparison with the body of research in structured expert judgment (see Goossens, et al., 2008). These questions asked experts to rank variables by assigning an order to each variable indicating influence in descending order. The most influential variable was then inserted into a rank correlation question typical of the Classical model. Experts were then asked to quantify the influence of the remaining variables as a percentage of the influence of the highest weighted variable. Rank correlation values were then calculated based on these responses using UNINET. Appendix E details the questions presented to SMEs.

On the basis that the overarching purpose of the elicitation process is to obtain a prior distribution, distributions were fitted to the elicited intervals and available for feedback to experts (Garthwaite et al., 2005). The EXCALIBUR software package was used for this function along with MS Excel in real-time during the elicitation as well as for post-elicitation analyses. The four-step process of obtaining intervals minimized the possibility of overly-simplistic intervals by eliminating the untenable belief that the parameter of interest lies at or close to the limits as may be suggested by only specifying a range. This process of over-fitting allowed more accurate evaluation of elicitation data once the collection and fitting of opinions was complete. Once elicitation was accomplished and fit to a minimally informative distribution, the adequacy of the elicited data was evaluated as an extension of the feedback process, by highlighting and reframing to experts the implications of the elicited values and confirming for each elicited value that these are satisfactory representations of the experts' beliefs. (O'Hagan

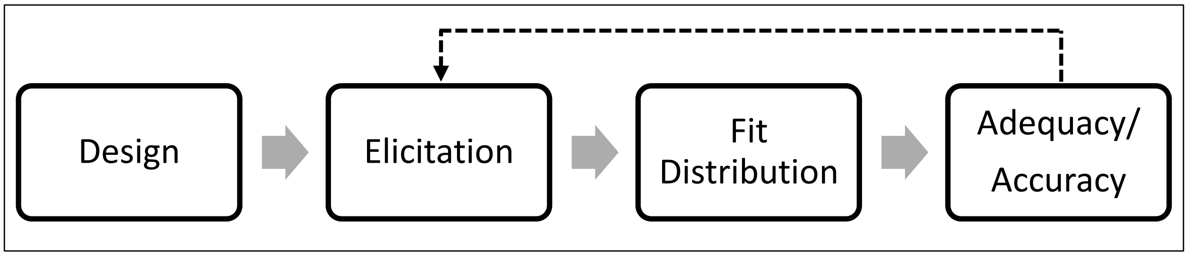


Figure 17. Expert elicitation process overview.

et al., 2006). Figure 17 illustrates the recursive process of elicitation as described here.

Assessment of the adequacy of elicited data has been mentioned briefly in previous discussions of feedback and over-fitting, both of which serve as tools for evaluation of expert- derived data. Feedback was presented to expert panelists to confirm that the assumed distribution form is a reasonable representation of each expert's ideas (Garthwaite, Kadane, & O'Hagan, 2005; Kynn, 2008). When distributions were inaccurate representations of an expert's beliefs, he or she was given the opportunity to revise earlier estimates. Humans are generally poor judges of probability by distribution alone (O'Hagan et al., 2006), so while feedback was utilized, the concept of over-fitting was used to expose inconsistent answers during the elicitation process. The four-step elicitation procedure discussed previously allowed this opportunity by asking for more information than is necessary to establish a parametric probability distribution. This not only reduced expert overconfidence, it permitted testing for coherence and refinement of the best-fit distribution to the elicited data (Garthwaite et al., 2005; Speirs-Bridge et al., 2010).

Using the basic format described by Figure 17, reliability and validity were continuously assessed. From the perspective of face and construct validity, iterative review of the protocol for elicitation with outside experts in structured expert judgment ensured that the content and structure of the questions was sound. The four-step procedure for elicitation and provision of feedback served as a means of assuring reliability, and bias was also controlled in this way as previously discussed. Though in the context of structured expert judgment, it is implausible to quantify these measures, the steps outlined here were carefully constructed to limit threats to reliability and validity, and the elicited data were found to be acceptable for further analysis.

Aggregation of elicited data. The literature review segment of this research and supplemental information in Appendix C covers in detail the elements of Cooke's classical model for elicitation and aggregation of expert opinion. As was described, the experts' data was subjected to the classical model weighting scheme based on the seed variable answers, and the weighted answers from all experts were pooled to provide a rational consensus judgment (Cooke, 1991; Cooke, 2009). The key feature of the Classical model is the performance-based aggregation of experts' uncertainty assessments, rewarding expert performance as opposed to consensus. As a measure of both calibration and information, which are related to the concepts of precision and accuracy, the Classical model prevents gaming the system through strictly proper scoring and rewards only an expert whose assessments perform well, are informative, and are in accordance with his or her true beliefs. Robustness analysis of calibration questions and of expert influence on the decision maker was also performed using the remove one-at-a-time method detailed by Cooke (1991). The elicitation and aggregation processes were managed in the EXCALIBUR (Goossens et al., 2008) software package using the Classical method algorithms discussed in more detail in Appendix C. The results section of this study presents in greater detail the aggregation of expert judgment using decision maker (DM) optimization as well as equal weighting for comparison (Cooke, 1991).

Once expert data were obtained through the structured elicitation process, minimally informative distributions fitted to the data, and aggregation of expert judgments completed, the data were inputted into UNINET for quantification of the model. Where data existed, as in the case of mechanical failure or some weather-related nodes, those data were also fitted to a distribution and entered into UNINET. Once the model was quantified within the software platform, sampling and subsequent analytic conditioning was performed. In the case of UNINET, real-time analytic sampling was conducted on a single value across probabilistic nodes, and sample-based conditioning was also used to conditionalize on specific points or intervals as appropriate (Lighttwist Software, n.d.).

Quantification and Interpretation of the Model

After completion of the first three methodological phases of the study, it was necessary to combine the information collected through the structured expert judgment process with the model structure developed in the second phase and informed by the first. This process was accomplished using the UNINET software package described previously, and expert-elicited uncertainty distributions and rank correlation information were entered into the model created and iteratively validated in phase two of this research. Following elicitation of uncertainty distributions and derivation of the DM for model target variables, UNINET was once again used to input model quantification parameters. Using the uncertainty analysis platform UNICORN (Kurowicka & Cooke, 2006) and Oracle Crystal Ball to analyze the output from EXCALIBUR following the structured expert judgment sessions allowed distributions shapes and parameters to be evaluated and verified for entry in the BBN model. This allowed confirmation of minimally informative distribution properties prior to entry in UNINET, and independent samples from the distributions were compared between platforms to ensure accuracy in data entry. Figure 18 shows a partial view of the distribution specification function within the UNINET software package. Each variable was quantified using a similar process within the program, with expected differences based on the type of distribution and whether the variable was to be quantified by existing data or expert-elicited data.

Rank correlation scores were calculated separately and entered into UNINET to support the model dependence structure. Figure 19 illustrates an example screen showing the entry of rank and correlation information into the model. Readers should note that the probabilities elicited from SMEs in the preceding phase differ from rank correlation, which is calculated separately. Recalling that experts were asked to characterize

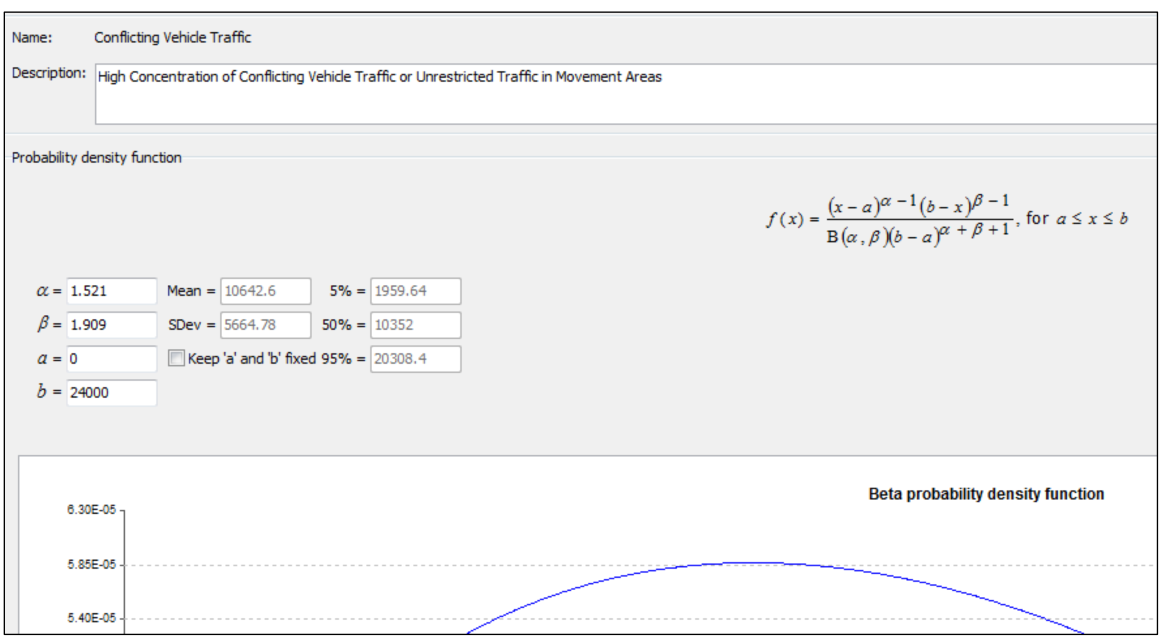


Figure 18. Partial screen capture of UNINET variable distribution entry.

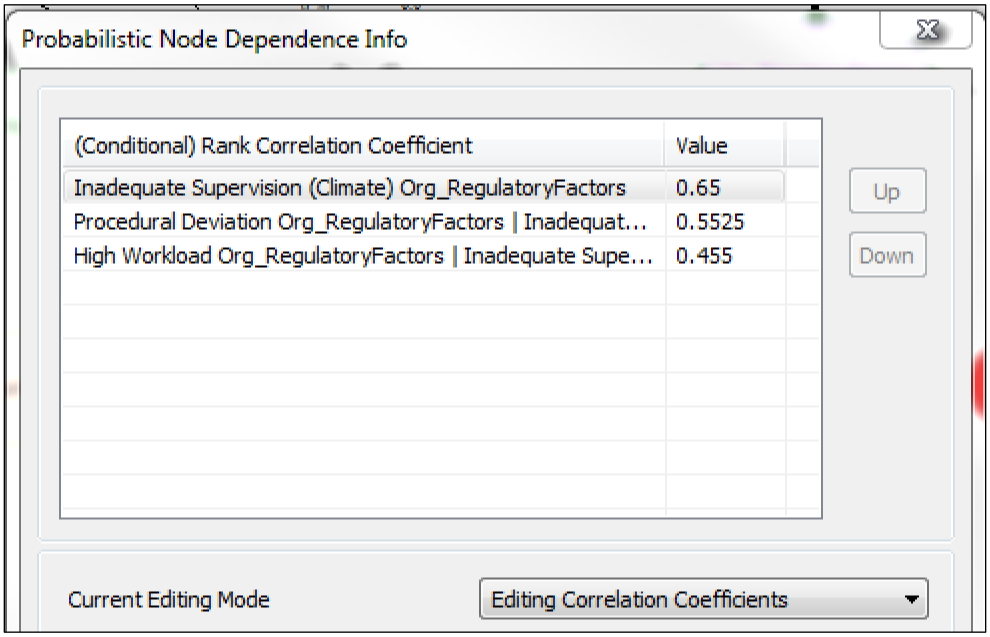


Figure 19. Entry window for rank and correlation information.

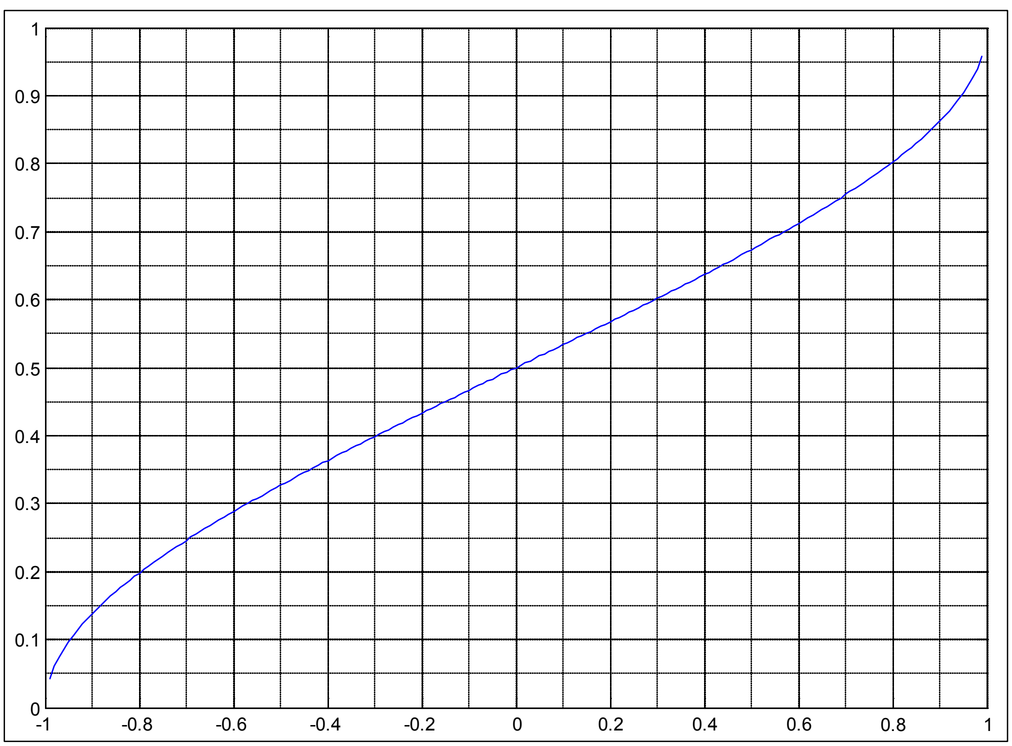


Figure 20. Probability of exceedance (y-axis) versus rank correlation (x-axis).

dependent relationships in the form of probabilities of exceedance, rank correlations were required to be calculated based on the relationship with the normal copula as shown in Figure 20. Because the conditional probability is transformed to rank correlation such that for any q in the exceedance probability equation (Equation 4), the probability is reflected as $P_1 = q$-1, any value other than 0.5 for q limits P and makes the choice of copula more impactful on conditional probability (Morales, Kurowicka, & Roelen, 2008). Because elicited probabilities were in the interval (0,1), and only values above 0.5 (given the choice of 0.5 for q) reflect as positively correlated, the absolute values given by SMEs were converted to the scale imposed by the rank correlation transformation rather than burdening experts with additional restrictions on their responses. In this process, intentional negative correlation values were converted to the interval (0, 0.5) and positive values were transformed to the interval (0.5, 1). This was confirmed with the SMEs to maintain integrity of the probability assignment task and the relative values intended by the experts. These values were then entered into UNINET as discussed here.

Figure 20 addresses only the issue of rank correlation when it is not conditioned by additional parents. In the case of multiple parent nodes as were present in the model in this study, conditional rank correlation must also be calculated from the exceedance probabilities elicited from experts. In the present case, the highest rank correlation identified in the elicitation was formulated as an exceedance probability question as a means of determining:

$$R_1^{e_i} = P\{X_5 \geq x_{5,q50}^{e_i} | X_1 \geq x_{1,q50}^{e_i}\} \tag{4}$$

The relationship between Equation 4 and rank correlation as plotted in Figure 20 was computed by integrating the bivariate normal density over the quantile's exceedance region, using Equation 5 to first determine product moment correlation, and then transforming product moment correlation to rank correlation via Equation 6 (see also Cooke, 1991; Cooke and Goossens, 1999; Kurowicka and Cooke, 2006; Morales, Kurowicka, and Roelen, 2008; and Morales, Cooke, and Kurowicka, 2008):

$$\frac{1}{1-q}\int_{\Phi^{-1}(q)}^{\infty}\int_{\Phi^{-1}(q)}^{\infty}\phi(x_1, x_2, \rho_{1,2})dx_1dx_2 \qquad (5)$$

where:

Φ^{-1} is the inverse normal cumulative density function;

$\phi(x_1x_2, \rho_{1,2})$ is the bivariate normal density;

ρ is the product moment correlation; and

q is the selected quantile for exceedance.

$$\rho = 2\sin\left(\frac{\pi}{6}r\right) \qquad (6)$$

where:

r is the rank correlation.

The remaining rank correlations for each ratio exceedance question set in the elicitation session were calculated as a function of their ratio with the next assessment in the rank hierarchy, as in:

$$R_2^{e_i} = \frac{r_{2,5}^{e_i}}{r_{1,5}^{e_i}} \qquad (7)$$

where:

$R_2^{e_i}$ is the elicited exceedance probability; and

$r_{1,5}^{e_i}$ is the rank correlation calculated in Equations 4 through 6.

which was bounded by the next higher ranked estimate.

The primary goal of the three fundamental phases of this research was to create and analyze a quantified Bayesian belief network model representative of the dynamic interaction of causal factors that lead to RIs. Essentially then, the desired outcome was a model that *supports* sensitivity analysis, including that of the model parameters, though such analysis was not completed in substantial depth here. In the scope of the present research, the desire for sensitivity analysis is purposely constrained to identification of those causal interactions, or paths, and factors that contribute most to the undesired outcome, in this case an RI event. The software UNIGRAPH and UNISENS (both developed in conjunction with UNINET and cooperatively with the Technical University of Delft) supports evaluation of the BBN simulation via graphical analysis using cobweb plots and through sensitivity measures including product moment correlation, rank correlation, regression coefficient, correlation ratio, and partial regression coefficient in conjunction with scatterplots (Lighttwist Software, n.d.). UNINET also allows conditioning of models on specific values or intervals, and the model was conditioned to evaluate causation in the ordinary sense, but also from the perspective of reverse propagation, the results of which are presented in the following chapter. Using these tools, the mechanics of the BBN model were evaluated to ensure that the model adequately supported further research, which may include evaluation of mitigation strategies or updating based on technological or infrastructural changes.

Methodological Validation

Several methods were used in an iterative process to develop the model described here, and as such, validation followed a similarly iterative progression. In cases where data were available, the population was used, and establishing the validity of the data was less concerning. However, where probabilistic data were used, it was generally infeasible to separate the data into sets to allow testing and evaluation. Instead, each phase of the modeling approach relied on several sources, including expert review and opinion, to triangulate the structure of the model and the elicitation of data. As detailed in the respective phases of the methodology, accepted mathematical algorithms exist for evaluating the structure and quality of the BBN as well as the expert opinion. These performance measures were carefully considered for each phase, and modeling processes and rationale that extended beyond a single phase were continuously evaluated for fitness of purpose and methodological transparency in cooperation with the expert panels used throughout. Where within-phase processes could not be verified using the techniques described, the phased approach was discontinued until such concerns were resolved and documented. The results presented in the following section present these processes in more detail for each phase of the study.

CHAPTER IV

RESULTS

This study addressed the understanding of causal factors and interactions leading to RI events in the U.S. In particular, attention was given to the lack of detailed data on many identified causal elements and relationships. A BBN model was formulated to combine sparse available data with data obtained through structured expert judgment and allow updating as new knowledge becomes available. This research was conducted through a mixed-method process of data acquisition, an iterative model building method, and a structured approach to model quantification through expert elicitation. Based on the phased methodological approach to this study and the differing outcomes of interest, results are presented separately for each phase.

Phase 1: Runway Incursion Data and Causal Factors

Phase one of this study sought to triangulate data from multiple sources with information gained through systematic literature review. As discussed previously, RI event data were collected from ASRS and RSO sources to supplement findings from the literature review. Table 7 summarizes the RSO data used by type of operation, including breakdown by Federal Aviation Regulation (FAR) Part. RSO data included 6,185 records for RI events after basic data screening operations. By contrast, using a search string structured as shown in Table 8 yielded 81 ASRS valid records for SME review. Seventy-one ASRS cases were retained after removal of reports where two or more raters agreed the case was an exception and should be removed from further review. ASRS reports were categorized according to type of operation as shown in Table 9.

Table 7. *RSO Data Summary of Operation Type.*

Type Operation	Count	Percentage
91	2918	47.2
N/A	1592	25.7
121	1095	17.7
135	232	3.8
129	106	1.7
Vehicle	97	1.6
Military	79	1.3
Maintenance	41	.7
Pedestrian	18	.3
125	6	.1
Total	6185	100.0

Table 8. *ASRS Search String Criteria*.

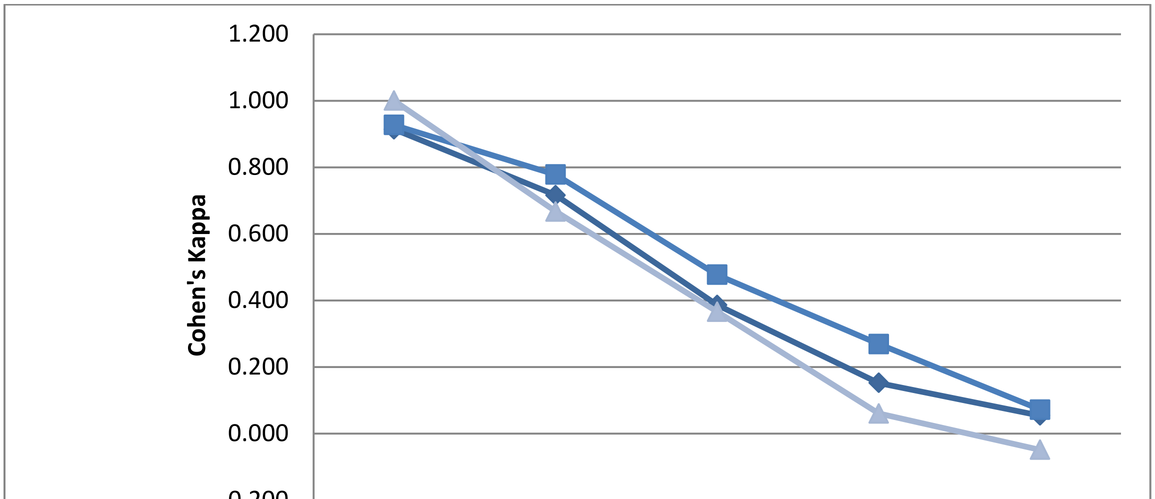
Search Field	Search Criteria
Date/Report Number:	January 1, 2008 through December 31, 2012
Event Type:	Ground Incursion: Runway *or* Taxiway
Reporter Function:	Flight Crew: Captain *or* Check Pilot *or* First Officer *or* Flight Engineer / Second Officer *or* Instructor *or* Other / Unknown or Pilot Flying *or* Pilot Not Flying *or* Relief Pilot *or* Single Pilot *or* Trainee
Text Contains:	"incursion" in ASRS narrative

Table 9. *ASRS Reports Summary of Operation Type*

Type Operation	Count	Percentage
121	32	39.5%
91	30	37.0%
Other	8	9.9%
91K	5	6.2%
Military	3	3.7%
135	3	3.7%
Total	81	100.0%

Over the study period, 2,918 Part 91 aircraft were involved in reported RI incidents as opposed to 1,095 and 232 aircraft operating under Parts 121 and 135, respectively. However, ASRS records indicated that these operation types were reversed in their frequency of RI involvement, with 43.2 percent of cases involving Part 121 or 135 aircraft versus 37 percent attributed to Part 91 operators. In part, the literature and SME input indicated that there is a higher incidence of self-reporting via ASRS reports among commercial operators than for those involved in Part 91 operations. This is confirmed through searching the entire ASRS database, which shows reports by Part 121 and 135 operators outnumber those by Part 91 operators by more than two to one. This apparent underreporting by noncommercial operators was not unexpected, and was one reason multiple streams of data were evaluated.

One such stream of data came in the form of SME identification of causal factors in ASRS-reported RI events. SMEs were tasked with review of 71 ASRS cases after data cleaning, and agreement between raters was evaluated by a modified application of Cohen's kappa as well as through union and intersection operations. Interrater reliability remained higher than the widely accepted (Leech, Barrett, & Morgan, 2008) minimum of 0.70 (shown in Figure 21) for matches of one and two ICAO causal codes; however, kappa decreased rapidly for cases of three, four, or five matching codes. This decreased convergence between rater pairs was likely a function of the possible combinations available to each SME for each ASRS case: up to 1,533,939 assuming five causal factors were used in a case. Reliability measures were supplemented with intersection and union operations to further understand rater responses. If all raters had used different, unique codes across a single case, the union should have reflected the maximum number of 15.



	1	2	3	4	5
Rater 1 to Rater 2	0.914	0.716	0.386	0.152	0.054
Rater 1 to Rater 3	0.927	0.778	0.477	0.269	0.072
Rater 2 to Rater 3	1.000	0.667	0.366	0.060	-0.049

Figure 21. Interrater agreement (*kappa*).

Instead, union most frequently ranged from three to five unique codes for a given case, which indicated some level of consolidation of causation, even if not in perfect agreement across raters as shown in Figures 22 and 23.

Among the raters, certain codes were used more than others, and in some cases were not used at all. Whereas the code describing a failure to obtain Automatic Terminal Information System (ATIS) details was used infrequently, crew failure to adhere to hold short instructions from ATC was the most commonly appearing code. Table 10 lists causal codes and gives count and percentage figures for each. A complete list of available codes is provided in Appendix F.

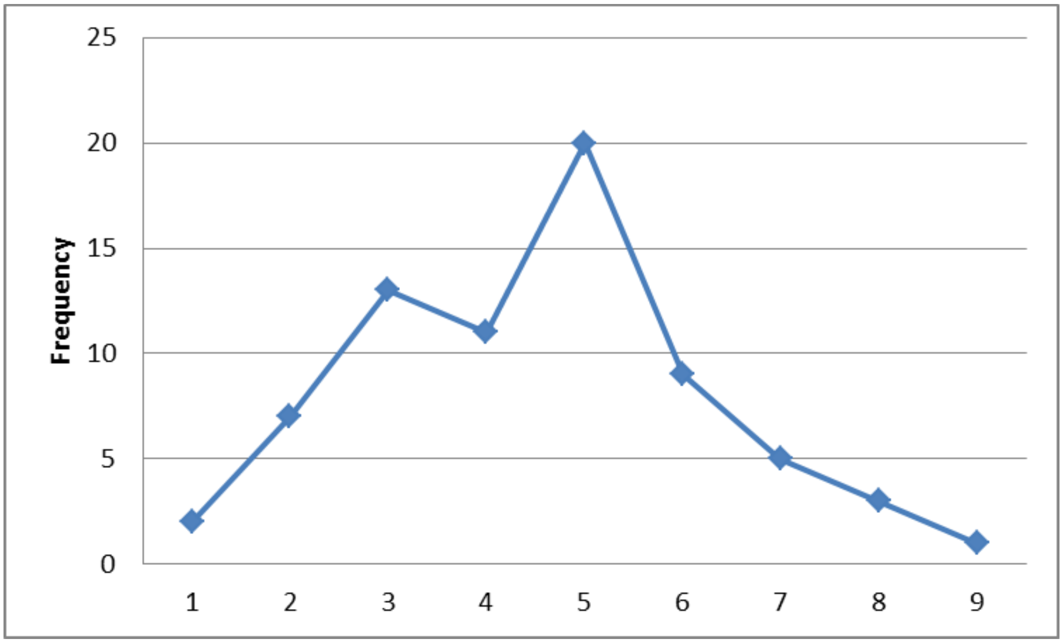


Figure 22. Union count by case.

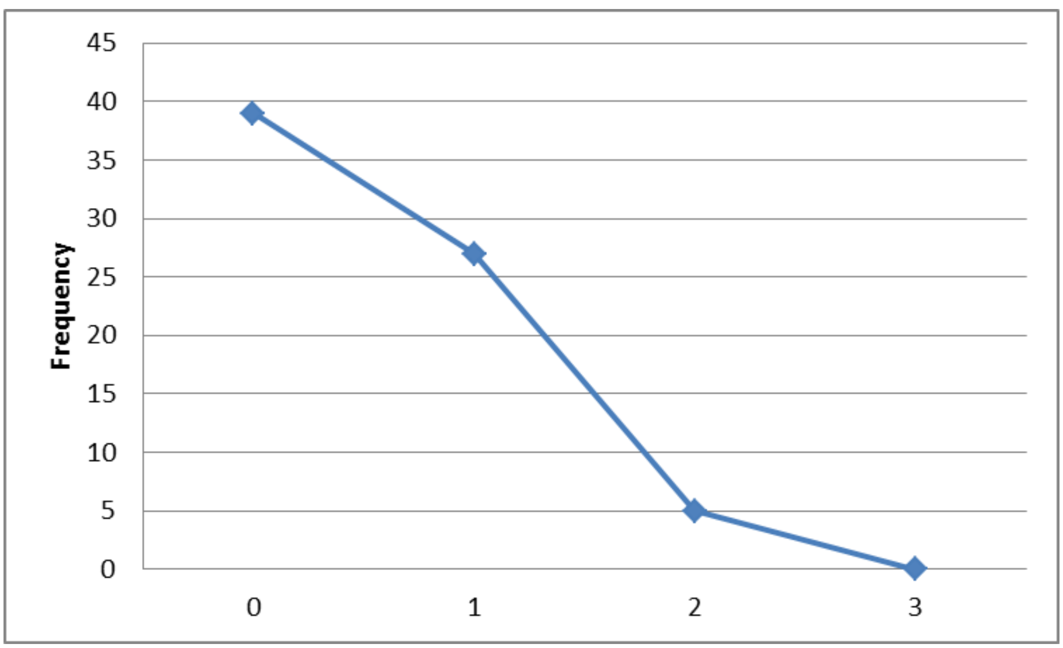


Figure 23. Intersected code frequency.

Table 10. *Use Count by Causal Code.*

Code	Count	Cum. Count	%	Cum. %
2.2.15	49	49	10.04	10.04
2.4.5	45	94	9.22	19.26
2.2.5	38	132	7.79	27.05
2.2.10	34	166	6.97	34.02
2.4.1	33	199	6.76	40.78
2.3.4	30	229	6.15	46.93
2.2.1	22	251	4.51	51.43
2.2.6	20	271	4.10	55.53
2.4.11	18	289	3.69	59.22
2.5.1	18	307	3.69	62.91
2.1.11	17	324	3.48	66.39
2.3.7	15	339	3.07	69.47
2.4.2	14	353	2.87	72.34
2.2.2	13	366	2.66	75.00
2.3.6	13	379	2.66	77.66
2.1.9	11	390	2.25	79.92
2.2.13	9	399	1.84	81.76
2.2.7	8	407	1.64	83.40
2.1.3	7	414	1.43	84.84
2.3.3	7	421	1.43	86.27
2.2.9	6	427	1.23	87.50
2.2.14	6	433	1.23	88.73
2.3.2	6	439	1.23	89.96
2.3.8	6	445	1.23	91.19
2.4.6	6	451	1.23	92.42
2.2.4	5	456	1.02	93.44
2.4.4	5	461	1.02	94.47
2.4.9	5	466	1.02	95.49
2.1.4	4	470	0.82	96.31
2.1.7	4	474	0.82	97.13
2.3.1	4	478	0.82	97.95
2.2.3	2	480	0.41	98.36
2.2.11	2	482	0.41	98.77
2.4.3	2	484	0.41	99.18
2.4.10	2	486	0.41	99.59
2.1.2	1	487	0.20	99.80
2.3.5	1	488	0.20	100.00

In addition to assigning ICAO causal codes to ASRS reports, SMEs provided comments on a many reports. These comments also added depth of understanding to cases where raters agreed or disagreed, and as shown in ASRS report 939675, where raters agreed on assignment of the 2.4.5 causal code, other themes were also apparent:

> *We were cleared to taxi via Taxiway D1, D, and A3, hold short of Runway 3/21 by Ground Control and told to contact Tower upon reaching the hold markings on Taxiway A3 at Runway 3/21. Upon contacting Tower we were instructed to hold short of Runway 11L. Upon receiving this instruction I assumed there was a separate hold line for Runway 11L and began taxiing across the hold line for Runway 3/21 at Taxiway A3. When crossing this hold line I realized that there was not a separate hold line for Runway 11L. I promptly turned right onto Runway 3/21 to leave the extended centerline for Runway 11L. An aircraft on final was asked to go-around. We had the airport diagram out as we were taxiing, but it was unclear as to whether there was a hold line for Runway 11L at Taxiway A3. Construction was occurring on Taxiway A, closing many of the exit taxiways off of Runway 11L. Aircraft were having to back taxi on the active runway to exit the runway at Taxiway D and Runway 11L thereby causing confusion and crowded conditions at the intersections around where Runways 11L/29R and 3/21 intersected.*

This particular report not only represented an example of agreement across experts on a causal code, it also highlighted the complex causation that characterized RI events by identifying lost situational awareness, airport unfamiliarity, construction, back-taxi, and intersection complexity as contributing factors. Beyond the emic perspective of causation, reports also offered an insider perspective on how flight crews identified avoidance procedures for the future, as in ASRS report 974660:

> *The corrective actions I need to take: Minimize distractions while in critical phase of flight/taxi-only monitor one frequency at a time. If any doubt exists, bring aircraft to an immediate stop and clarify instruction. Ensure both crew members have their heads up looking outside during all critical phases*.

SMEs also used the 2.2.5 causal code for unfamiliar airport layout with frequency, as in ASRS report 969670, where all SMEs agreed on the code's use:

> *Upon exiting Runway 15 at DCA, we were instructed by Tower to hold short of Runway 19 at Taxiway M. The First Officer read back the instructions which I then repeated to him. The hold short line for 15 and 19 are extremely close together (there is insufficient room for larger aircraft to exit 15 fully and still remain short of 19 on Taxiway M). By the time I had acknowledged the hold short instructions we had cleared the first hold short line (exiting Runway 15), and about to cross the hold short line for Runway 19. I had confused this hold short line for the 15 hold short line, and passed the Runway 19 hold short line before realizing my error. ATC repeated the hold short instructions. Once clear of conflicting traffic, ATC cleared us to cross Runway 19 and taxi to the gate. My limited familiarity with DCA airport and high workload caused the momentary confusion which led to the incursion. The approach and runway exiting plan was briefed thoroughly during the approach briefing. We had the airport diagram to refer to and knew our taxi route, which included the 'Hot Spots' (one happens to be at taxiway M on the other side of Runway 19 (the terminal side)). We still*

nevertheless managed to miss the hold short line, meaning that additional vigilance was needed. I would suggest A 'Hot Spot' or a note on the 10-9 airport diagram at Taxiway M between Runway 15/33 and 1/19 could provide an additional safety measure. Additionally, it should be noted that there is insufficient room for larger aircraft to exit 15 fully and still remain short of 19 on Taxiway M.

As in the previous examples, the emic view of the event addressed important interactions leading to the RI, and reviewing cases with universal SME agreement or comments allowed for a more complete understanding of RSO data and findings of previous research and a more intuitive approach to development of the model.

The frequency of code use was interesting in that it suggested certain combinations of codes that grouped together thematically. Unused codes were also telling in this way. Figure 24 illustrates the relative proportion with which codes were assigned by raters. Figure 25 graphically presents some of the most commonly emerging themes based on causal code assignment and rater comments, with word size within the figure indicating the relative frequency with which certain words or phrases were identified by the SME panel. The literature initially suggested that human factors, organizational, and technological domains existed to describe RI causation. These domains were confirmed by the emerging themes from ASRS review, and they were expanded to include weather and the operational environment.

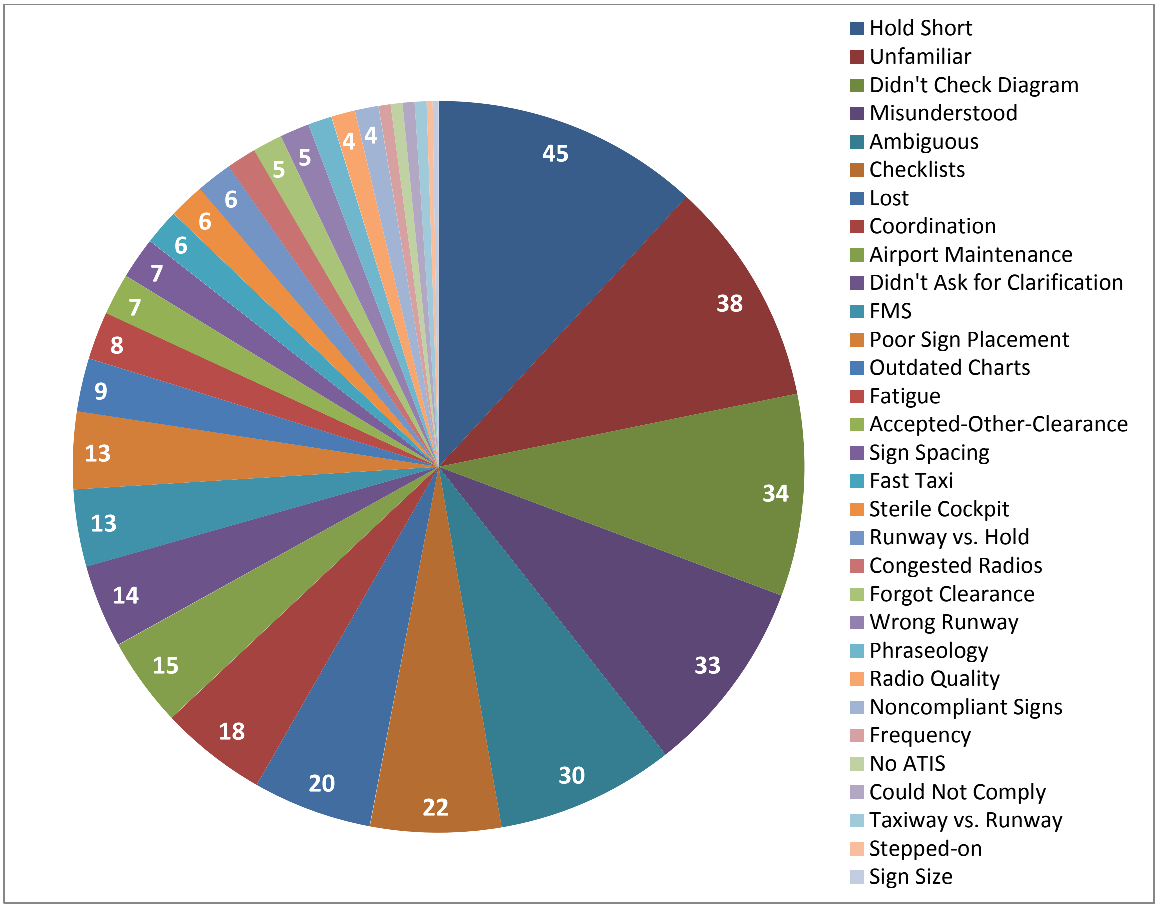


Figure 24. Proportional ASRS causal code assignment by SMEs.

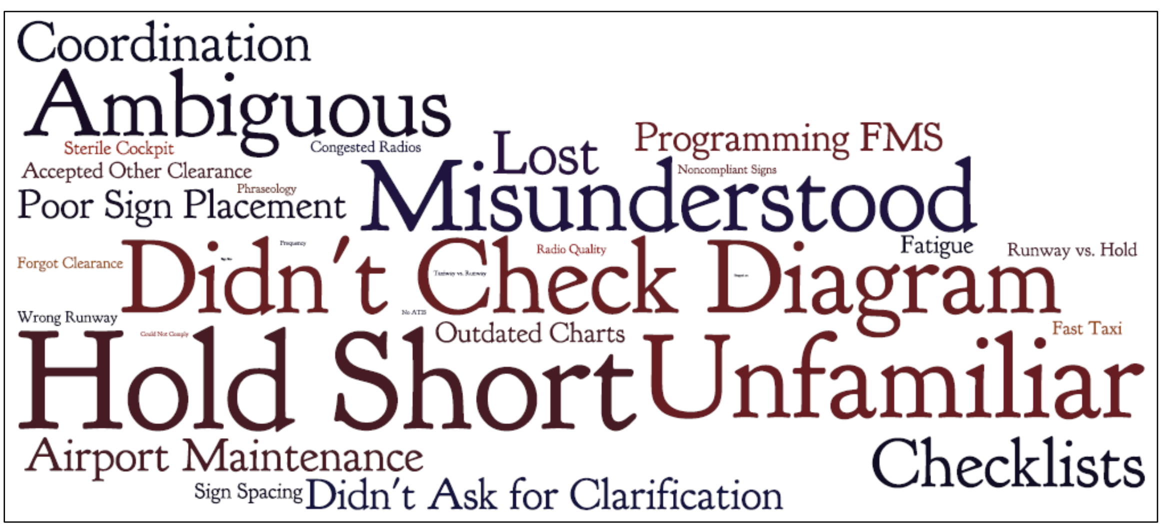


Figure 25. Word cloud of SME causal codes and comments.

Phase 2: Belief Network Model

Review of the literature on RI occurrence and prevention provided a list of domains, concepts, and factors relative to causation of incursion-type incidents. Analysis of RSO data alongside the ASRS data and SME-coded, casewise causal factors and comments allowed for a database of causal factors to be created on which to base the influence diagram and network model. Using this set of variables, an initial model structure was constructed from 58 variables within the organizational, operational, human factors, weather, and technological domains, and after iterative review by SMEs, the model was arranged as shown in Figure 26. This SME review was conducted individually, with feedback and a structured series of questions following the general principles of the Delphi method as outlined in the previous chapter. The parsimonious representation of the model, which was derived from expert consensus and supported by the literature and data reviewed in phase one, includes 27 nodes and is shown in Figure 27. This model retained the same basic dependence structure and thematic elements as the complete model, and is compatible with the results from phase one while easing the computational burden of verification and validation of what is demonstrably a new representation of RI causation and dependence. Furthermore, this approach supports the additive philosophy of this study in that verification and validation of the parsimonious model leads naturally to expansion of such efforts into the more granular, detailed model. Aggregating nodes within each domain were connected by converging directed edges to a central, probabilistic node that addresses the requirement of the RI definition for the "incorrect presence of an aircraft, vehicle or person" (EASA, 2011, p. v; ICAO, 2007, p.

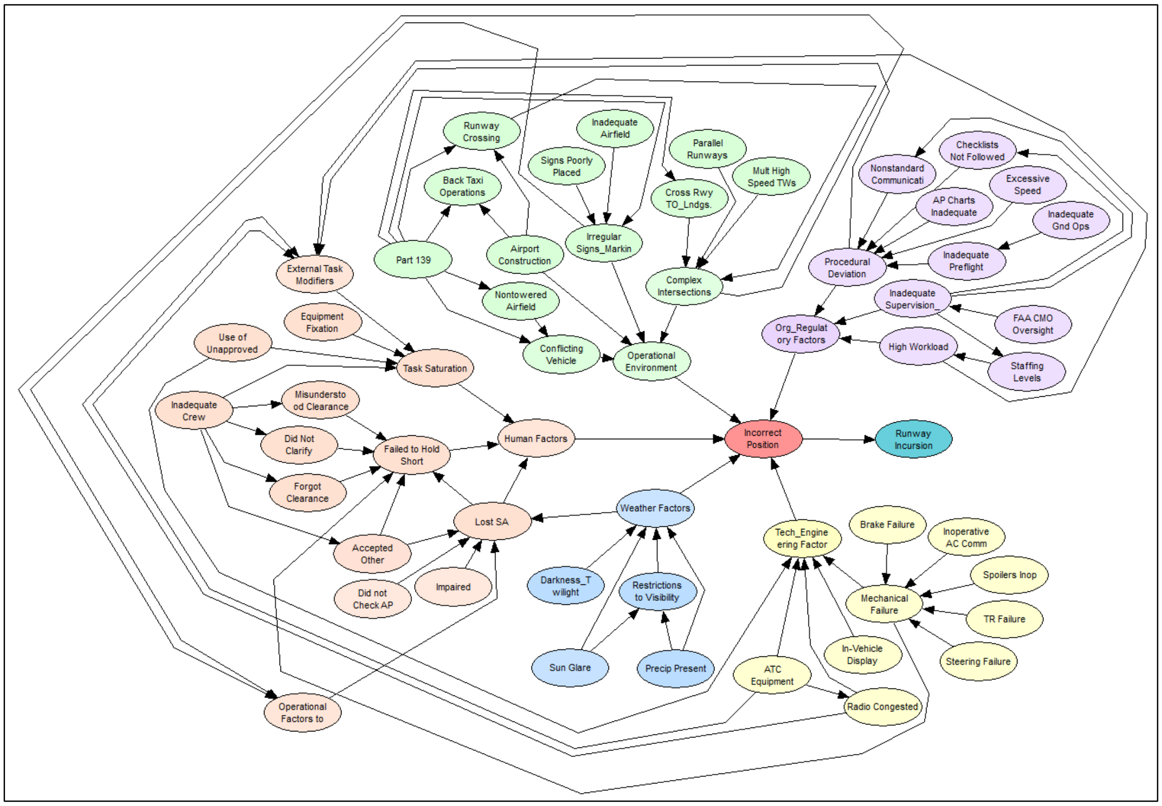


Figure 26. Full BBN model structure after SME review.

vii). Inter-domain edges connect related factors, illustrating the conditioning of one or more nodes on the variable state of another node or nodes.

Aside from nodes capturing airport construction, 14 CFR Part 139, non-towered airfields, darkness, visibility, precipitation, and mechanical failure, the model was quantified with data generated through expert judgment, as outlined in the following section. A complete list of variables and their definitions is included in the elicitation protocol and in Appendix J. Appendix K includes more detailed figures of the domain segments for each model. As discussed previously, the model was initially qualified with artificial distributions and dummy rank coefficients to test for function and correctness of dependence relationships. In this context, the model functioned as expected.

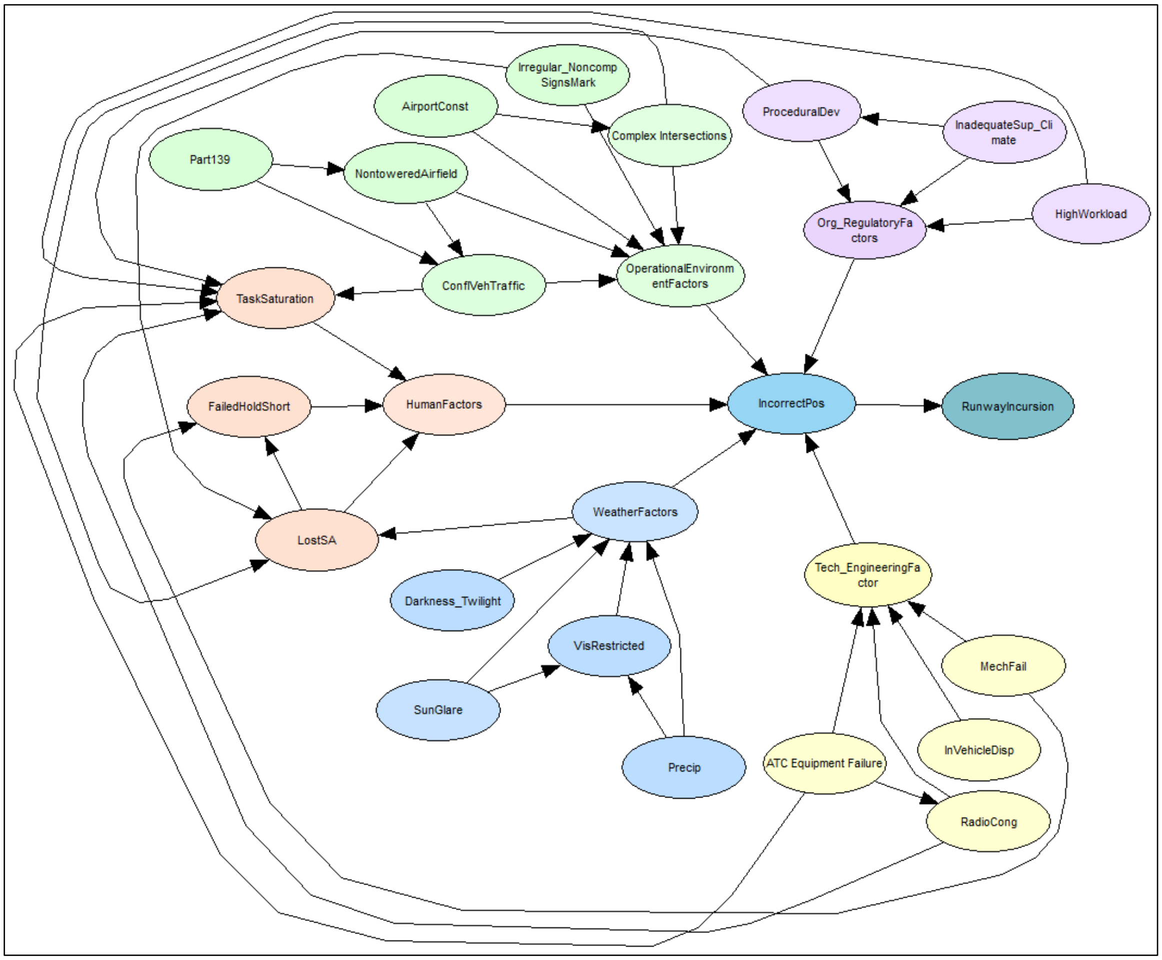


Figure 27. Parsimonious BBN structure after SME review.

Phase 3: Structured Elicitation

Structured expert elicitation was conducted to determine marginal distributions for variables where data were sparse or unavailable, to verify and validate model structure, and to quantify dependencies among interacting variables. Experts were presented with questions to which answers were known with certainty, and their responses were used to calibrate expert performance weights, as described in detail in Cooke (1991) and Cooke and Goossens (2000). Global and equal weight decision maker (DM) performance and information scores for the six experts across the 10 calibration questions used for this study are given in Table 11. These weights were applied to the marginal distribution and rank correlation elicitations within the scoring rules of the Classical model to achieve rational consensus among the expert judgments and calculate a DM based on the composite of expert performance and information. Decision maker responses to the seed variable (calibration) questions are given as range graphs in Figure 28 with the realization for each question indicated by the red hash mark and the DM range bounded by the 5th and 95th percentiles and the most likely value shown as the asterisk within the range. Range graphs for all experts over all questions are shown in Appendix I. Robustness analysis on seed items as well as on experts was completed, the results of which are reported as Figures 29 and 30, respectively.

Of note, and discussed in greater detail in the following chapter, is that a single expert was used in the derivation of the global-weight DM. Equal weighting, as the name implies, treats all experts' responses equally in calculating the DM. The results in Table 11 indicate that the Classical model global weighting and scoring method resulted in a higher calibration score, which can be interpreted similarly to *p*-value, though it is relative to 1 rather than 0 (values closer to 1 indicate better performance). In this way, calibration score is a reflection of an expert's assessments as compared to seed variable realizations. It is a means of expressing the degree to which the data supports the hypothesis that an expert's probability estimates are accurate (Aspinall, 2011). In this study, one expert (Expert B) was found to exhibit a relatively higher likelihood of uncertainty distributions reflecting true values, as shown in the calibration score column of Table 11. This result is discussed in more detail in the following chapter. The higher calibration score of the global-weight DM compared to the equal-weight DM indicated

Table 11. *Results of Scoring Experts.*

Expert ID	Calibration Score	Mean Rel. Total	Mean. Rel. Realization	Un-Norm. Weight	Norm. Weight w/o DM	Norm. Weight w/ DM
A	.0750	1.8684	1.1045	0	0	0
B	.7069	1.0943	.9585	.6776	1	.5
C	.0471	1.8177	1.2731	0	0	0
D	.0008	.8648	1.2680	0	0	0
E	.0063	1.8577	1.9258	0	0	0
F	.0471	1.7909	1.0974	0	0	0
Global	.7069	1.0943	.9585	.6776	--	.5
Equal	.5503	.3595	.2771	.1525	--	.1469

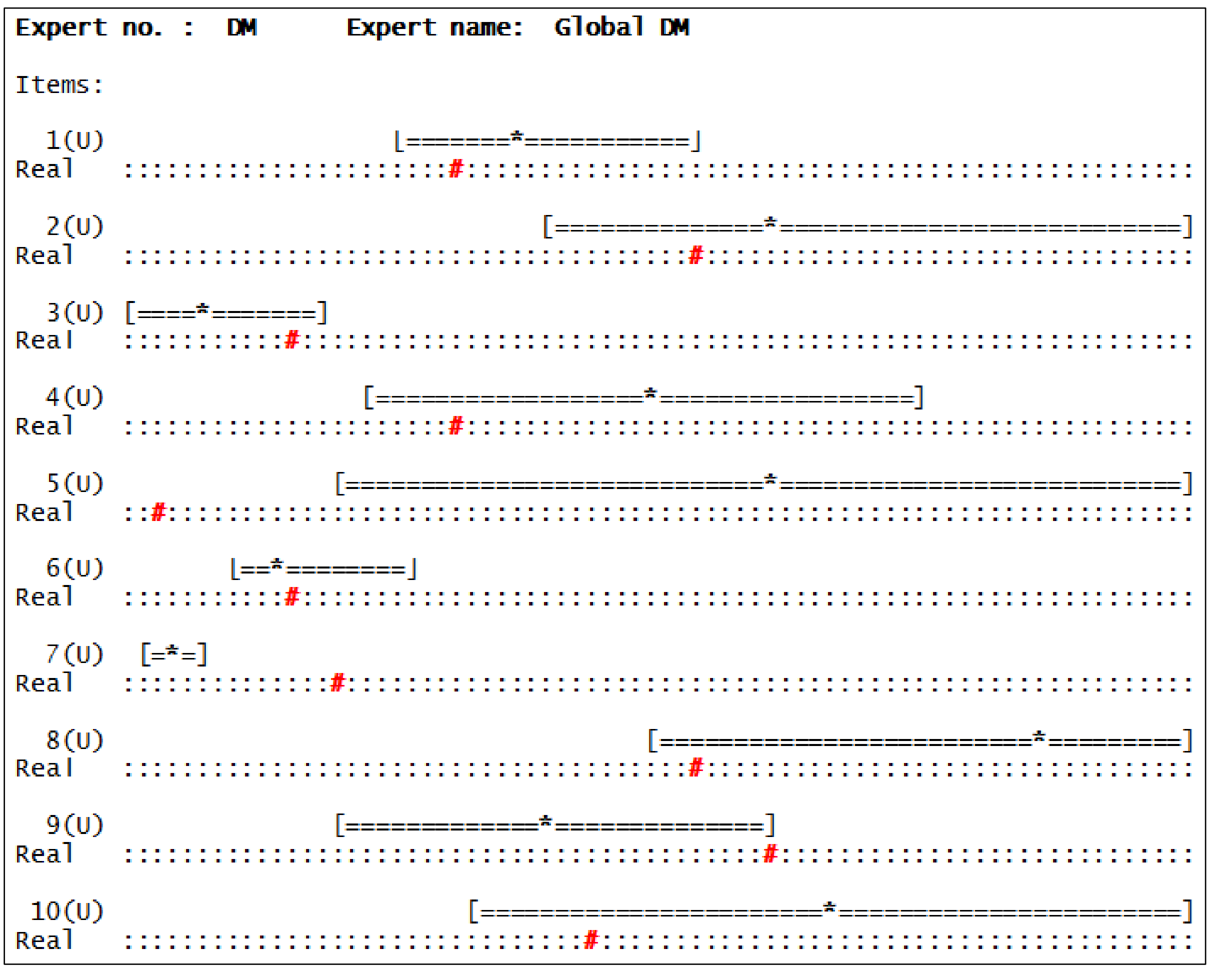


Figure 28. Range graph of Global DM assessments of seed variables (calibration questions).

Nr.	Id of excl. item	Rel.info/bgr. total	Rel.info/bgr. realization	Calibr.	Rel.info/or.DM total	Rel.info/or.DM realization
1	CQ1	1.0954574	0.94740486	0.73051077	0	0
2	CQ2	1.1150644	1.0149404	0.73051077	0	0
3	CQ3	1.081416	0.89904022	0.47917092	0	0
4	CQ4	1.1135941	1.0098757	0.73051077	0	0
5	CQ5	1.1242975	1.0467429	0.59248388	0	0
6	CQ6	1.0785764	0.88925999	0.47917095	0	0
7	CQ7	1.0431175	0.76712328	0.59248388	0	0
8	CQ8	1.1088237	0.99344468	0.73051077	0	0
9	CQ9	1.1067692	0.98636776	0.47917095	0	0
10	CQ10	1.1195968	1.0305519	0.73051077	0	0
11	None	1.0942901	0.95847511	0.70694172		

Figure 29. Item-wise robustness analysis.

Nr.	Id excl.exp	Rel.info/bgr. total	Rel.info/bgr. realization	Calibr.	Rel.info/or.DM total	Rel.info/or.DM realization
1	A	1.047	0.859	0.7069	0	0
2	B	0.7668	0.3085	0.4735	1.811	1.424
3	C	1.039	0.9005	0.7069	0	0
4	D	0.9619	0.9585	0.7069	0	0
5	E	1.044	0.9043	0.7069	0	0
6	F	1.009	0.7952	0.7069	0	0
7	None	1.094	0.9585	0.7069	0	0

Figure 30. Expert-wise robustness analysis.

that the global-weight DM performed best, and should be used in favor of an equal-weight DM (or other methods of consensus) to quantify the BBN model. Though the individual calibration scores are included in Table 11, they provide detail on the DM only and were not used individually to quantify the model.

On the basis of the global-weight pooling of the Classical model, a DM was derived, and the uncertainty assessments of the DM are shown in Figure 28 with respect to the true values, or realizations, of the seed variables. Readers will note that in two cases, the DM failed to capture the true value of the calibration variables. This was expected based on the responses of the most highly-weighted expert, as presented in Appendix I, which also includes DM uncertainty distributions across the set of target variables. Although this is the case, the DM correctly captures the seed variable value in 80 percent of the calibration questions.

Robustness analysis was performed on a case-wise and on an item-wise basis to evaluate the impact of each question and expert on the global-weight DM. As shown in Figure 29, removal of seed variable questions influences the DM calibration score. As an example, removal of CQ3 from the 10-question seed variable pool reduced the calibration score of the global-weight DM to 0.4792 from 0.7069. As higher calibration scores (closer to 1.0) are more desirable, this was an indicator of the relative importance of CQ3 to the derivation of the DM. Figure 28 illustrates the results of a similar method of analysis with focus on the influence of individual experts. The DM was based on the assessments of Expert B as shown in Table 11, and the relative influence of this expert is confirmed in Figure 30. Removal of Expert B from the analysis would have resulted in a lowered calibration score for the global-weight DM from 0.7069 to 0.4735.

Model Quantification

Quantification of the model consisted in part of calculation and entry of rank correlation data as well as evaluation, validation, and input of DM uncertainty distributions into the model. A complete correlation matrix, which shows all correlation

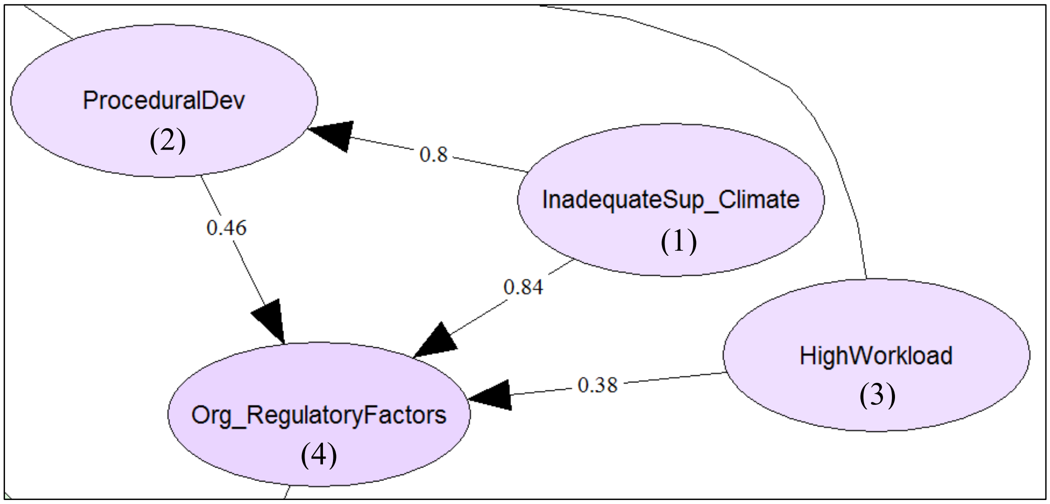


Figure 31. Organizational /regulatory subnet with rank correlation coefficients.

Table 12. *Dependence Information.*

Probability		Rank Correlation	
P_1	0.65	r4,1	0.84
P_2	0.55	r4,2| 1	0.46
P_3	0.46	r4,3 | 1, 2	0.38

pairs and rank correlation values, is included as Appendix L. Illustrating the premise of the rank correlation and the difference in coefficients are Figure 29 and Table 12, which focus on the results of a single domain subnet.

Field data and data from the structured expert judgment phase were also entered into the model as described previously. Figure 32 shows a high level overview of how the model treats the uncertainty distributions and functions as a continuous, dynamic network. Appendix M includes more detailed information on the quantified nodes and rank correlation for each domain subnet. In Figure 32, the central nodes (child nodes) of each domain, the node representing incorrect position, and the RI probability node are represented as histograms. The baseline probability of the model for RI occurrence was

0.000525, or roughly 525 RI occurrences per million flight operations. Relative to the RI rate per million operations figures reported by the FAA, which was nearly 19 in the most recent Runway Safety Report (see Figure 1), the model performance appears to overestimate the occurrence of RI events by a substantial margin. However, a key feature of the model is that it aims to predict occurrences at all airports and inclusive of all operations types within the US, not solely those with an ATC facility, and thus this result is not altogether unexpected or unusual. Given this, the model was acceptable in its base-rate prediction, and was validated for the purposes of further analysis.

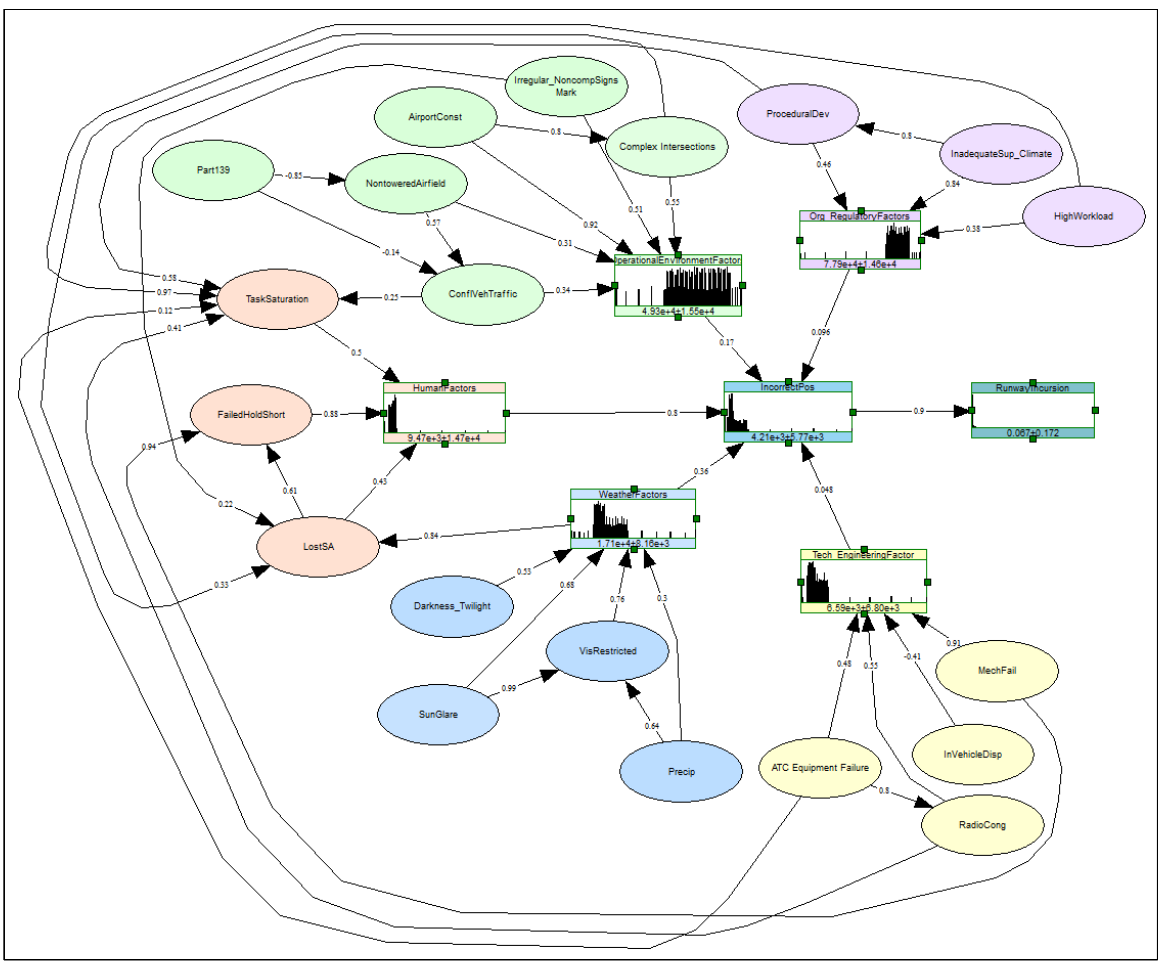


Figure 32. Final, quantified model with domain nodes, improper position, and runway incursion as histograms.

Figure 33 is a closer view of the central part of the model as shown in Figure 27, though in Figure 33 the model has been conditionalized (analyzed based on a manually selected point sample for one or more variables in the model) on the RI variable to examine the effect of reverse propagation of RI occurrence on the causal, dependent model structure. The updated histograms based on the conditioning show black while the original, unconditionalized distributions are displayed as grey. The model was also conditionalized to propagate evidence as it would normally occur. Based on the data and SME input from earlier phases, *complex intersection*s and *task saturation* were identified as having strong correlation directly to RI events or to other causal factors. In Figure 34, the results of conditionalizing on these two variables at values substantially above

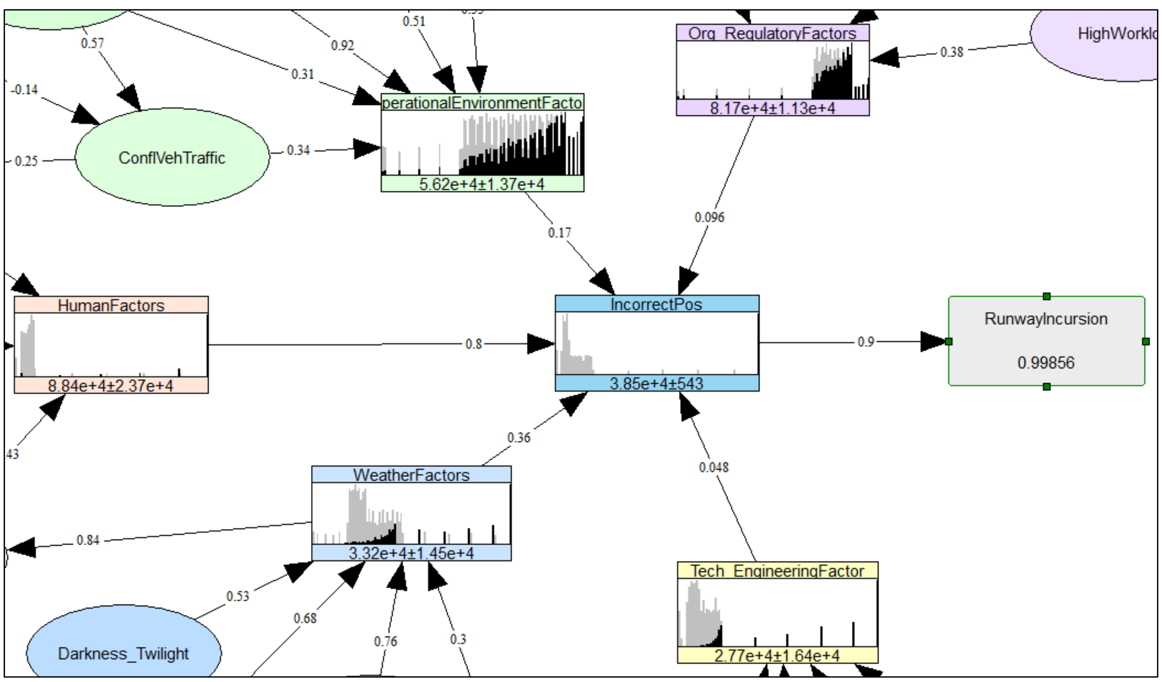


Figure 33. Final model conditionalized on RI occurrence.

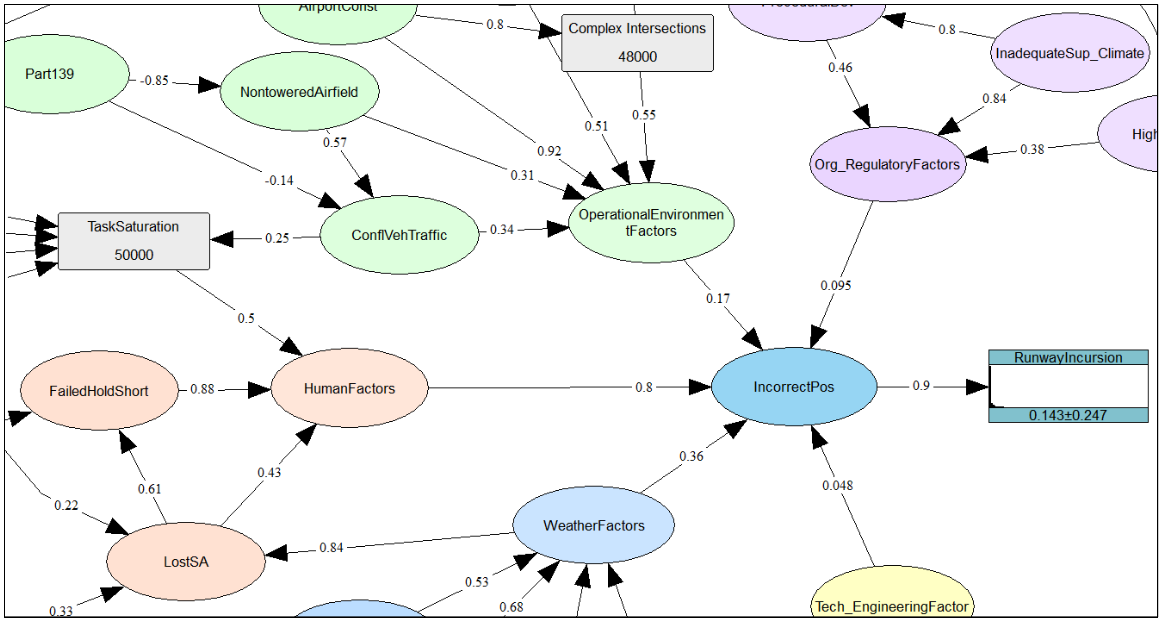


Figure 34. Propagation of evidence through the model.

median are shown. At the extreme values entered for the variables of interest, the probability of RI increased from 0.0005 to 0.0107. This attribute of propagating evidence through the network is one of the defining characteristics of a BBN, as discussed earlier, and allows both diagnostic and predictive reasoning. Some of the more telling results of back propagating on the occurrence of an RI are illustrated in Table 13, which shows selected relative increases from baseline resulting from conditionalizing on RI occurrence for variables exhibiting high rank correlation within their domain. When an RI is forced in the model, the variables in Table 13 increase as shown, indicating to some extent the strength of association with an RI event.

Using UNISENS and UNIGRAPH, sensitivity analysis was performed to assess the prominent interactions between model variables, with particular attention paid to those interfaces that occurred across domains. Analysis of the entire model was

Table 13. *Effect on Selected Variables of RI Occurrence.*

Variable	Approx. % Increase from Baseline
Inadequate Supervision/Climate	92%
Complex Intersection	28%
Airport Construction Present	62%
Task Saturation	57%
Failure to Hold Short	164%
Sun Glare	251%

consistent with data obtained during the literature review and in phases one and three of this study, showing that human-centric errors such as lost situational awareness and task saturation were influential individual factors. By looking at each domain in terms of sensitivity analysis and identifying associations, however, causal paths became clearer. Within the organizational and regulatory domain, *procedural deviation* was most predictive of abnormal factors in that area. *Procedural deviation* was joined to *task saturation* by a directed arc in the model, as were several other variables, indicating these connections may warrant further investigation. Table 14 gives the most influential single variable within each domain and presents linear least squares fitted regression coefficient and correlation ratio, the squared product moment correlation that maximizes the correlation value, for each. Correlation ratio was used here to interpret the ratio of variance of variable Y given X and the variance of Y. As indicated in Table 14, mechanical failure has an unusually high relative regression coefficient, though this is explained by the combined effects that a failure in a braking or steering system may have, combined with the rarity of these events.

To look at these interactions in more detail, and bearing in mind that *human factors* were shown to have the strongest causal influence on RI events, additional

Table 14. *Regression and Correlation Coefficients for Selected Variables.*

Predicted Variable	Base Variable	Correlation Ratio	Regression Coefficient
Org_RegulatoryFactors	ProceduralDev	0.5529	0.6636
OperationalEnvironmentFactors	AirportConst	0.7906	0.8590
HumanFactors	FailedHoldShort	0.5495	2.1218
WeatherFactors	SunGlare	0.5747	0.7735
Tech_EngineeringFactor	MechFail	0.5765	14679.6211

sensitivity analysis was performed on all other random variables as predictors of abnormal human factors conditions. *Mechanical failures* ranked highest from a regression standpoint, a result that was not surprising given the disruptive nature of systems failures on operator performance noted in the literature and by SMEs. Somewhat unanticipated however was that restrictions to visibility, including *sun glare*, had even more strength of influence than factors and constructs such as airport construction or high workload. Arguably, weather related factors are less variable and known with greater certainty that the organizational and operational environment factors that follow them in magnitude of association, and thus illustrate that sensitivity analysis alone does not explain the model completely, which must instead be evaluated on the basis of many factors within the model. It was used, however, to further validate the model by allowing variable sensitivity and covariance to be evaluated once again for dependence and retention in the model structure to verify that parsimony was achieved as intended.

UNIGRAPH output lent further illumination to the results discussed previously by graphically presenting selected dependence information. In Figure 35, the left most variable is *runway incursion*, with high values of runway incursion selected (above 0.75) and the resultant values for additional model variables shown. Quite logically, the figure illustrates that for high values of RI occurrence, *failure to hold short* and *incorrect position* are also high. Also worth noting in Figure 33 is that *procedural deviation* and *task saturation* are also quite high. To a lesser extent, *airport construction* and *conflicting vehicle traffic* are also elevated above median. However, *darkness/twilight* and *ATC equipment failure* are both generally lower values (centered about the median) when compared to high RI occurrence. Returning to the model with this information, additional evidence propagation was implemented and is shown in Figure 34. For the final round of evidence propagation, variables identified in the previous results were

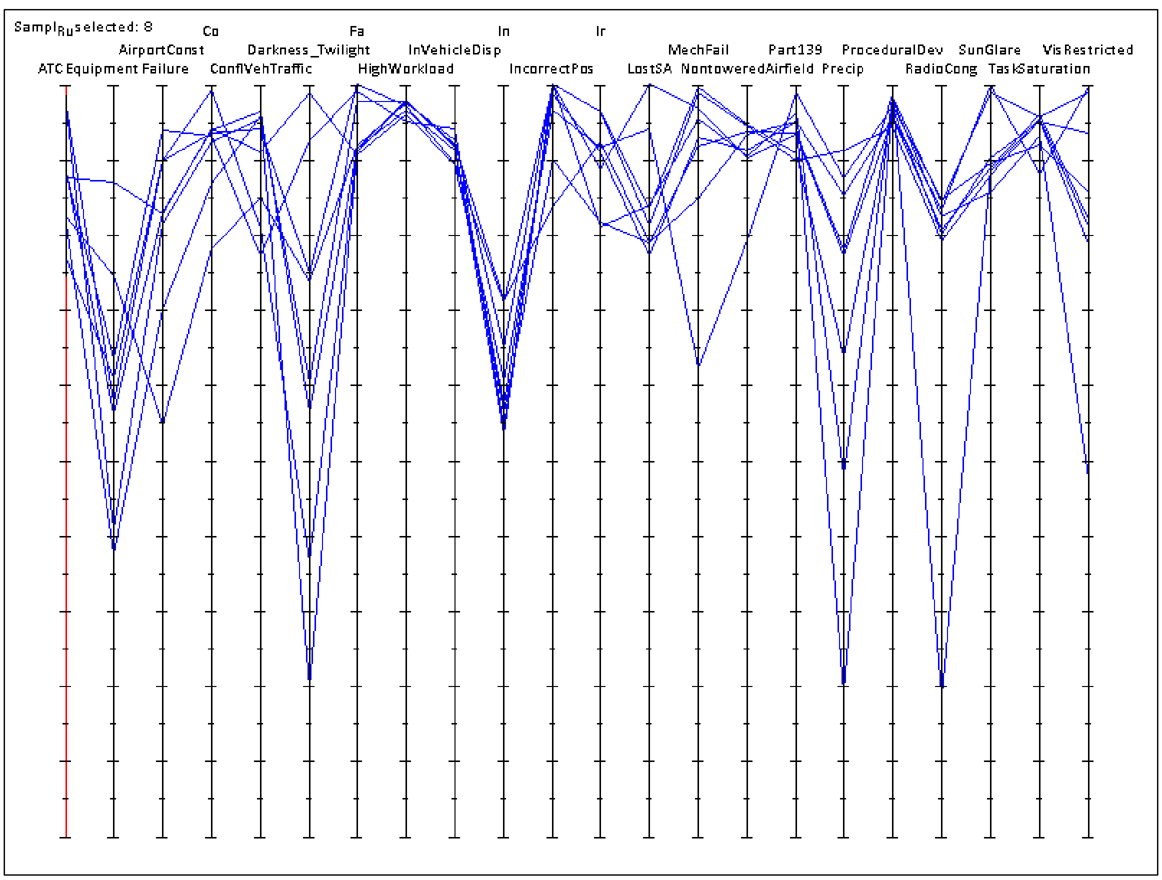


Figure 35. Cobweb plot of selected model variables.

conditionalized at their 95th percentile values, indicating that the SMEs would be surprised if the true value was in excess of this number. Figure 36 shows the model with conditionalized variables as grey blocks, and the target variable, *runway incursion*, as a histogram. Conditionalizing on these variables showed the combined result of their interaction, and represented what has been often referred to as a causal path in this study

The model and high-level sensitivity analysis completed in this study indicated that the variables conditionalized in Figure 36 had the greatest impact in combination of

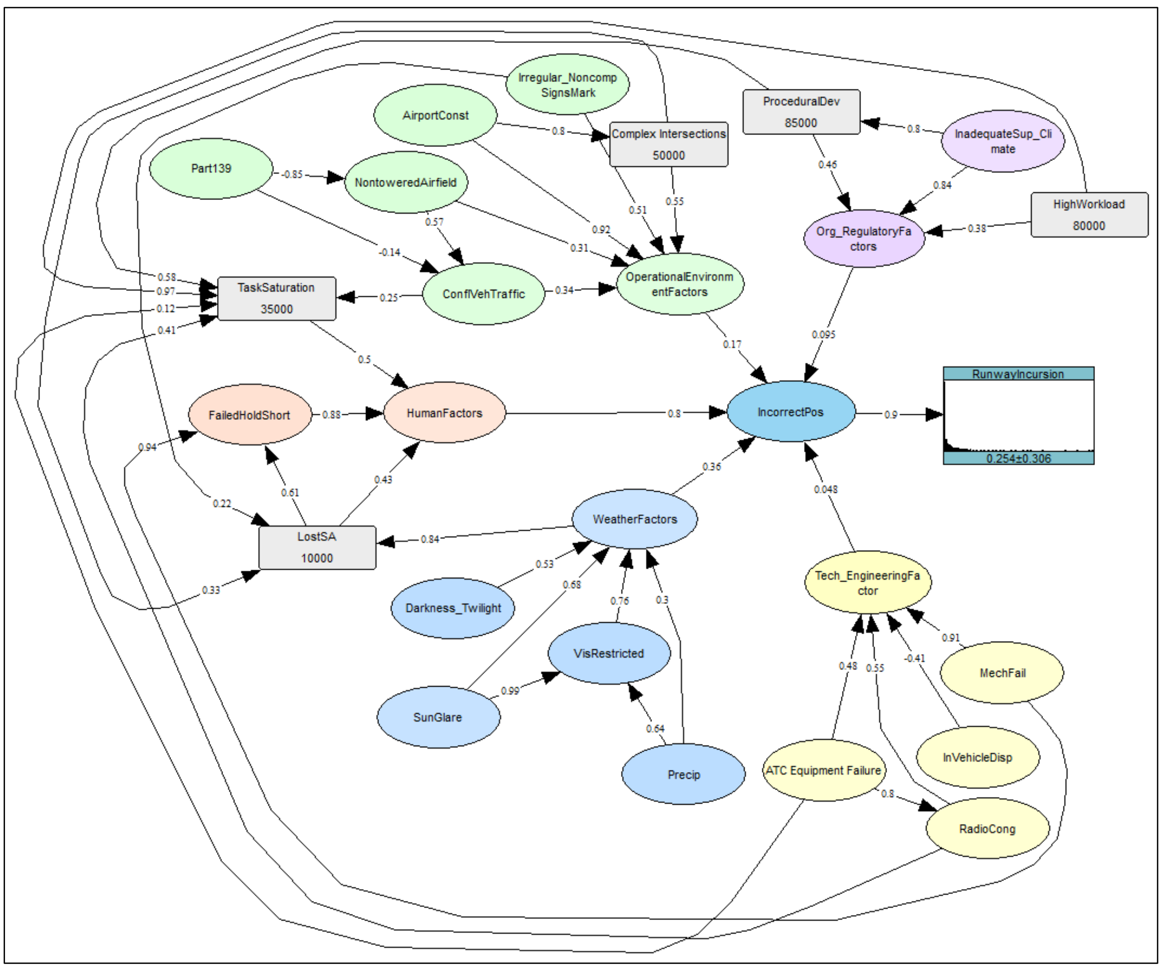


Figure 36. Model conditionalized on indicated variables to propagate evidence.

those variables for which direct intervention is possible (as compared to weather or construction, as examples of conditions that are not inherently controllable to the same extent). By propagating 95th percentile values for these variables, the rate of occurrence for RI events was observed to increase to approximately 99,000 per million operations. This number was alarmingly high, and is a forced manipulation of the model. However, the resultant probability provided some indication as to the strength of influence of the variables in the parsimonious model acting in dynamic collaboration.

One of the most important points to be considered when evaluating the model was that identification of causal paths and interactions between variables was not a static or single-point process. Instead, assessing the model on the basis of rank correlation, sensitivity analysis, and both forward and reverse propagation of evidence was a requirement to achieve a holistic view of model performance and to evaluate the model against what is known. By approaching the model in this way, validity was also addressed, in that many results were evaluated in the course of drawing conclusions about causal interactions.

CHAPTER V

DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

The primary focus of this study was to develop and validate a holistic, system-level model for U.S. RI event causation that could serve as the basis for future model development and evaluation of RI reduction proposals, particularly by agencies such as the FAA. Two research questions refined this purpose, asking: What are the interacting causal factors that lead to RIs in the U.S.; and, can runway incursions in the U.S. and their dynamic causal factors and interactions be modeled through the use of a Bayesian belief network supported by expert-elicited data? To that end, this research introduces a mixed-method approach for informing the model content and structure, develops a BBN representation of RI causation in an iterative, operationally-oriented process, and builds upon established methodology for structured expert judgment to create an intuitive protocol for quantifying the model. The following discussion addresses some of the theoretical and practical implications of this study, and then presents conclusions and recommendations for further research.

Discussion

Results of this study are reported in Chapter 3, and the following discussion offers further interpretation of the results, with a focus on practical implications of the outcome of the study.

Phase 1: Runway incursion data and causal factors. It is demonstrated that some discrepancies between RSO and ASRS data exist, and these differences are largely accounted for within the literature. As an example, RSO data indicates clearly that a preponderance of RI events involved aircraft operating under 14 CFR Part 91as compared with commercial (14 CFR Parts 135 and 121) operations. Within the ASRS data, this trend is reversed. The use of SMEs to review ASRS reports provides depth of understanding when reviewing these differences, and reinforces the need to consider the underlying sources and motivations that affect RI data. In the context of incorporating emic, thematic perspectives from pilot-reported RI events, the model benefited as domains were explored and expanded from an initial three (human factors, technology, and environment) to five. Remembering that the number of possible causal factor iterations per case was potentially over one million, achieving acceptable interrater reliability for one and two causal matches across rater pairs (and relatively strong results for three matches) is rather notable, and indicates consistency in interpretation of causation. This agreement among raters leads to confidence in the identified causal themes, and allows domains to be expanded and further specified to include human factors, operational and regulatory elements, weather factors, technological and engineering influences, and operational environment factors that may act independently or in dynamic interaction to contribute to RI occurrence.

Phase 2: Belief network model. A secondary research question to the matter of RI causation was whether or not RIs could be modeled using a continuous BBN supported by structured expert judgment. The matter of expert judgment is addressed in the following section, but the results presented here demonstrate that in fact, RI events can be modeled for causal interpretation using BBNs. The review of literature clearly reveals that RIs are not isolated or static events, but are instead the result of a complex series of dynamic interactions. BBNs were proposed as an appropriate methodology for investigating RI causation, not only because they support dynamic probabilistic reasoning, but because quantifying them with expert-elicited data is a natural feature. The results of the study shows that this was indeed a critical component for consideration, given that many of the constructs or factors that were identified in the first phase were not supported through available and sufficiently detailed data, requiring expert opinion for insertion into the model. An unstated, but important consideration of this study is the operationalization of methods so that use of modeling methods such as BBNs can be made more accessible to a greater number of users. From a practical standpoint, the precision of the model is sufficiently accurate to reflect what is known about RI events, and it achieves this accuracy over a broader scope of causation than previous studies. Additionally, the model structure uncovers more detailed information about the dynamic causation of such events, revealed through sensitivity analysis of the BBN.

Phase 3: Structured expert judgment. Structured expert elicitation played a pivotal role in the quantification of the BBN model, and the results achieved here do more than validate the use of the Classical model in the context of RI causation. The format used to elicit the rank correlation and conditional probability information from experts appears to be the first use of such a technique to quantify a continuous BBN. Confirming Cooke and Goossens (1999) assertion that experts are not opposed to performance measurement and in fact react quite well to answering questions in a format that captures uncertainty, the SMEs in the study unanimously commented that the structured elicitation methods were productive, insightful, and generally supported intuitive reasoning. In training exercises using probability assessment in lieu of describing influence as a ratio, all six experts agreed that the ratio method allowed them to express the same ideas regarding rank correlation and conditionality, but in a more natural way. As a result of this more instinctual assessment, the process was also completed more quickly than when the more common, probabilistic elicitation of the Classical model was used. Further extending the work of Roelen et al. (2008), which used a single expert, probability assessments in the elicitation process were made in the ordinary manner, but were normalized to function with the constraints of the normal copula. By approaching the SMEs in this way, the need to provide feedback in terms of rank correlation is reduced, which was helpful given the difficulty for experts to translate the transformed numbers.

As noted in the earlier presentation of results, the DM derived from the structured expert judgment sessions utilized only one of the six experts who took part in the elicitation. Although in the context of a consensus-based method this may be considered unacceptable, in this study, the scientific derivation of the DM provides support for the identification of a single SME for inclusion in the DM, and this is not altogether unusual. Van Der Fels-Klerx, Cooke, Nauta, Goossens, and Havelaar (2005), experienced a similar outcome, and such results, though possibly underreported, are not problematic (Dr. R. Cooke, personal communication, January 15, 2013). The purpose of the Classical model is to provide a structured methodology for weighting experts based on accuracy and precision. Just as a medical testing laboratory may only have a single instrument for testing DNA, the resultant DM from this study contains only one SME. Like the lab, which almost certainly would ensure a machine's accuracy through careful testing and calibration, the DM in the Classical model is also tested and calibrated to ensure accuracy and optimal performance (Dr. M. Wittmann, personal communication, July 23, 2013).

Essentially, when a DM is based on a single, or few, experts, it is often less an indication that the other experts' performance was subpar, but that their accurate responses are in fact contained within the responses of the expert whose assessments form the DM.

Model completion, testing, and evaluation. The model developed in this study represents the first known attempt to characterize RI causation in the US across all domains using a BBN. From the perspective of causation, reverse propagation is revealing in showing the adverse effects of interference with operator (flight crew of vehicle operator) visibility. This is interesting, but not necessarily unexpected. The effect of inadequate supervision and safety climate also exercised substantial influence in the model, and affected procedural deviation probability as well as the occurrence of task saturation because of interconnected edges, illustrating the multifaceted causation of RI events. When examining the combination of variables that most affected RI occurrence, human factors and organizational factors dominated, followed by the effects of abnormal operational environment conditions. From a practical perspective, operational issues are superficially easier to address, but treating the root causes of organizational and human factors is a more complex and layered problem. The model supports investigation into efforts to do so, however; and it allows end users to evaluate changing organizational environments or training efforts to improve latent human factors inadequacies. At an operational level, this model translates to a useful tool for evaluation of future mitigation efforts by regulators or even operators with the requisite sophistication and data to do so.

Conclusions

Consistent with recommendations that RI research include pilot, ATC, and airport influences (Torres, Metscher, & Smith, 2011), the model developed as a result of this study demonstrates that it is possible to gain important insight into causation of RI events through holistic, dynamic modeling methodologies such as BBNs. As Hendrickson (2009) suggests, a greater understanding of aviation accident factors, especially those that are principally rooted in human error, can be uncovered through text analysis of operator reports. The mixed-method approach to deriving cross-domain variables for inclusion in the model presented in this study is shown to be effective, and is acceptable to experts who evaluated the data through an incremental review process. This methodology expands the current body of knowledge with respect to both RI investigation and model building by enhancing current methods that often focus solely on quantitative information. The specific methods presented here for review of ASRS narratives and evaluation of rater responses creates a more refined and systematic process, improving upon some of the more loosely-defined text analysis that appears in the literature. In deriving supporting data for the model from varied sources that included ASRS pilot reports, the emic and etic perspectives were combined to create a more inclusive picture of RI causation than has been developed in past studies, especially considering the more detailed model shown in Figure 21. From the perspective of addressing RIs as a complex and interdisciplinary problem, the combining of quantitative and qualitative data here enables a fuller understanding of the dynamic interactions and dependencies of complex operational incidents such as RIs (Stroeve, Blom, & Bakker, 2013).

Developing the content and structure of the model iteratively allowed for the model to be made sequentially more compact (as in Figure 25), easing the process of verification and validation, quantification, and assignment of rank correlation for a preliminary model. The general procedural guidance from Marcot, et al. (2006) formed a foundation for the construction of the model structure and refinement of content, but certain liberty was taken given the lack of available data and a fundamental desire to make creating such a model more operationally accessible. This development method also resulted in a more comprehensive model (Figure 21), which was intended as a next step in the quantification process and which allows a more detailed assessment of variable interaction from the perspective of sensitivity analysis or what-if scenarios. The ability to work with a more compact model in the initial stages of development makes this study less theoretical in nature, and more translational, in that the iterative model building is somewhat less resource-intensive than ordinarily academic BBNs and is therefore more apt to be extended beyond systems-level application to the operator level. The model developed in this study showed that human factors, as expected, play a pivotal role in RI causation. Differentiating it from previous work in this area, however, is that this research shows more intuitively and dynamically than previous studies how causal factors within and across domains interact to affect the probability of an RI event.

The elicitation of marginal distributions showed that more specific, objective variables are preferable whenever possible. Though it is clear that empirical data should be used when they are available, it is the case in this study that many of the factors identified as causal to RI events are constructs for which achieving objectivity is difficult and some subjectivity must be accepted in the course of inquiry. As has been shown in other studies (e.g. Roelen, van Baren, Smeltink, Lin, and Morales, 2007), a lack of objectivity manifests itself in difficulty among experts in estimating the combined influence of variables as in the present case. In this study, the method by which marginal, rank, direction and strength of correlation were elicited was shown to alleviate this issue. In addition, this method of elicitation allowed for structured expert judgment to be completed in less time and more intuitively while maintaining accuracy. Structured expert judgment is an established methodology, but its use here is refined and extended, and it shows that such methods lend more scientific scrutability to otherwise inaccessible data, most notably in the organizational and human factors domains.

This research is purposely constrained to include only U.S. aviation operations with respect to RI event analysis and prediction. The resultant model and expert judgment information are both useful and clearly generalizable to that population. Beyond aviation operations in the U.S., however, the methodology presented here for structuring the model and for eliciting expert judgment is applicable to construction and quantification of any continuous, nonparametric BBN that uses structured expert elicitation. The particular protocol for expert judgment is relevant to a broad audience, and even to applications outside aviation. Given the difficulty in quantifying organizational and human factors data, expert knowledge plays an important role in furthering study of these concepts through modeling. While previous studies have often identified human factors as a critical causal element to RI events, this study extends these observations by developing and validating a platform for causal inference across domains, and with consideration of interacting causal components. Specifically, the interactions between human factors, organizational influences, and factors within the operational environment are evaluated in dynamic interaction to show the increased threat of RI occurrence when these elements functioned in combination. Reverse propagation of RI occurrence through the model also illustrates the effects of causal factors such as sun glare and supervisory issued on RI causation. Most notably, the model stands as a basis for future investigation of RI mitigation efforts. This expansion of the relevant knowledge of RI causation and reduction is accomplished while also reducing the elicitation and model structuring burden to make the methods detailed here more operationally relevant.

Recommendations

The scope of this study necessarily limited the extent to which the model presented was analyzed, expanded, and revised. As such, extensions to this research logically converge on a continuation of this study involving the complete model presented here, or a variation of the models in this research. Validation of the complete, detailed BBN model as respects frequency of occurrence and evaluation of causal interactions of increased complexity would contribute meaningfully to the body of RI knowledge, as suggested by Biernbaum and Hagemann (2012).

Extensions to the present study may also focus on more complete validation of the elicitation methodology used here, with attention paid most closely to the question structure used to obtain expert opinion of rank correlation and the mathematics involved in assuring a positive definite correlation matrix as well as some of the other limitations that arise in quantifying a continuous, nonparametric model through structured elicitation. Verifying the information in the model presented here is tenuous in many cases because insufficient data exists for many of the causal factors that make up the model. In part, this points to the scarce availability of data describing the detailed factors that combine to cause RIs and to a demand for availability of more specific causal data so that subjectivity may be reduced to the greatest extent possible.

Future research may also be directed toward converging frequency data with severity, considering that the present study concerns itself only with the *occurrence* of an RI, not to the extent with which a conflict manifests as an incident or accident. Research may also be extended toward further testing of this methodology with the aim of reducing subjectivity to the greatest extent possible. Primarily, efforts should be focused on further development of the rank correlation elicitation methods and on quantifying the complete model, or some variation of it, presented here. This need again underscores the lack of data describing the detailed factors that combine to cause RIs (GAO, 2011) and to a demand for availability of more specific causal data, a problem partially addressed by this study, but a persistent issue to which further research should be focused.

Finally, additional work to validate potential factors for inclusion in the model is warranted, as is evaluation of RI mitigation programs and technology. Given thorough verification and validation of a more complete model, what-if analysis can be explored to test the effectiveness of various strategies for RI reduction, reducing the need for testing mitigation efforts in real time within the national airspace system.

This study has illustrated in a dynamic, intuitive platform the interconnected nature of RI causation, and has validated a model for evaluating RI events across a number of causal domains. BBN models are shown to be an appropriate means of investigating RI causation, and structured expert judgment is demonstrated as a natural and informative methodology in the presence of uncertainty and sparse data. Future research should address causation in more detail, and should extend models to include insertion of mitigation efforts to evaluate effectiveness in stemming the growing potential for RI events in the U.S.

REFERENCES

Adam, G., & Kelley, D. (1994). *Reports by airline pilots on airport surface operations: Part 1. Identified problems and proposed solutions for surface navigation and communications* (MTR-94W0000060). McLean, VA: MITRE Corporation.

Adam, G., & Kelley, D. (1996). *Reports by airline pilots in airport surface operations: Part 2. Identified problems and proposed solutions for surface operational procedures and factors affecting pilot performance* (MITRE Report No. MTR-94W0000060v2). McLean, VA: MITRE Corporation.

Akkermans, D. E. (1989, October). *Crane Failure estimates at DSM*. Paper presented at the Expert Judgment in Risk and Reliability Analysis; Experience and Perspective, ESRRDA Conference. Brussels, Belgium.

Albert, I., Donnet, S., Guihennec-Jouyaux, C., Low-Choy, S., Mengersen, K., & Rousseau, J. (2012). Combining expert opinions in prior elicitation. *Bayesian Analysis, 7*(3), 503-532.

Ale, B. J. M., Bellamy, L. J., van der Boom, R., Cooper, J., Cooke, R., Goossens, L. H. J.,...Spouge, J. (2009). Further development of a Causal model for Air Transport Safety (CATS): Building the mathematical heart. *Reliability Engineering & System Safety, 94*(9), 1433-1441.

Apostolakis, G. (1990). The concept of probability in safety assessments of technological systems. *Science, 250*(4986), 1359-1364.

Aspinall, W. (2011, July). Expert elicitation and judgment pooling using the Cooke Classical Model and EXCALIBUR. Retrieved from http://risk2.ewi.tudelft.nl/ejwshop_2011_files/Aspinall_Briefing_Note_on_Classical_Model_and_EXCALIBUR_Jul_2011.pdf

Aspinall, W. P., & Cooke, R. M. (1998). Expert judgment and the Montserrat Volcano eruption. In A. Mosleh & R. A. Bari (Eds.), *Proceedings of the 4th International Conference on Probabilistic Safety Assessment and Management PSAM4: Vol. 3* (pp. 2113-2118). New York, NY

Bakker, M. R. A. (2004). *Quantifying operational risks within banks according to Basel II* (Doctoral dissertation, TU Delft, Delft, The Netherlands).

Bales, R. A., Gilligan, M. R., & King, S. G. (1989). *An analysis of ATC-related runway incursions, with some potential technological solutions* (MTR-89W00021). McLean, VA: MITRE Corporation.

Bedford, T., & Cooke, R. (2001). *Probabilistic risk analysis: Foundations and methods.* New York, NY: Cambridge University Press.

Bedford, T. J., & Cooke, R. M. (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics, 30*(4), 1031-1068.

Bedford, T., Quigley, J., & Walls, L. (2006). Expert elicitation for reliable system design. *Statistical Science, 21*(4), 428-450.

Bellantoni, J., & Kodis, R. (1981). *An analysis of runway-taxiway transgressions at controlled airports* (FAA-EM-81-5/DOT-TSC-FAA-81-9). Cambridge, MA: Transportation Systems Center.

Biernbaum, L., & Hagemann, G. (2012). *Runway incursion severity risk analysis* (DOT-VNTSC-FAA-12-13). Cambridge, MA: Volpe National Transportation Systems Center.

Billings, C. E., & O'Hara, D. B. (1978). Human factors associated with runway incursions. *ASRS Eighth Quarterly Report, NASA TM-78540*.

Borsotto, M., Zhang, W., Kapanci, E., Pfeffer, A., & Crick, C. (2006). A junction tree propagation algorithm for Bayesian networks with second-order uncertainties. In *Proceedings of the 18th IEEE International Conference on Tools with Artificial Intelligence. ICTAI '06* (pp. 455-464).

Brooker, P. (2011). Experts, Bayesian belief networks, rare events and aviation risk estimates. *Safety Science, 48*(8-9), 1142-1155.

Brown, A. J., & Aspinall, W. P. (2004). Use of expert opinion elicitation to quantify the internal erosion process in dams. In *The 13th Biennial British Dams Society Conference: University of Kent, Canterbury*. Canterbury, UK

Brown, J., Goossens, L. H. J., Harper, F. T., Haskin, E. H., Kraan, B. C. P., Abbott, M. L.,...Rood, A. (1997). *Probabilistic accident consequence uncertainty analysis: Food chain uncertainty assessment* (NUREG/CR-6523). Washington, D.C.: U.S. Nuclear Regulatory Commission.

Cantor, A. B. (1996). Sample size calculations for Cohen's kappa. *Psychological Methods, 1*(2), 150-153.

Cardosi, K. M. (2001). *Runway safety: It's everybody's business* (DOT-FAA/AR-01/66). Washington, DC: US Department of Transportation.

Cardosi, K., & Yost, A. (2001). *Controller and pilot error in airport operations: A review of previous research and analysis of safety data* (DOT/FAA/AR-00/51). Washington, DC: U.S. Department of Transportation.

Cardosi, K. M. (2005). Runway Safety. In B. Kirwan, M. Rodgers, & D. Schafer (Eds.), *Human factors impacts in air traffic management (*pp. 43-70). Aldershot, UK: Ashgate.

Chang, Y., & Wong, K. (2012). Human risk factors associated with runway incursions. *Journal of Air Transport Management, 24*, 25-30.

Clemen, R. T. (1986). Calibration and the aggregation of probabilities. *Management Science, 32*(3), 312-314.

Clemen, R. T. (2008). Comment on Cooke's classical method. *Reliability Engineering and System Safety, 93*, 760-765.

Clemen, R. T., & Lichtendahl, Jr., K. C. (2002). Debiasing expert overconfidence: A Bayesian calibration model. In E. J. Bonano, A. L. Camp, M. J. Majors, & R. A. Thomson (Eds.), *Probabilistic Safety Assessment and Management (PSAM6): 6th International Conference on Probabilistic Safety Assessment and Management*. Oxford, UK: Pergamon.

Clemen, R. T., & Winkler, R. L. (1999). Combining probability distributions from experts in risk analysis. *Risk Analysis, 19*(2), 187-203.

Commercial Aviation Safety Team [CAST] (n.d.). Background. Retrieved December 12, 2011, from http://www.cast-safety.org/about_vmg.cfm

Commercial Aviation Safety Team [CAST] (2007). *Wrong Runway Departures*. Washington, DC: U.S. Department of Transportation.

Controllers on Airport Surface Operations: The Causes and Prevention of Runway Incursions (Report No. MTR- 98W0000033). McLean, VA: The MITRE Corporation.

Congdon, P. (2003). *Applied Bayesian modeling.* West Sussex, UK: John Wiley & Sons.

Conrady, S., & Jouffe, L. (2011, February 11). Introduction to Bayesian networks. Retrieved from http://www.conradyscience.com/index.php/introduction-to-bayesian-networks

Cooke, R. M. (1991). *Experts in uncertainty.* New York, NY: Oxford University Press.

Cooke, R. M. (2004). The anatomy of the squizzel. *Reliability Engineering & System Safety, 85*, 313-319.

Cooke, R. M. (2009). Obtaining distributions from groups for decisions under uncertainty. In T. M. Williams, K. Samset, & K. J. Sunnevag (Eds.), *Making essential choices with scant information: front-end decision making in major projects (*pp. 257-274). London, UK: Palgrave-Macmillan.

Cooke, R. M., & Goossens, L. J. H. (1999). *Procedures guide for structured expert judgment* (EUR 18820). Luxembourg, Luxembourg: European Commission.

Cooke, R. M., & Goossens, L. H. (2006). TU Delft expert judgment data base. *Reliability Engineering & System Safety, 93*(5), 657-674.

Cooke, R. M., & Probst, K. M. (2006). *Conference summary: Highlights of the expert judgment policy symposium and technical workshop*. Paper presented at the Expert Judgment Policy Symposium and Technical Workshop. Washington, D.C..

Cowell, R. (1999). Introduction to inference for Bayesian networks. In M. I. Jordan (Ed.), *Learning in graphical models (*pp. 9-26). Cambridge, MA: MIT Press.

Dabipi, I. K., Burrows-McElwain, J. B., & Hartman, C. (2010, October). *Low cost runway incursion detection system for general aviation airports*. Paper presented at the 40th ASEE/IEEE Frontiers in Education Conference. Washington, DC.

Darwiche, A. (2009). *Modeling and reasoning with Bayesian networks.* New York, NY: Cambridge University Press.

Darwiche, A. (2010). Bayesian networks. *Communications of the ACM, 53*(12), 80-90.

Dawotola, A. W., van Gelder, P. H. A. J. M., & Vrijling, J. K. (2011). Decision analysis framework for risk management of crude oil pipeline system. *Advances in Decision Sciences, 2011*, 1-17.

DeWispelare, A. R., Herren, T. L., & Clemen, R. T. (1995). The use of probability elicitation in the high-level nuclear waste regulation program. *International Journal of Forecasting, 11*(1), 5-24.

European Aviation Safety Agency [EASA] (2011, April). *European action plan for the prevention of runway incursions: Edition 2.0*. Brussels, Belgium: EUROCONTROL.

Federal Aviation Administration [FAA] (2001). *FAA runway safety blueprint, 2002-2004*. Washington, DC: Federal Aviation Administration

Federal Aviation Administration [FAA] (2004a, August). *FAA runway safety report*. Washington, DC: Federal Aviation Administration.

Federal Aviation Administration [FAA] (2006). *Design and installation details for airport visual aids* (Advisory Circular 150/5340-30G). Washington, DC: U.S. Deaprtment of Transportation.

Federal Aviation Administration [FAA] (2007). *FAA adopts ICAO definition for runway incursions* (DOT/Federal Aviation Administration). Washington, D.C.: Department of Transportation.

Federal Aviation Administration [FAA] (2009a). Runway Safety - Northeast Region: Definitions. Retrieved October 2, 2012, from http://www.faa.gov/airports/runway_safety/ane/def/

Federal Aviation Administration [FAA] (2009b, November 20). Runway safety - Runway incursions. Retrieved June 4, 2012, from http://www.faa.gov/airports/ runway_safety/news/runway_incursions/

Federal Aviation Administration [FAA] (2010a). *Annual runway safety report 2010*. Washington, DC: Federal Aviation Administration.

Federal Aviation Administration [FAA] (2010b). Serious runway incursions cut in half for second straight year [Press release]. Retrieved July 17, 2012, from http://www.faa.gov/news/press_releases/news_story.cfm?newsId=11959&omniRss=press_releasesAoc

Federal Aviation Administration [FAA]. (2010c), *Safety Management Systems for Aviation Service Providers* (Advisory Circular 120-92A). Washington, DC: Government Printing Office.

Federal Aviation Administration [FAA] (2011a, March). NextGen implementation plan. Washington, DC: Author.

Federal Aviation Administration [FAA] (2011b, June 10). *Runway incursion prevention actions* (SAFO 11004). Washington, DC: U.S. Department of Transportation.

Federal Aviation Administration [FAA] (2011c, September). National runway safety plan 2012-2014. Retrieved September 10, 2012, from http://www.faa.gov/ airports/ runway_safety/news/ publications/ media/2011_ATO_ Safety_National_ Runway _Safety_Plan.pdf

Federal Aviation Administration [FAA] (2011d). *Runway Status Light Engineering Briefing* (Engineering Brief #64D). Washington, DC: U.S. Department of Transportation.

Federal Aviation Administration [FAA] (2012a). *FAA aerospace forecast: Fiscal years 2012-2032* (HQ-121545). Washington, DC: U.S. Department of Transportation.

Federal Aviation Administration [FAA] (2012b). *Parts 91 and 135 single pilot, flight school procedures during taxi operations* (Advisory Circular No: 91-73B). Washington, DC: U.S. Department of Transportation.

Federal Aviation Administration [FAA] (2012c). *Parts 91, 121, 125, and 135 flightcrew procedures during taxi operations* (Advisory Circular No: 120-74B). Washington, DC: U.S. Department of Transportation

French, S. (2011). Aggregating expert judgment. *Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales, 105*(1), 181-206.

Gamerman, D., & Lopes, H. F. (2006). *Markov chain Monte Carlo: Stochastic simulation for Bayesian inference* (2nd ed.). Boca Raton, FL: Chapman & Hall.

Garthwaite, P. H., Kadane, J. B., & O'Hagan, A. (2005). Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association, 100*(470), 680-701.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall.

Genest, C., & Zidek, J. V. (1986). Combining probability distributions: A critique and annotated bibliography. *Statistical Science, 1*(1), 114-135.

Gill, J. (2008). *Bayesian methods: A social and behavioral sciences approach* (2nd ed.). Boca Raton, FL: Chapman & Hall.

Goossens, L. H. J., Cooke, R. M., Hale, A. R., & Rodić-Wiersma, Lj. (2008). Fifteen years of expert judgment at TU Delft. *Safety Science, 46*(2), 234-244.

Goossens, L. H. J., & Harper, F. T. (1998). Joint EC/USNRC expert judgement driven radiological protection uncertainty analysis. *Journal of Radiological Protection, 18*(4), 249-264.

Government Accountability Office [GAO] (2008). *FAA has increased efforts to address runway incursions* (GAO-08-1169T). Washington, DC: U.S. Government Printing Office.

Government Accountability Office [GAO] (2011). *Enhanced oversight and improved availability of risk-based data could further improve safety* (GAO-12-24). Washington, DC: U.S. Government Printing Office.

Hanea, A. M., Kurowicka, D., & Cooke, R. M. (2006). Hybrid method for quantifying and analyzing Bayesian belief nets. *Quality and Reliability Engineering International, 22*(6), 709-729.

Hanninen, M., & Kujala, P. (2010). The effects of causation probability on the ship collision statistics in the Gulf of Finland. *International Journal on Marine Navigation and Safety of Sea Transportation, 4*(1), 79-84.

Hanninen, M., & Kujala, P. (2012). Influences of variables on ship collision probability in a Bayesian belief network model. *Reliability Engineering and System Safety,* (102), 27-40.

Heidegger, M., & Sadler, T. (1988). *The essence of truth: on Plato's parable of the cave and the Theaetetus.* New York, NY: Continuum.

Hendrickson, S. M. L. (2009). *The wrong Wright stuff: Mapping human error in aviation* (Doctoral dissertation, University of New Mexico).

Hoelzer, K., Oliver, H. F., Kohl, L. R., Hollingsworth, J., Wells, M. T., & Wiedmann, M. (2012). Structured expert elicitation about Listeria monocytogenes cross-contamination in the environment of retail deli operations in the United States. *Risk Analysis, 32*(7), 1139-1156.

HUGIN Expert A/S (n.d.). Case Stories. Retrieved November 8, 2012, from http://www.hugin.com/case-stories

International Civil Aviation Organization [ICAO] (2007). *Manual on the prevention of runway incursions* (Doc 9870/AN/463). Montreal, Québec, CA: International Civil Aviation Organization.

International Civil Aviation Organization [ICAO] (2009). *Safety management manual (SMM)*, 2nd ed. (Doc 9859 AN/474). Montreal, Canada: ICAO.

James, A., Low Choy, S., & Mengersen, K. (2010). Elicitator: An expert elicitation tool for regression in ecology. *Environmental Modelling & Software, 25*(1), 1219-145.

Jensen, F. V. (2009). Bayesian networks. *WIREs Computational Statistics, 1*(3), 307-315.

Jones, D. (2002). *Runway incursion prevention system evaluation*. Paper presented at the 21st Digital Avionics Systems Conference. Irvine, CA.

Jones, D. (2005). *Runway incursion prevention system testing at the Wallops Flight Facility*. Paper presented at the SPIE Defense & Security Symposium. Orlando, FL.

Jones, D., & Prinzel, L. J. (2006). *Runway incursion prevention for general aviation operations*. Paper presented at the 25th Digital Avionics Systems Conference. Portland, OR.

Joslin, R. E., Goodheart, B. J., & Tuccio, W. A. (2011). A mixed method approach to runway incursion rating. *International Journal of Applied Aviation Studies, 11*(2), 13-36.

Kaplan, S. (1992). 'Expert information' versus 'expert opinions.' Another approach to the problem of eliciting/combining/using expert knowledge in PRA. *Reliability Engineering & System Safety, 35*(1), 61-72.

Kelley, D. R., & Steinbacher, J. G. (1993). *Pilot surface incident safety study* (Report No. MTR-92000116). McLean, VA: The MITRE Corporation.

Kelley, D. R., & Jacobs, G. (1998). *Reports by airport traffic control tower controllers on airport surface operations* (Report No. MTR-98W0000033). McLean, VA: The MITRE Corporation.

Kelly, D. L., & Smith, C. L. (2009). Bayesian inference in probabilistic risk assessment: the current state of the art. *Reliability Engineering and System Safety, 94*(2), 628-643.

Kelly, D., & Smith, C. (2011). *Bayesian inference for probabilistic risk assessment: A practitioner's guidebook.* London, UK: Springer.

Kjaerulff, U. B., & Madsen, A. L. (2008). *Bayesian networks and influence diagrams: A guide to construction and analysis.* New York, NY: Springer.

Knol, A. B., Slottje, P., van der Sluijs, J. P., & Lebret, E. (2010). The use of expert elicitation in environmental health impact assessment: a seven step procedure. *Environmental Health, 9*, 19-34.

Knott, B., Gannon, A., & Rench, M. (2000). *Runway incursion: Human factors in runway incursions* (RA-2000-001A). WPAFB, OH: CSERIAC Program office.

Koenig, R. L. (1995). FAA report surveys U.S. airline pilots to discover factors that promote runway incursions. *Flight Safety Foundation Airport Operations, 21*(4), 1-8.

Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: Principles and techniques.* Cambridge, MA: MIT Press.

Kurowicka, D., & Cooke, R. M. (2006). *Uncertainty analysis with high dimensional dependence modelling.* West Sussex, UK: Wiley.

Kynn, M. (2006). Designing elicitor: software to graphically elicit expert priors for logistic regression models in ecology. Retrieved November 1, 2012, from http://www.citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.135.5924

Kynn, M. (2008). The ‘‘heuristics and biases’’ bias in expert elicitation. *Journal of the Royal Statistical Society, 171*(1), 239-264.

Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological Forecasting and Social Change, 73*(5), 467-482.

Lechner, K. W., & Luxhøj, J. T. (2005). Probabilistic causal modeling of risk factors contributing to runway collisions: Case studies. *Human Factors and Aerospace Safety, 5*(3), 185-216.

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2008). *SPSS for intermediate statistics: Use and interpretation* (3rd ed.). New York, NY: Taylor & Francis.

Levine, E. S. (2012). Estimating conditional probabilities of terrorist attacks: Modeling adversaries with uncertain value tradeoffs. *Risk Analysis, 32*(2), 294-303.

Lighttwist Software (n.d.). *UNINET help.* . Retrieved June 19, 2012, from http://www.lighttwist.net/wp/wp-content/doc/UninetHelp.pdf

Lin, S. W., & Bier, V. M. (2008). A study of expert overconfidence. *Reliability Engineering and System Safety, 93*(5), 711-721.

Low-Choy, S., & McBride, M. (2011, November). *Preconference tutorial: Expert elicitation methods for Bayesian networks*. Paper presented at the Australian Bayesian Network Modelling Society Conference. Brisbane, Australia.

Luxhøj, J. T. (2003). Probabilistic causal analysis for system safety risk assessments in commercial air transport. In *Proceedings of the Workshop on Investigating and Reporting of Incidents and Accidents (IRIA)* (pp. 17-38). Williamsburg, VA

Luxhøj, J. T., & Coit, D. W. (2006). *Modeling low probability/high consequence events: An aviation safety risk model*. Paper presented at the Reliability and Maintainability Symposium. Piscataway, NJ.

Lynch, S. M. (2007). *Introduction to applied Bayesian statistics and estimation for social scientists.* New York, NY: Springer.

Lynch, S. M. (2012, March). *Application of Bayesian statistics in social science*. Paper presented at the Applies Statistics Workshop on Bayesian Analysis. Long Beach, CA.

Marcot, B. G., Steventon, D., Sutherland, G. D., & McCann, R. K. (2006). Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. *Canadian Journal of Forest Research, 36*(12), 3063-3074.

Martin, T., Burgman, M., Fidler, F., Kuhnert, P., Low-Choy, S., MacBride, M., & Mengersen, K. (2012). Elicitation of expert knowledge in conservation science. *Conservation Biology, 26*(1), 29-38.

McLean, D., & Monro, I. (2004). Runway incursions and prevention systems. *Aircraft Engineering and Aerospace Technology, 76*(5), 472-478.

Moertl, P. M., & McGarry, K. (2011). *Using a cockpit display of traffic information with indicators and alerts to prevent runway incursions* (Case Number: 11-0965). McLean, VA: MITRE Corporation.

Morales, O., Kurowicka, D., & Roelen, A. (2008). Eliciting conditional and unconditional rank correlations from conditional probabilities. *Reliability Engineering & System Safety, 93*(5), 699-710.

Morales, O., Cooke, R. M., & Kurowicka, D. (2008, July). *EEMCS final report for the Causal modeling for Air Transport Safety (CATS) project*. Delft, The Netherlands: Delft Institute of Applied Mathematics.

Mosleh, A., Bier, V. M., & Apostolakis, G. (1988). A critique of current practice for the use of expert opinions in probabilistic risk assessment. *Reliability Engineering & System Safety, 20*, 63-85.

Mosleh, A., Dias, A., Eghbali, G., & Fazen, K. (2004). An integrated framework for identification, classification, and assessment of aviation systems hazard. In C. Spitzer, U. Schmocker, & V. Dang (Eds.), *Probabilistic Safety Assessment and*

Management (PSAM7) and European Safety and Reliability Conference (ESREL) (pp. 2384-2389). London, UK: Springer.

Nadkarni, S., & Shenoy, P. P. (2001). A Bayesian network approach to making inferences in causal maps. *A Bayesian network approach to making inferences in causal maps, 128*, 479-498.

Nadkarni, S., & Shenoy, P. P. (2004). A causal mapping approach to constructing Bayesian networks. *Decision Support Systems, 38*(2), 259-281.

Napoles, O. M. (2010). *Bayesian belief nets and vines in aviation safety and other applications* (Unpublished Dissertation). Delft University of Technology, Delft, The Netherlands.

National Aeronautics and Space Administration [NASA] (2002, August). Probabilistic risk assessment procedures guide for NASA managers and practitioners. Retrieved November 22, 2011, from http://www.hq.nasa.gov/office/codeq/doctree/preaguide.pdf

National Transportation Safety Board [NTSB] (1986). *Special investigation report: Runway incursions at controlled airports in the United States* (NTSB/SIR-86/01). Washington, D.C.: National Transportation Safety Board.

National Transportation Safety Board [NTSB] (1990). *Aircraft accident report* (NTSB/AAR-91/05). Washington, D.C.: National Transportation Safety Board.

Neri, A., Aspinall, W. P., Cioni, R., Bertagnini, A., Baxter, P. J., Zuccaro, G.,...Woo, G. (2008). Developing an Event Tree for probabilistic hazard and risk assessment at Vesuvius. *Journal of Volcanology and Geothermal Research, 178*(3), 397-415.

Ng, K. Y., & Earley, P. C. (2006). Culture+ Intelligence Old Constructs, New Frontiers. *Group & Organization Management, 31*(1), 4-19.

National Transportation Safety Board [NTSB] (n.d.). Most wanted list. Retrieved from http://www.ntsb.gov/safety/mwl.html

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J.,...Rakow, T. (2006). *Uncertain judgements: Eliciting expert probabilities.* West Sussex, UK: John Wiley & Sons.

O'Hagan, T., & Oakley, J. (2010, September 10). SHELF. Retrieved September 28, 2012, from http://tonyohagan.co.uk/shelf/

Ouchi, F. (2004, February). *A literature review on the use of expert opinion in probabilistic risk analysis* (World Bank Policy Research Working Paper 3201). Washington, D.C.: World Bank Group.

Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). New York, NY: Cambridge University Press.

Pearl, J., & Russell, S. (2003). Bayesian analysis. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks (*pp. 157-160). Cambridge, MA: MIT Press.

Qing, X. (2002). *Risk analysis for real estate investment* (Doctoral dissertation, TU Delft, Delft, The Netherlands).

Rankin II, W. B. (2007). *Runway incursions: A critical examination of airport driver training methods*. *Dissertation Abstracts International: Section A. Humanities and Social Sciences*.

Rankin II, W. B. (2008). Runway incursions: An industry examination of FAA initiatives and objectives. *International Journal of Applied Aviation Studies, 8*(2), 225-240.

Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Safety Science, 27*(2-3), 183-213.

Reason, J. (1997). *Managing the risks of organizational accidents.* Aldershot, UK: Ashgate.

Renooij, S. (2001). Probability elicitation for belief networks: Issues to consider. *The Knowledge Engineering Review, 16*(3), 255-269.

Renooij, S., & Witteman, C. L. M. (1999). Talking probabilities: communicating probabilistic information with words and numbers. *International Journal of Approximate Reasoning, 22*, 169-194.

Roelen, A. (2008). *Causal risk models of air transport: Comparison of user needs and model capabilities.* Amsterdam, The Netherlands: IOS Press.

Roelen, A. L. C., van Baren, G. B., Morales, O., & Krugla, K. (2008). *A generic maintenance technician performance model for application in a causal model of air transport* (NLR-CR-2008-445). Amsterdam, Netherlands: NLR.

Roelen, A., Wever, R., Hale, A., Goossens, L., Cooke, R., Lapuhaa, R.,...Valk, P. (2003). Causal modelling for integrated safety at airports. In *Proceedings of ESREL 2003 - European Safety and Reliability Conference, the Netherlands* (pp. 1321-1327).

Rubino, G., & Tuffin, B. (Eds.). (2009). *Rare event simulation using Monte Carlo methods.* Chichester, UK: Wiley.

Schönefeld, J., & Möller, D. P. F. (2012). Runway incursion prevention systems: A review of runway incursion avoidance and alerting system approaches. *Progress in Aerospace Science, 51*(May), 31-49.

Shonkwiler, R. W., & Mendivil, F. (2009). *Explorations in Monte Carlo methods.* New York, NY: Springer.

Siu, N. O., & Kelly, D. L. (1998). Bayesian parameter estimation in probabilistic risk assessment. *Reliability Engineering & System Safety, 62*(1-2), 89-116.

Speirs-Bridge, A., Fidler, F., McBride, M., Flander, L., Cumming, G., & Burgman, M. (2010). Reducing overconfidence in the interval judgments of experts. *Risk Analysis, 30*(3), 512-523.

Steinbacher, J. G. (1991). *An analysis of ATC-related runway incursions in the national airspace system* (WP-91W00234). McLean, VA: MITRE Corporation.

Stolzer, A. J., Halford, C. D., & Goglia, J. J. (2008). *Safety management systems in aviation.* Aldershot, UK: Ashgate.

Stroeve, S. H., Blom, H. A., & Bakker, G. J. (2013). Contrasting safety assessments of a runway incursion scenario: Event sequence analysis versus multi-agent dynamic risk modelling. *Reliability Engineering and System Safety, 109*(January 2013), 133-149.

Tarrel, R. J. (1985). *Non-airborne runway conflicts: the causes and effects of runway transgressions* (NASA Contractor Report 177372). Moffett Field, CA: Ames Research Center.

Torres, K. R., Metscher, D. S., & Smith, M. (2011). A correlational study of the relationship between human factor errors and the occurrence of runway incursions. *International Journal of Professional Aviation Training & Testing Research, 5*(1), 1-25.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4175), 1124-1131.

Tyshenko, M. G., El Saadany, S., Oraby, T., Darshan, S., Aspinall, W., Cooke, R.,...Krewski, D. (2011). Expert elicitation for the judgment of prion disease risk uncertainties. *Journal of Toxicology and Environmental Health, Part A: Current Issues, 74*(2-4), 261-285.

Valdés, R. M. A., Comendador, F. G., Gordún, L. M., & Nieto, F. J. S. (2011). The development of probabilistic models to estimate accident risk (due to runway overrun and landing undershoot) applicable to the design and construction of runway safety areas. *Safety Science, 49*(5), 633-650.

Van Der Fels-Klerx, H. J., Cooke, R. M., Nauta, M. N., Goossens, L. H. J., Havelaar, A. H. (2005). A structured expert judgment study for model of Campylobacter transmission during broiler chicken processing. *Risk Analysis, 25*(1), 109-124.

Vernaleken, C., Urvoy, C., & Klingauf, U. (2007). Prevention of Runway Incursions due to closed runways or unsuitable runway choices by enhanced crew situational awareness and alerting. In J. G. Gerly & J. J. Guell (Eds.), *Enhanced and Synthetic Vision.* Orlando, FL: SPIE.

Wang, C. (2007). *Hybrid causal logic methodology for risk assessment* (Doctoral dissertation). Available from ProQuest Dissertations and Theses. (3297344)

Winkler, R. L., Wallsten, T. S., Whitfield, R. G., Righmond, H. M., Hayes, S. R., & Rosenbaum, A. S. (1995). An assessment of the risk of chronic lung injury attributable to long-term ozone exposure. *Operations Research, 43*(1), 19-27.

Wood, L. E., & Ford, J. M. (1993). Structuring interviews with experts during knowledge elicitation. In K. M. Ford & J. M. Bradshaw (Eds.), *Knowledge acquisition as modeling (*Vol. 1, pp. 71-90). New York, NY: John Wiley & Sons.

Wu, J. S., Apostolakis, G., & Okrent, D. (1990). Uncertainties in system analysis: Probabilistic versus nonprobabilistic theories. *Reliability Engineering & System Safety*, 30, 163-181.

Zio, E. (2009). Reliability engineering: Old problems and new challenges. *Reliability Engineering and System Safety, 94*(2), 125-141.

Zuschlag, M. (2005, September). *Violations of Temporary Flight Restrictions and Air Defense Identification Zones: An analysis of airspace violations and pilot report data* (NASA/CR—2005-213923). Washington, D.C.: National Aeronautics and Space Administration.

APPENDIX A

Operational Example of a Bayesian Belief Network

Suppose we wish to evaluate the probability of an unintentional runway overrun in a particular type of aircraft utilized frequently by a theoretical flight department. In traditional PRA, a series of contributing factors to runway overruns may be identified through the study of past accidents, hazard reports and through elicitation of expert opinion. Considering these pieces of information, an estimate of the risk of overrun may be made. However, this estimate is first deterministic, and second addresses the risk of an undesired event only as a point estimate with no accommodation for uncertainty. Furthermore, this method, which has arguably been oversimplified for the purpose of discussion here, ignores the inherent suitability of the problem to Bayesian methods. This is evident in the monotonic behavior of deductive reasoning as is generally used in the formulation of estimation of risk. In a deductive logic-based system, there is no provision for dynamically asserting or retracting assumptions as knowledge of the domain changes. The conditional nature of contributing factors and their interaction is generally acknowledged, but is ignored in simplified computations of risk. In a Bayesian model-based system, revisable degrees of belief account for the ever-changing state of domain knowledge and allow a much more commonsense approach to reasoning about the outcome of a postulated risk. It is beyond the scope of this example to create a full model, much less one that has been populated with probabilistic distributions or that addresses the full range of contributing factors. Instead, the following discussion addresses the basic process of model creation as an example of the power of Bayesian models for probabilistic inference. As in traditional PRA, contributing factors to an

Table 15. *Example Contributing Factors to Runway Overrun.*

Contributing Factors
Raining
Snowing
Low visibility
Brake system malfunction
Auto braking disengaged
Brake over-temp
A/C over weight
Wildlife on runway

undesired event must be identified. For the purpose of this discussion, an abbreviated list of such factors is provided as Table 15.

After identifying contributing or causal factors, conditional probability tables must be created for each node in the network. It is helpful to first create a directed acyclic graph of the network in order to identify child-parent dependencies. One may find that it is useful to begin with the node being assessed and work backward to develop parent nodes and dependencies as appropriate. In Figure 37, a simple Bayesian network is presented for selected factors identified in Table 15. Generally, useful BBNs will have many more nodes and may have multiple-level parent-child dependencies; however, the model in Figure 19 is sufficient for demonstrating the basic principles of the analysis. Worth noting here is that if the entire sequence of potential contributing factors to the example event were to be modeled in any other way, the result would likely be a computationally impossible problem. However, as was previously discussed, BBNs circumvent this issue and can model the problem compactly through the assumption that each variable becomes independent of its non-descendants once the value of the parent node is known (Darwiche, 2010).

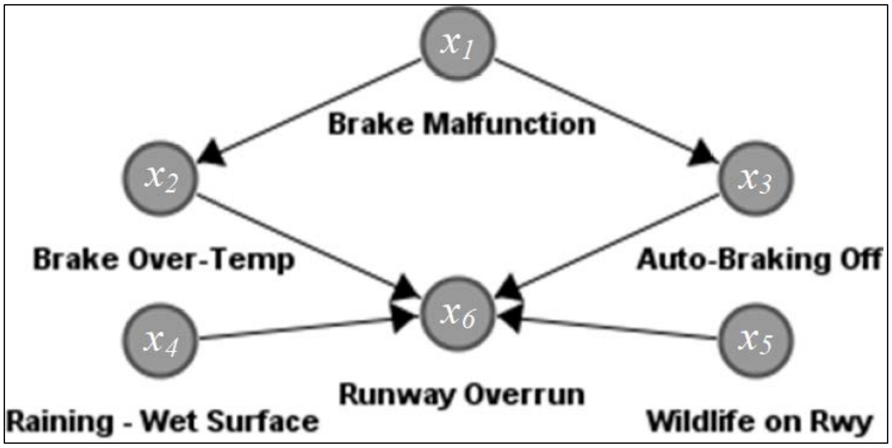


Figure 37. BBN assessing the probability of a runway overrun.

This network graph, as would any well-advised risk assessment, relies on verification of the model, generally via domain experts who can verify that the conditional independencies and influence paths in the diagram are valid assumptions. Where no consensus on node relationship or interdependency can be reached, the graph and conditional interdependencies must be revised. It is beyond the scope of this literature review to discuss the mechanics of this procedure beyond describing the broad idea that implicit dependencies can be assessed by creating a moral graph, as shown in

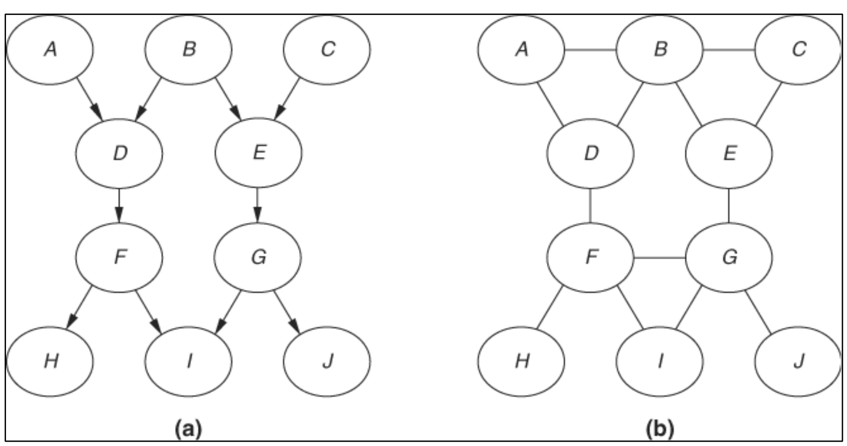


Figure 38. BBN (a) and its associated moral graph (b) (Jensen, 2009).

Figure 38, to accompany the belief net, and adjustments may be made for the imperfect way influence paths represent conditional independence. For more detailed descriptions of this process, readers may consult the Methodology section to follow as well as Darwiche (2009), Bedford and Cooke (2001), Kelly and Smith (2011), or Jensen (2009).

Returning to the network graph in Figure 37, the computational advantage of the BBN approach to the model becomes clearer upon examination of the joint distribution, which is specified in full as:

$$P(x_1, \ldots, x_n) = \prod_i P\,(x_1|pa_1) \tag{8}$$

where:

P is probability;

$(x_1, \ldots, x_n)$ is the values for some variables X_1 to X_n;

$\prod_i P$ is the product notation for the conditional distribution; and

pa_1 is the set of values for the parents of X_1.

And as specified in the Figure 37 example as:

$$P(x_1, x_2, x_3, x_4, x_5, x_6) = P(x_1)\,P(x_2|x_1)\,P(x_3|x_1)\,P(x_6|x_2, x_3)\,P(x_6|x_4)\,P(x_6|x_5) \tag{9}$$

In this example, the computational compactness previously discussed becomes apparent, as the number of parameters required increases in a linear fashion, whereas the joint distribution grows exponentially (Pearl & Russell, 2003). Also discussed earlier is the property of Bayesian network graphs such that independence is maintained only for parent nodes as nondescendants are eliminated as in Equation 10:

$$P(x_6|x_1, x_2, x_3) = P(x_6|x_2, x_3) \tag{10}$$

Assuming creation of a network structure and appropriate validation is complete; a practitioner must assign probabilistic values to each node. This may be accomplished by way of conditional probability tables where some value of prior probability is known, or by a probability distribution function from which values can be sampled in the case of uncertainty. Each method is reviewed briefly here to illustrate the basic principles, and readers are directed to Gamerman and Lopes (2006), Kelly and Smith (2011), Bedford and Cooke (2001), and Darwiche (2009) for more detailed information.

In the model presently discussed, there are six nodes. For each node, the states must be defined, and conditional probability given parent nodes must be established. As an example, in Node 1, *brake malfunction* takes the states *yes* and *no*. Node 2 has the states *yes* or *no* as does Node 3, and so on. Tables 16 and 17 illustrate how conditional probability may be presented as a conditional probability table (CPT). It is important to note that the probability estimates in these tables appear as point estimates; however, a probability distribution accounting for uncertainty could be framed similarly. Figure 39 appropriately presents a visualization of the probability distribution of runway stopping location.

Table 16. *Conditional Probability of Brake Over-temp Given Brake Malfunction.*

Brake Malfunction	Brake Over-temp	$\Theta_{X2\|X1}$
Yes	Yes	.50
No	Yes	.01

Note: $\Theta_{x2|x1}$ is CPT for variable X_2 and its parent, X_1 in this and following CPTs.

Table 17. *Conditional Probability of Runway Overrun given x_2, x_3.*

Brake Over-temp	Auto-braking Off	Runway Overrun	$\Theta_{x6\|x2, x3}$
Yes	Yes	Yes	.1200
Yes	No	Yes	.0010
No	Yes	Yes	.0004
No	No	Yes	$1*10^{-6}$

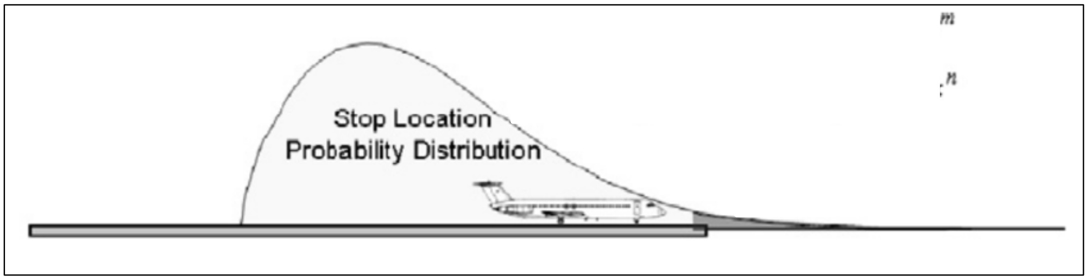


Figure 39. Probability distribution of runway stopping distance (Valdes et al., 2011).

Though the example suggested here is brief by necessity, it illustrates the intuitive design of BBNs in aviation PRA settings, where the development of event sequences and fault trees is commonplace. From a practitioner's point of view, the process must be undertaken in much the same way as the building of a database or even the design of an SMS. That is, the bulk of the effort is confined to the design phases of the model, and with modern software and hardware developments, the model can be run and evaluated with relative ease once the necessary framework has been laid.

APPENDIX B

Human Subject Protocol Application Form

INSTRUCTIONS

Please answer the **10** questions on the ***Human Subject Protocol Application Form*** as completely and thoroughly as possible. Include more lines where necessary. Upon completion, email to hollerat@erau.edu at the Pre-Award office.

Include any supporting documentation along with a complete copy of your ***Informed-Consent Form***, any other tests, instruments, or surveys, as well as any proposal for funding. It is incumbent upon the researcher to demonstrate that the Principal Investigator is qualified to perform the study, every possible step has been taken to reduce risk to the participants, and that adequate benefit will come from the study to offset the risks.

The answers to the questions need not be long, but they should be sufficiently detailed so that the reviewer can accurately assess the risks and benefits associated with your study.

Embry-Riddle Aeronautical University
Human Subject Protocol Application Form

Project Title:
____

Principal Investigator:
____
(If student, list advisor's name as investigator)

List all Other Investigators:

____

____

____

Submission Date: ______________________

Beginning Date: ____________________ Expected End Date: _______________

Type of Project:
____

Type of Funding Support (if any):
____

Please answer the following questions and provide a brief explanation of the answer for each. Include more lines where necessary.

1. Briefly describe the background and purpose of the research.

   ____

2. Describe in detail each condition or manipulation to be included with the study.

   ____

3. What measures or observations will be taken in the study? If any questionnaires, tests, or other instruments are used, provide a brief description and include a copy for review (computer programs may require demonstration at the request of the IRB).

____

4. Describe the possible risks and benefits (if any) to the subjects and describe how the experimental design will limit risks.

____

5. Describe the methods to be used in securing the informed consent of the subjects. If an informed consent form is to be used, attach to this form. See Informed Consent information sheet for more information on Informed Consent requirements.

____

6. Will participant information be anonymous (not even the researcher can match data with names), confidential (names or any other identifying demographics can be matched, but only members of the research team will have access to that information. Publication of the data will not include any identifying information), or public (names and data will be matched and individuals outside of the research team will have either direct or indirect access. Publication of the data will allow either directly or indirectly, identification of the participants). Justify the classification and describe how privacy will be ensured/protected.

____

7. If video/audio recordings are part of the research, please describe how that data will be stored or destroyed.

____

8. Are students being required to participate in this research as part of a class project or as a class assignment? What are the alternatives to be offered in the event that a student(s) choose not to participate? If so, please list the class(es) and faculty members involved and justify this situation in light of APA ethical guidelines of the APA Publication Manual.

____

9. Are participants going to be paid for their participation? If yes, describe your policy for dealing with participants who 1) Show up for research, but refuse informed consent; 2) Start but fail to complete research.

___

10. Approximately how much time will be required of each participant?

___

APPENDIX C

Cooke's Classical Model: Elicitation and Aggregation

Within the classical model, two quantitative measures of expert performance exist: *information* and *calibration*. Whereas information is a measure of the concentration of the distributed expert opinions, calibration is a measure of the likelihood that s set of experimental results would correspond to those assessed by the experts (Cooke & Goossens, 2006). In determining calibration, Cooke (2009) suggests that each expert be asked to provide 5%, 50%, and 95% values. The range for each quantity is divided into four inter-quartile intervals such that p_1 = 0.05: less than or equal to the 5% value, p_2 = 0.45: greater than the 5% value and less than or equal to the 50% value, p_3 = 0.45: greater than the 50% value and less than or equal to the 95% value, and p_4 = 0.5: greater than the 95% value and less than or equal to the 100% value:

p= (0.05, 0.45, 0.45, 0.05)

$s_1(e) = \#\{\, i \mid x_i \leq 5\% \text{ quantile}\}/N$
$s_2(e) = \#\{\, i \mid 5\% \text{ quantile} < x_i \leq 50\% \text{ quantile}\}/N$
$s_3(e) = \#\{\, i \mid 50\% \text{ quantile} < x_i \leq 95\% \text{ quantile}\}/N$
$s_4(e) = \#\{\, i \mid 95\% \text{ quantile} < x_i \,\}/N$
$s(e) = (s_1, \ldots s_4)$

Cooke (1991) notes that the sample distribution depends on expert *e* and assuming independent draws from a distribution with the quantiles described previously by the expert is a chi-square test statistic for goodness of fit with three degrees of freedom as in the following, in which $I(s(e)|p)$ denotes relative information of distribution *s* with respect to *p*for expert *e*:

$$2NI(s(e)|p) = 2N \sum_{i=1..4} s_i \, ln(s_i/p_i) \qquad (6)$$

where:

e is the expert;

s is the distribution;

p is probability of concurrence

The decision maker in the classical model scores each expert e as the likelihood of the hypothesis, "the inter-quantile interval containing the true value for each variable is drawn independently from probability vector p" (Cooke, 2009). The calibration score is calculated by:

$$\text{Calibration score}(e) = p\text{-}value = \text{Prob}\{2NI(s(\text{e})|\text{p}) \geq r \mid H_e\}$$

where:

r is the value from (6) based on the observed values of $x_1 \ldots x_N$;

H_e is the hypothesis that a deviation at least as large as r is observed on N realizations id the hypothesis is true; and

s is the sample distribution.

The information element of the expert score weighting requires a density be associated with each of the quantile assessments gained from the experts (and discussed in the preceding section). The classical model offered by Cooke (2009) uses the k% overshoot rule: "for each item we consider the smallest interval $I = [L, U]$ containing all the assessed quantiles of all experts and the realization, if known" (p. 265). Cooke extends this interval to:

$$I^* = [L^*, U^*];\ L^* = L - k(U\text{-}L)/100;\ \ U^* = U + k(U\text{-}L)/100. \tag{7}$$

where:

I is the interval;

U is the upper interval limit;

L is the lower limit;

k is the overshoot value chosen by the researcher.

In the discussion that follows, mention is made of the role that analyst experience plays in the quality of the weighting process. In Equation 7, the value of k is chosen by the researcher, and Cooke notes that a large value tends to "suppress the relative differences in information scores" (2009, p. 265). With this in mind, the process continues, with the information score calculated for each rater as by:

Inf (e) =average relative information with respect to background = $(1/N) \sum_{\mathrm{i}} = 1.._{N} I(f_{e,i}|g_i)$

where:

g_i is the background density for I and $f_{e,i}$ is expert e's density function for item i.

The combination of the information and calibration scores serves as a weighting mechanism for each expert and is dependent upon the value of α, as shown in Equation 8 (Cooke, 1991, 2009).

$$w_{\alpha}(\mathrm{e}) = \mathrm{Cal}\ (\mathrm{e}) \times \mathrm{Inf}\ (\mathrm{e}) \times 1_{\alpha}(\mathrm{Cal(e)} \geq \alpha) \qquad (8)$$

The classical model, and the evaluation of the quality of expertise is essentially a function of linear pooling and the derivation of weights as a product of information and calibration. To avoid what Cooke and Goossens (2006) refer to as "haphazard" influence on the decision maker by an individual expert, a set of scoring rules is also imposed within the classical model (p. 7). Performance of the weighting scheme (more information on the details of which can be found in Cooke (2004 & 2009) and Cooke & Goosens (2006)) is largely dependent on the experience of the researcher in practical use (Cooke & Goossens, 2006). As such, the research proposed herein must account for bias and performance issues in this regard, and it is likely that this limitation will propagate through the model results, creating a limitation to the generalizability of the data. To attempt to limit this potential shortcoming, the software package *EXCALIBUR*, which accomodates the elements of the classical model and will calculate information and calibration scores to assign weights after the elicitation results are inserted, can be used.

Structured expert judgment requires:

- Scrutability/accountability: All data, including experts' names and assessments, and all processing tools are open to peer review and results must be reproducible by competent reviewers.
- Empirical control: Quantitative expert assessments are subjected to empirical quality controls.
- Neutrality: The method for combining/evaluating expert opinion should encourage experts to state their true opinions, and must not bias results.
- Fairness: Experts are not pre-judged, prior to processing the results of their assessments (Cooke & Goossens, 2006, p. 3).

APPENDIX D

Structured Expert Judgment Subject Matter Expert Profiles

Expert A is a Captain for a 14 CFR Part 121 regional airline operator. He has extensive line pilot experience in airline operations at large primary commercial airports. He has over 7700 hours of flight time and 18 years of aviation experience. His aviation background includes 14 CFR Part 135 cargo operations and 14 CFR Part 141 flight instruction. He also holds a bachelor's degree in Aviation Science and a minor in Aviation Safety.

Expert B Is a Designated Pilot Examiner in Northern California. He conducts check rides for Private, Instrument, Commercial, ATP, and instructor licenses and certificates. He is also a Chief Pilot for an International Charter Operator and has extensive background in developing and writing manuals and training courses. He has 20 years of aviation experience with over 5500 flight hours. He instructs actively as well as currently flies multiple types of aircraft in 14 CFR Part 91 and 14 CFR Part 135 operations. His background also includes intensive operations in 14CFR91 Subpart K.

Expert C is the Operations Training Supervisor for a large hub airport where one of his duties is overseeing the 14 CFR Part 139.329 driver certification program at the airport. He has worked in airport operations for the last 7 years and in aviation training for the last 12. He has participated in a number of Runway Safety Action Teams and has been a Subject Matter Expert in FAA SRM panels on airfield safety. He also has been the lead airport representative in Airport/FAA study focusing on the airfield driver human factors.

He has presented on Runway Safety both in the US and Internationally. He is a CFI/CFII and holds a MEL pilot certificate.

Expert D is an airline captain for a large regional carrier in the United States. He is also the Safety Management System (SMS) Program Manager tasked with development and implementation of the SMS within his organization. He has an extensive background in safety management, incident/accident investigation, and human factors. He has over 4,300 hours of flight experience and was previously a training center evaluator (TCE) for a major flight university conducting practical tests for Private, Instrument, Commercial, and instructor certificates.

Expert E has been involved in military and civilian aviation for 45 years, serving as a pilot and engineer prior to joining the FAA in 1985. He acted within the FAA as an airport engineer, program and technical manager, and as a Region Runway Safety Manager as well as the Acting Director for Runway Safety at the national level.

Expert F has been working in the Air Traffic Control and Airport Management field since 1984. During the past 28 years, he has managed the research, development, and implementation of several national aviation projects for the FAA. He is the inventor of several new tools for pilots, airports and aircraft worldwide. He has instructed internationally on aviation related materials and is in high demand for this field. He has been instrumental in designing and redesigning airport layout plans. Many new enhancements to the XXXXX International Airport can be attributed to his fuel savings/efficiency modifications.

APPENDIX E

Structured Expert Judgment Elicitation Protocol

Structured Expert Judgment
Elicitation Protocol

for

IDENTIFICATION OF CAUSAL PATHS AND PREDICTION OF RUNWAY INCURSION RISK
USING BAYESIAN BELIEF NETWORKS

by

Benjamin Jeffry Goodheart

Embry-Riddle Aeronautical University
Daytona Beach, Florida

Transportation Research Board
Airport Cooperative Research Project
Washington, D.C.

July 2, 2013

Experts,

Thank you for your participation in this study, which attempts to probabilistically model the factors that contribute to runway incursions in the U.S.

In quantifying the probability distribution of the variables in the network model developed in this study, structured expert judgment is used, following the Classical Model developed by Cooke (1991).

The questions that follow relate to aviation operations in the United States, and are intended to include ground, ATC, and flight operations as appropriate. U.S. aviation metrics are generally expressed per 100,000 operations, and unless otherwise noted, this rate may be assumed throughout.

Throughout the following exercise, you will be asked to provide our estimate of the 5th, 50th, and 95th percentile of a particular measure. As a brief review, these percentiles correspond to a probability density function as in the figure below, and describe the bounds of a distribution (not necessarily normally distributed as shown below) using the basic concepts of:

- The 50th percentile (median) of the distribution, i.e. given 100 samples of a variable value, 50 would be expected to fall below and 50 would fall above the median value.
- The 5th percentile value, which can be interpreted as: it would surprise you if more than 5 out of 100 samples have a value *lower* than this value.
- The 95th percentile value, which can be interpreted as: it would surprise you if more than 5 out of 100 samples have a value *higher* than this value.

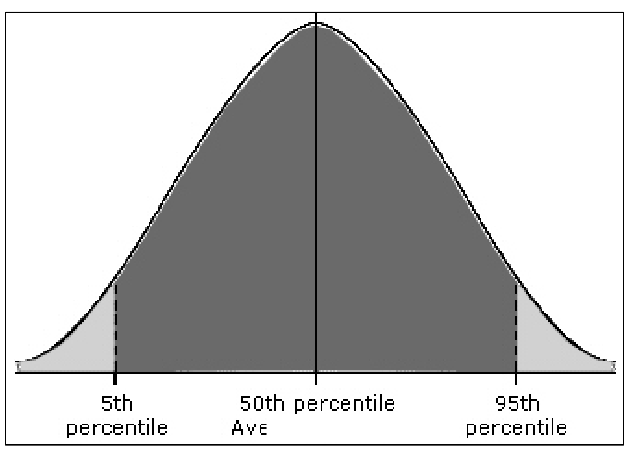


Figure 1. Normal distribution with percentiles.

Included in the elicitation are variables whose actual values are known. These variables are used to assist in measurement and validation of expert performance in quantifying uncertainty. A good assessment of uncertainty is statistically accurate and informative, two characteristics that are evaluated relative to the known variables and expert responses.

Use the accompanying spreadsheet, which describes the model variables and provides additional information, to support your responses to the following questions.

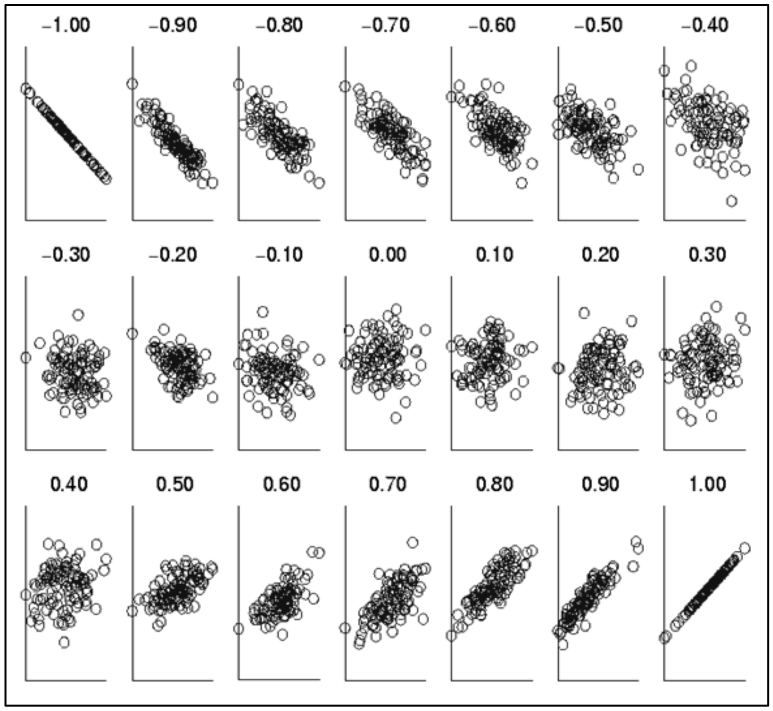
To understand how probabilities and dependencies will be assessed from your responses, consider the following example. If we assume a population of airline flight crew employees having a median age is 45 years old, a median gross salary of $90,000, and a median experience on three different aircraft within the current fleet, we can perform some basic estimates. If we look at the portion of the population that lies below the median, that is younger than 45, we can reasonably expect their mean salary to reflect a value lower than the median for the whole population. An appropriate question might be: Suppose we have 1,000 employees who are younger than 45 years old, how many of those would have an annual salary of less than $90,000. If your answer is 700, then the probability is expressed as:

$$P(\text{annual salary} \leq x_{50} = \$90{,}000 \mid \text{pilot age} \leq y_{50} = 45) = 0.70$$

Investigating further, we may want to look at the probability that the crewmember's salary is less than the median value given that the pilot is younger than 45 and has experience on fewer than three aircraft types, the median experience level given above. We would now express that probability as (where X, Y, and Z represent flight crew salary, age, and experience, respectively):

$$P(X \leq x_{50} \mid Y \leq y_{50}, Z \leq z_{50})$$

As part of the process of ranking variable interaction, you will be asked to express your opinion in terms of correlation. Recall that correlation can be positive or negative, and can vary in strength of association, as shown here:


−1.00
−0.90
−0.80
−0.70
−0.60
−0.50
−0.40
−0.30
−0.20
−0.10
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00

Expert Profiles

Expert A is a Captain for a 14CFR121 regional airline operator. He has extensive line pilot experience in airline operations at large primary commercial airports. He has over 7700 hours of flight time and 18 years of aviation experience. His aviation background includes 14CFR135 cargo operations and 14CFR141 flight instruction. He also holds a bachelor’s degree in Aviation Science and a minor in Aviation Safety.

Expert B Is a Designated Pilot Examiner in Northern California. He conducts check rides for Private, Instrument, Commercial, ATP, and instructor licenses and certificates. He is also a Chief Pilot for an International Charter Operator and has extensive background in developing and writing manuals and training courses. He has 20 years of aviation experience with over 5500 flight hours. He instructs actively as well as currently flies multiple types of aircraft in 14CFR91 and 14CFR135 operations. His background also includes intensive operations in 14CFR91 Subpart K.

Expert C is the Operations Training Supervisor for a large hub airport where one of his duties is overseeing the Part 139.329 driver certification program at the airport. He has worked in airport operations for the last 7 years and in aviation training for the last 12. He has participated in a number of Runway Safety Action Teams and has been a Subject Matter Expert in FAA SRM panels on airfield safety. He also has been the lead airport representative in Airport/FAA study focusing on the airfield driver human factors. He has presented on Runway Safety both in the US and Internationally. He is a CFI/CFII and holds a MEL pilot certificate.

Expert D is an airline captain for a large regional carrier in the United States. He is also the Safety Management System (SMS) Program Manager tasked with development and implementation of the SMS within his organization. He has an extensive background in safety management, incident/accident investigation, and human factors. He has over 4,300 hours of flight experience and was previously a training center evaluator (TCE) for a major flight university conducting practical tests for Private, Instrument, Commercial, and instructor certificates.

Expert E has been involved in military and civilian aviation for 45 years, serving as a pilot and engineer prior to joining the FAA in 1985. He acted within the FAA as an airport engineer, program and technical manager, and as a Region Runway Safety Manager as well as the Acting Director for Runway Safety at the national level.

Expert F has been working in the Air Traffic Control and Airport Management field since 1984. During the past 28 years, he has managed the research, development, and implementation of several national aviation projects for the FAA. He is the inventor of several new tools for pilots, airports and aircraft worldwide. He has instructed internationally on aviation related materials and is in high demand for this field. He has been instrumental in designing and redesigning airport layout plans. Many new enhancements to the XXXXX International Airport can be attributed to his fuel savings/efficiency modifications.

Calibration Questions

CQ1	**General**	
During the period between 1988 and 1999, the total number of runway incursions reported in the U.S. increased by what percent? (Not limited to 100% maximum)		
Answer: 171%		
Source: NASA ASRS Callback Summary Issue No. 263 (http://asrs.arc.nasa.gov/publications/ callback/cb_253.htm)		
5th percentile	**50th percentile**	**95th percentile**

CQ2	**General**	
During the period between 1988 and 1999, the number of U.S. runway incursions resulting from pilot deviations (PD) increased by what percent? (Not limited to 100% maximum)		
Answer: 267%		
Source: NASA ASRS Callback Summary Issue No. 263 (http://asrs.arc.nasa.gov/publications/ callback/cb_253.htm)		
5th percentile	**50th percentile**	**95th percentile**

CQ3	**General**	
How many airline departures per year occurred in the U.S. during the period from April 1, 2012 through March 31, 2013?		
Answer: 8,796,000		
Source: Bureau of Transportation Statistics (http://www.transtats.bts.gov/)		
5th percentile	**50th percentile**	**95th percentile**

CQ4	**General**	
What was the reported General Aviation accident rate for FAA FY2011 expressed as accidents per 100,000 flight hours?		
Answer: 6.51		
Source: NTSB 2011 Annual Aviation Safety Statistics (http://www.ntsb.gov/news/ 2012/120427.html)		
5th percentile	**50th percentile**	**95th percentile**

CQ5	General	
For the same period in CQ4, what was the reported Part 121 airline accident rate expressed as accidents per 100,000 flight hours?		
Answer: 0.175		
Source: NTSB (http://www.ntsb.gov/data/table5_2012.html)		
5th percentile	**50th percentile**	**95th percentile**

CQ6	General	
The FAA identifies airport "hot spots" as "a location on an airport movement area with a history of potential risk of collision or runway incursion, and where heightened attention by pilots and drivers is necessary." Considering there are approximately 5,170 public-use airports (503 with Part 121 airline service), how many hot spots have been identified		
Answer: 601		
Source: FAA (http://www.faa.gov/airports/runway_safety/hotspots/hotspots_list/)		
5th percentile	**50th percentile**	**95th percentile**

CQ7	General	
How many Runway Incursions were reported in the U.S in the period from January 1, 2013 through June 30, 2013?		
Answer: 569		
Source: FAA Office of Runway Safety (http://www.faa.gov/airports/runway_safety/statistics/year/?fy1=2013&fy2=2012)		
5th percentile	**50th percentile**	**95th percentile**

CQ8	General	
Of the Runway Incursions reported in CQ7, what percentage of these were pilot deviation events?		
Answer: 62%		
Source: FAA Office of Runway Safety (http://www.faa.gov/airports/runway_safety/statistics/year/?fy1=2013&fy2=2012)		
5th percentile	**50th percentile**	**95th percentile**

CQ9	General	
Of the Runway Incursions reported in CQ7, what percentage of these were vehicle/pedestrian deviation events?		
Answer: 20%		
Source: FAA Office of Runway Safety (http://www.faa.gov/airports/runway_safety/ statistics/year/?fy1=2013&fy2=2012)		
5th percentile	**50th percentile**	**95th percentile**

CQ10	General	
What percentage of aviation accidents may be attributed, at least in part, to human error?		
Answer: 70%		
Source: Shappell, et al, 2005 (http://www.hf.faa.gov/docs/508/docs/gaHFACS2005.pdf)		
5th percentile	**50th percentile**	**95th percentile**

Elicitation Questions

Q1	**Procedural Deviation**	
Consider 100,000 randomly chosen flights within the U.S. under the general model conditions. On how many of these flights will a *DEVIATION FROM PROCEDURE* be committed by the flight crew? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q2	**Inadequate Supervision**	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. In what proportion these operations might *INADEQUATE SUPERVISION*, as defined in the materials provided, be observed? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q3	**High Workload**	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. What percentage of these operations involves *HIGH WORKLOAD* as defined in the materials provided? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimate.		
5th percentile	**50th percentile**	**95th percentile**

Q4	**Organizational/Regulatory Factor**	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. What percentage of these operations will experience an abnormal *ORGANIZATIONAL/REGULATORY FACTOR* as defined in the materials provided? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimate.		
5th percentile	**50th percentile**	**95th percentile**

Q5	**Complex Intersection**	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. What proportion of these operations involves navigating an airport intersection defined as a *COMPLEX INTERSECTION*? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q6	Irregular or Noncompliant Signs or Markings	
Consider 100,000 randomly chosen operations within the U.S. under the general model conditions. What proportion of these operations experiences IRREGULAR OR NONCOMPLIANT SIGNS OR MARKINGS during movement to or from the runway area? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q7	Airport Construction	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. On how many of these operations might an operator encounter *AIRFIELD CONSTRUCTION IN PROGRESS*, as defined in the materials provided? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q8	Part 139	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. How many of these operations might utilize a *14CFR PART 139 AIRFIELD*? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q9	Non-Towered Airfield	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. How many of these operations might utilize an *NON-TOWERED AIRFIELD*? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q10	Conflicting Vehicle Traffic	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. What proportion of these operations might experience *CONFLICTING VEHICLE TRAFFIC*, as defined in the materials provided? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q11	Operational Environment Factor	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. What percentage of these operations will experience an abnormal *OPERATIONAL ENVIRONMENT FACTOR* as defined in the materials provided? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimate.		
5th percentile	**50th percentile**	**95th percentile**

Q12	Task Saturation	
Consider 100,000 randomly chosen aviation operations within the U.S. under the general model conditions. On how many of these operations might *TASK SATURATION*, as defined in the materials provided, be observed? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q13	Failure to Hold Short	
Consider 100,000 randomly chosen flight/airport vehicle operations within the U.S. under the general model conditions. On how many of these operations might a *FAILURE TO HOLD SHORT*, as defined in the materials provided, be observed? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q14	Lost Situational Awareness	
Consider 100,000 randomly chosen operations within the U.S. under the general model conditions. On how many of these operations might a *LOSS OF SITUATIONAL AWARENESS*, as defined in the materials provided, be observed? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q15	Human Factors	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. What percentage of these operations will experience an abnormal *HUMAN FACTOR* as defined in the materials provided? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimate.		
5th percentile	**50th percentile**	**95th percentile**

Q16	Sun Glare	
Consider 100,000 randomly chosen operations within the U.S. under the general model conditions. On how many of these operations might *SUN GLARE*, as defined in the materials provided, affect the operator(s)? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q17	Weather Factors	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. What percentage of these operations will experience an abnormal *WEATHER FACTOR* as defined in the materials provided? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimate.		
5th percentile	**50th percentile**	**95th percentile**
Q18	**ATC Equipment Failure**	
Consider 100,000 randomly chosen operations within the U.S. under the general model conditions. On how many of these operations might an *ATC EQUIPMENT FAILURE*, as defined in the materials provided, occur? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q19	In-Vehicle Display	
Consider 100,000 randomly chosen operations within the U.S. under the general model conditions. On how many of these operations might an *IN-VEHICLE DISPLAY*, as defined in the materials provided, be used by an operator(s)? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q20	Radio Congestion	
Consider 100,000 randomly chosen operations within the U.S. under the general model conditions. On how many of these operations might *RADIO CONGESTION*, as defined in the materials provided, occur? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimated distribution.		
5th percentile	**50th percentile**	**95th percentile**

Q21	Technical/Engineering Factors	
Consider 100,000 randomly chosen flight operations within the U.S. under the general model conditions. What percentage of these operations will experience an abnormal *TECHNICAL/ENGINEERING FACTOR* as defined in the materials provided? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimate.		
5th percentile	**50th percentile**	**95th percentile**

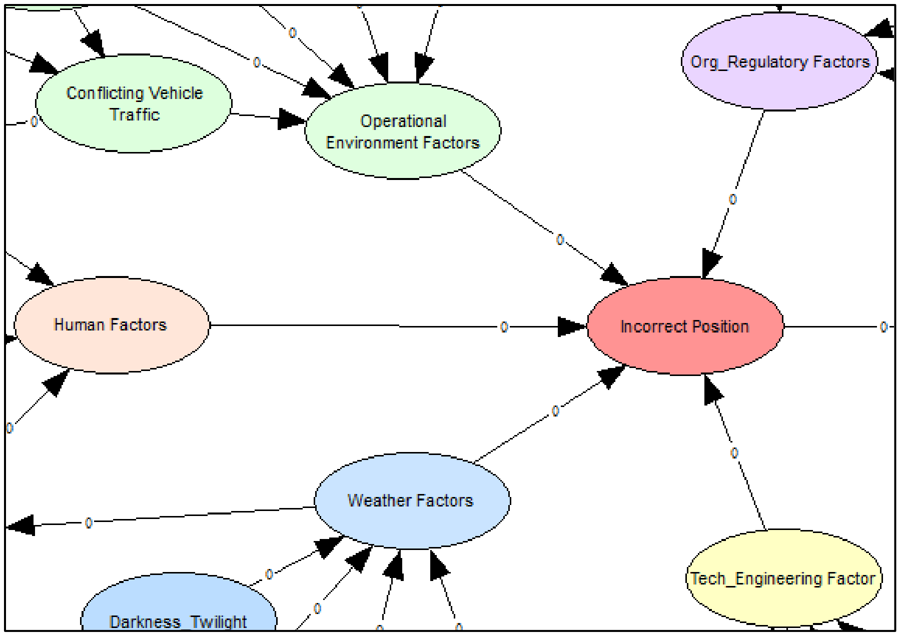
Q22	Incorrect Presence	
Consider 100,000 randomly chosen airport surface operations within the U.S. under the general model conditions. What percentage of these operations will, during movement on the airfield surface, occupy an *INCORRECT PRESENCE* as defined in the materials provided? Express your uncertainty by providing a 5th, 50th, and 95th percentile of your estimate.		
5th percentile	**50th percentile**	**95th percentile**

Rank Correlation Questions

Incorrect Presence Factor Rank	
Variable	**Relative Rank**
Organizational/Regulatory	
Operational Environment	
Human Factors	
Weather	
Technological/Engineering	

RQ1	**Incorrect Presence Rank Correlation**
If 50,000 flight operations from the sample in Question 22 are randomly chosen, then the number of those operations where an *INCORRECT PRESENCE* occurs should be approximately half the median value from Question 22. Instead of randomly selecting these operations, suppose that only flights where *XXXXXX* is above its median value are chosen (Question X). Given this situation, what is the probability that instances of *INCORRECT PRESENCE* will be above half the 50th percentile estimate from Question 22? Given these conditions, what portion of these 50,000 will experience more than the median number of instances of *INCORRECT PRESENCE* in Question 2?	
Probability	**Portion (Count)**

Incorrect Presence Factor Influence			
Variable	**Rank**	**Influence as a % of highest ranked variable (Does not need to add up to 100%)**	**Direction of correlation (positive/negative)**
Organizational/Regulatory			
Operational Environment			
Human Factors			
Weather			
Technological/Engineering			



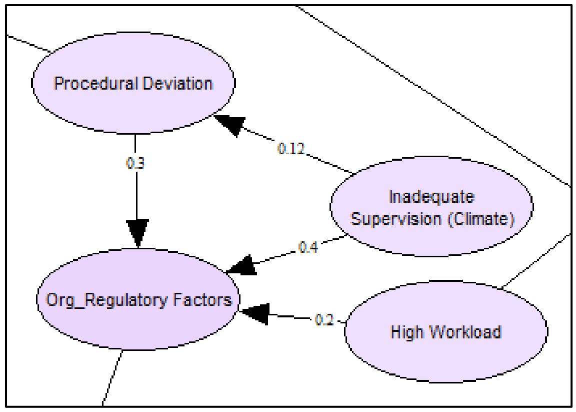
Conflicting Vehicle Traffic
Operational Environment Factors
Org_Regulatory Factors
Human Factors
Incorrect Position
Weather Factors
Darkness_Twilight
Tech_Engineering Factor

RQ2	Procedural Deviation Rank Correlation
If 50,000 flight operations from the sample in Question 1 are randomly chosen, then the number of those operations where a *DEVIATION FROM PROCEDURE* occurs should be approximately half the median value from Question 1. Instead of randomly selecting these operations, suppose that only operations where *INADEQUATE SUPERVISION* is above its median value are chosen (Question 2). Given this situation, what is the probability that the number of *DEVIATIONS FROM PROCEDURE* will be above half the 50th percentile estimate from Question 1? Given these conditions, what portion of these 50,000 will commit more than the median number of *DEVIATIONS FROM PROCEDURE* in Question 1?	
Probability	**Portion (Count)**

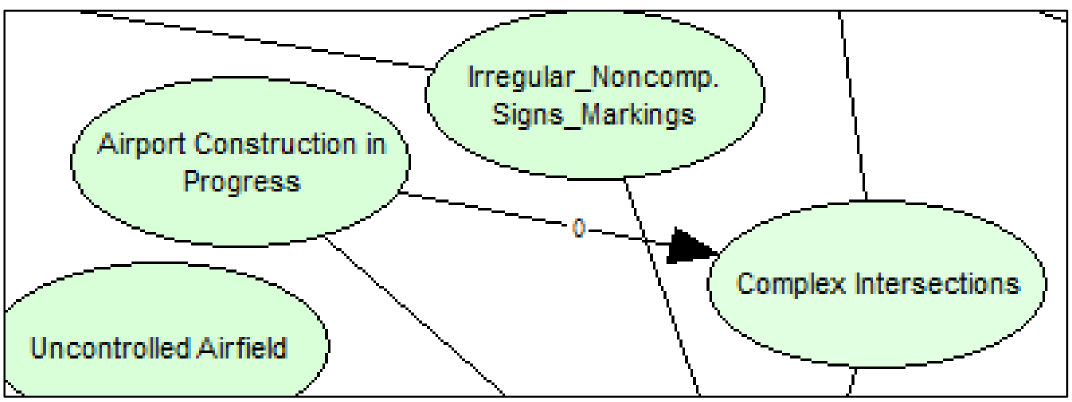
Organizational/Regulatory Factor Rank	
Variable	**Relative Rank**
High Workload	
Inadequate Supervision/Climate	
Procedural Deviation	

RQ3	Organization/Regulatory Factors Rank Correlation
If 50,000 flight operations from the sample in Question 4 are randomly chosen, then the number of those operations where an *ORGANIZATIONAL/REGULATORY* error occurs should be approximately half the median value from Question 4. Instead of randomly selecting these operations, suppose that 50,000 operations where *XXXXXXX* is at or above its median value are chosen (Question X). Given this situation, what is the probability that the number of *ORGANIZATIONAL/ REGULATORY* errors will be above half the 50th percentile estimate from Question 4? Given these conditions, what portion of these 50,000 will commit more than the median number of *ORGANIZATIONAL/REGULATORY* errors in Question 4?	
Probability	**Portion (Count)**

Organizational/Regulatory Factor Influence			
Variable	**Rank**	**Influence as a % of highest ranked variable (Does not need to add up to 100%)**	**Direction of correlation (positive/negative)**
High Workload			
Inadequate Supervision/Climate			
Procedural Deviation			



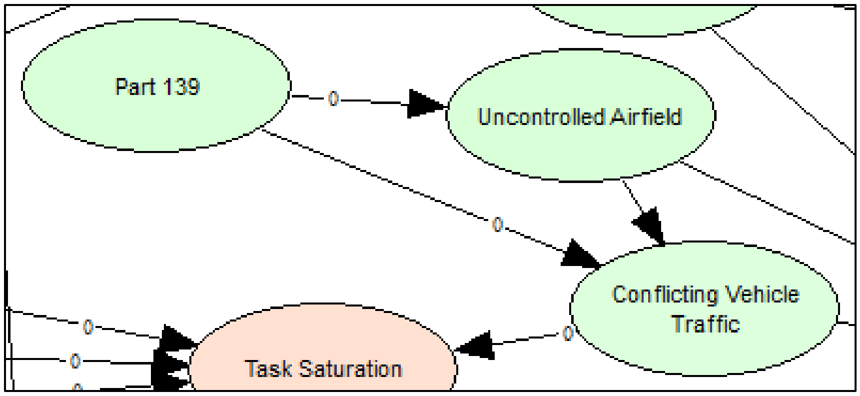
RQ4	Complex Intersection Rank Correlation
If 50,000 flight operations from the sample in Question 5 are randomly chosen, then the number of those operations where a *COMPLEX INTERSECTION* is encountered should be approximately half the median value from Question 5. Instead of randomly selecting these operations, suppose that 50,000 operations where *AIRFIELD CONSTRUCTION* is above its median value are chosen (Question 7). Given this situation, what is the probability that the number of *COMPLEX INTERSECTION* encounters will be above half the 50th percentile estimate from Question 5? Given these conditions, what portion of these 50,000 will be required to navigate more than the median number of *COMPLEX INTERSECTIONS* in Question 5?	
Probability	**Portion (Count)**



RQ5	Non-Towered Airfield Rank Correlation
If 50,000 flight operations from the sample in Question 9 are randomly chosen, then the number of those operations at an *NON-TOWERED AIRFIELDS* airfield should be approximately half the median value from Question 9. Instead of randomly selecting these operations, suppose that 50,000 operations where *PART 139* is below its median value are chosen (Question 8). Given this situation, what is the probability that the number of operations at *NON-TOWERED AIRFIELDS* will be above half the median value from Question 9? Given these conditions, what portion of these 50,000 will be at more than the median number of *NON-TOWERED AIRFIELDS* in Question 9?	
Probability	**Portion (Count)**

RQ6	Vehicle Traffic (I) Rank Correlation
If 50,000 flight operations from the sample in Question 10 are randomly chosen, then the number of those operations where *CONFLICTING VEHICLE TRAFFIC* is encountered should be approximately half the median value from Question 10. Instead of randomly selecting these operations, suppose that 50,000 operations where *NON-TOWERED AIRFIELDS* are encountered at above the median value are chosen (Question 9). Given this situation, what is the probability that the number of *CONFLICTING VEHICLE TRAFFIC* encounters will be above half the 50th percentile estimate from Question 10? Given these conditions, what portion of these 50,000 will be required to navigate more than the median number of *CONFLICTING VEHICLE TRAFFIC* interactions in Question 10?	
Probability	**Portion (Count)**

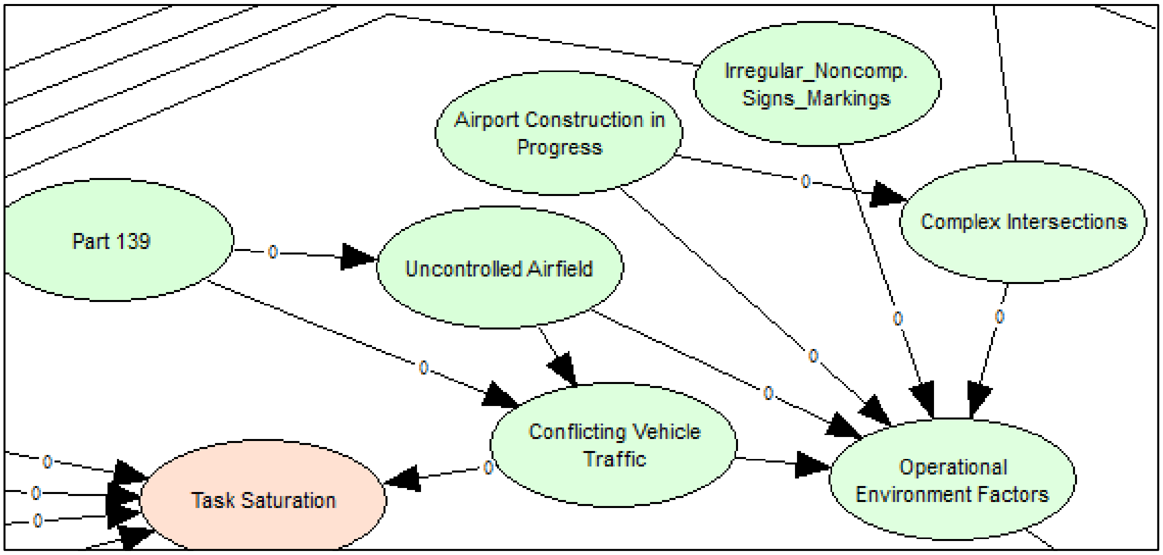
RQ7	Vehicle Traffic (II) Rank Correlation
If 50,000 flight operations from the sample in Question 10 are randomly chosen, then the number of those operations where *CONFLICTING VEHICLE TRAFFIC* is encountered should be approximately half the median value from Question 10. Instead of randomly selecting these operations, suppose that 50,000 operations below the median value of *PART 139* airfields are chosen (Question 8). Given this situation, what is the probability that the number of *CONFLICTING VEHICLE TRAFFIC* encounters will be above half the 50th percentile estimate from Question 10? Given these conditions, what portion of these 50,000 will be required to navigate more than the median number of *CONFLICTING VEHICLE TRAFFIC* interactions in Question 10?	
Probability	**Portion (Count)**



Operational Environment Factor Rank	
Variable	**Relative Rank**
Complex Intersections	
Irregular Signs/Markings	
Non-towered Airfield	
Airport Construction	
Conflicting Vehicle Traffic	

RQ8	Operating Environment Rank Correlation
If 50,000 operations from the sample in Question 11 are randomly chosen, then the number of those operations where an *OPERATING ENVIRONMENT* factor is present should be approximately half the median value from Question 11. Instead of randomly selecting these operations, suppose that only flights where *XXXXX* is above its median value are chosen (Question X). Given this situation, what is the probability that *OPERATING ENVIRONMENT* factors are above half the 50^{th} percentile estimate from Question 11? Given these conditions, what portion of these 50,000 will be above the median from Question 11?	
Probability	**Portion (Count)**

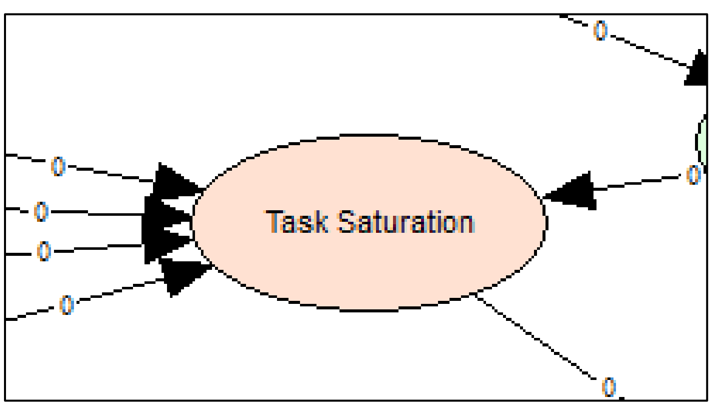
Operational Environment Factor Influence			
Variable	**Rank**	**Influence as a % of highest ranked variable (Does not need to add up to 100%)**	**Direction of correlation (positive/negative)**
Complex Intersections			
Irregular Signs/Markings			
Non-Towered Airfield			
Airport Construction			
Conflicting Vehicle Traffic			


Irregular_Noncomp.
Signs_Markings
Airport Construction in
Progress
Complex Intersections
Part 139
Uncontrolled Airfield
Conflicting Vehicle
Traffic
Operational
Environment Factors
Task Saturation
0
0
0
0
0
0
0
0
0
0
0
0

Task Saturation Factor Rank	
Variable	**Relative Rank**
High Workload	
Procedural Deviation	
ATC Equipment Failure	
Radio Congested	
Conflicting Vehicle Traffic	

RQ9	Human Factors Rank Correlation
If 50,000 flight operations from the sample in Question 12 are randomly chosen, then the number of those operations where *TASK SATURATION* occurs should be approximately half the median value from Question 15. Instead of randomly selecting these operations, suppose that only flights where a *XXXXX* is above its median value are chosen (Question X). Given this situation, what is the probability that instances of *TASK SATURATION* will be above half the 50th percentile estimate from Question 1? Given these conditions, what portion of these 50,000 will experience more than the median number of instances of *TASK SATURATION* in Question 12?	
Probability	**Portion (Count)**

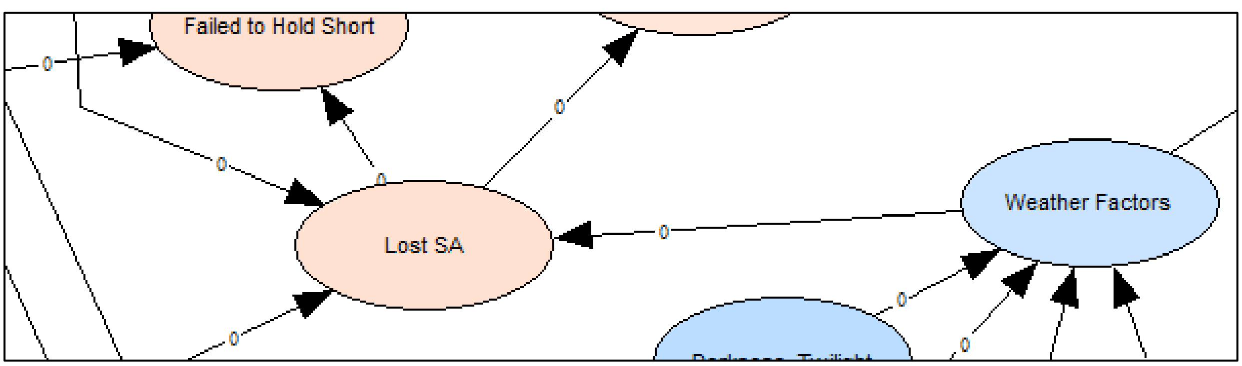
Task Saturation Factor Influence			
Variable	**Rank**	**Influence as a % of highest ranked variable (Does not need to add up to 100%)**	**Direction of correlation (positive/negative)**
High Workload			
Procedural Deviation			
ATC Equipment Failure			
Radio Congested			
Conflicting Vehicle Traffic			



Lost Situational Awareness Rank	
Variable	**Relative Rank**
Weather	
Complex Intersections	
Irregular Signs/Markings	

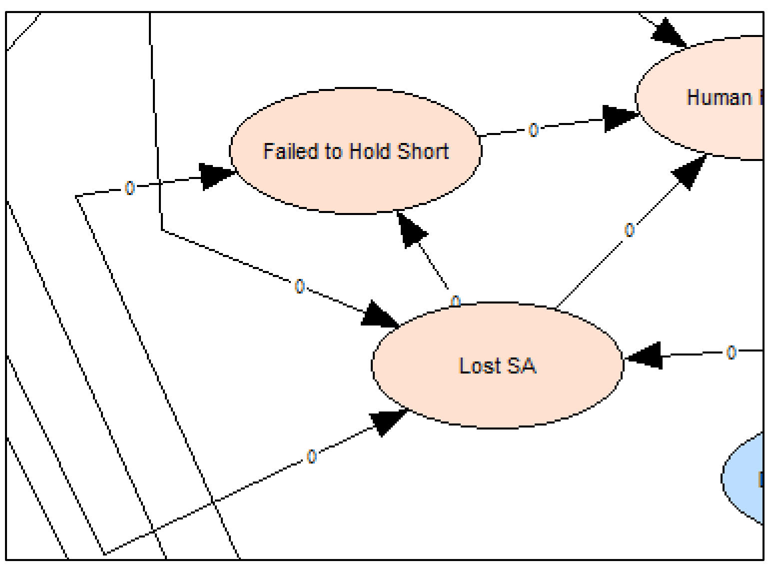
RQ10	**Lost Situational Awareness Rank Correlation**
If 50,000 flight operations from the sample in Question 14 are randomly chosen, then the number of those operations where *LOST S/A* occurs should be approximately half the median value from Question 14. Instead of randomly selecting these operations, suppose that only flights where a *XXXXXX* is above its median value are chosen (Question X). Given this situation, what is the probability that instances of *LOST S/A* will be above half the 50th percentile estimate from Question 14? Given these conditions, what portion of these 50,000 will experience more than the median number of instances of *LOST S/A* in Question 14?	
Probability	**Portion (Count)**

Lost Situational Awareness Influence			
Variable	**Rank**	**Influence as a % of highest ranked variable (Does not need to add up to 100%)**	**Direction of correlation (positive/negative)**
Weather			
Complex Intersections			
Irregular Signs/Markings			



RQ11	Failure to Hold Short Rank Correlation (I)
If 50,000 flight operations from the sample in Question 13 are randomly chosen, then the number of those operations where a *FAILURE TO HOLD SHORT* occurs should be approximately half the median value from Question 13. Instead of randomly selecting these operations, suppose that only flights where *LOST S/A* is above its median value are chosen (Question 14). Given this situation, what is the probability that *FAILURE TO HOLD SHORT* will be above half the 50^{th} percentile estimate from Question 13? Given these conditions, what portion of these 50,000 will experience more than the median number of instances of *FAILURE TO HOLD SHORT* in Question 13?	
Probability	**Portion (Count)**

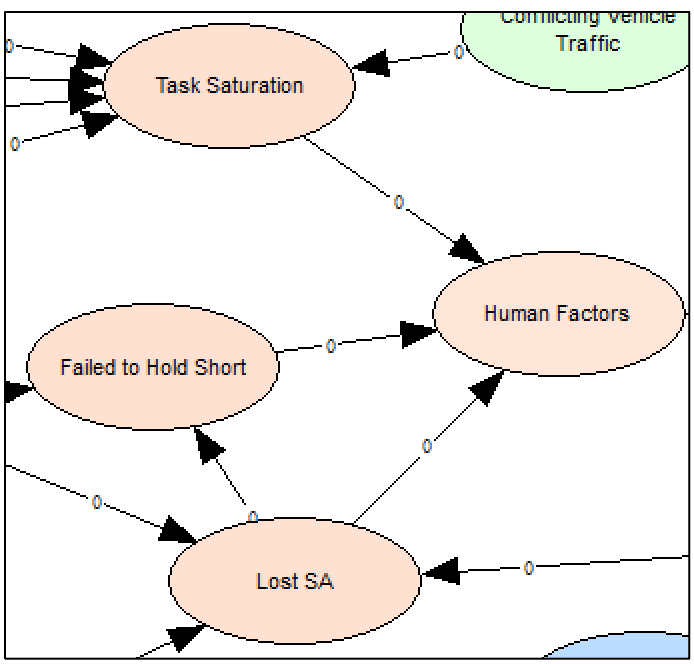
RQ12	Failure to Hold Short Rank Correlation (II)
If 50,000 flight operations from the sample in Question 13 are randomly chosen, then the number of those operations where a *FAILURE TO HOLD SHORT* occurs should be approximately half the median value from Question 13. Instead of randomly selecting these operations, suppose that only flights where *MECHANICAL FAILURE* is above its median value are chosen. Given this situation, what is the probability that *FAILURE TO HOLD SHORT* will be above half the 50^{th} percentile estimate from Question 13? Given these conditions, what portion of these 50,000 will experience more than the median number of instances of *FAILURE TO HOLD SHORT* in Question 13?	
Probability	**Portion (Count)**



Human Factors Rank	
Variable	**Relative Rank**
Task Saturation	
Failed to Hold Short	
Lost Situational Awareness	

RQ13	Human Factors Rank Correlation
If 50,000 flight operations from the sample in Question 15 are randomly chosen, then the number of those operations where a *HUMAN FACTORS* error occurs should be approximately half the median value from Question 15. Instead of randomly selecting these operations, suppose that only flights where a *XXXXX* is above its median value are chosen (Question X). Given this situation, what is the probability that instances of *HUMAN FACTORS* error will be above half the 50th percentile estimate from Question 15? Given these conditions, what portion of these 50,000 will experience more than the median number of instances of *HUMAN FACTORS* error in Question 15?	
Probability	**Portion (Count)**

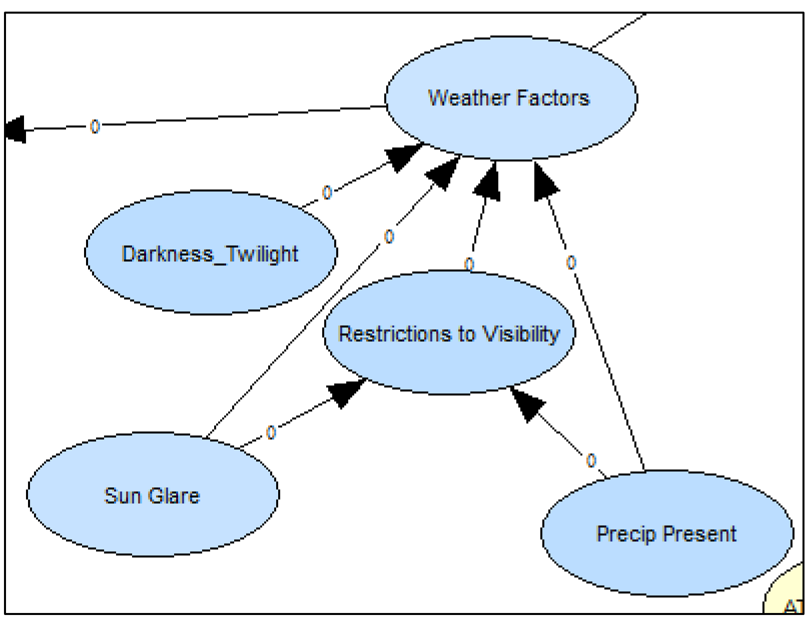
Human Factors Influence			
Variable	**Rank**	**Influence as a % of highest ranked variable (Does not need to add up to 100%)**	**Direction of correlation (positive/negative)**
Task Saturation			
Failed to Hold Short			
Lost Situational Awareness			



Weather Factor Rank	
Variable	**Relative Rank**
Darkness/Twilight	
Sun Glare	
Visibility Restriction	
Precipitation	

RQ14	Weather Rank Correlation
If 50,000 operations from the sample in Question 17 are randomly chosen, then the number of those operations where a *WEATHER FACTOR* occurs should be approximately half the median value from Question 17. Instead of randomly selecting these operations, suppose that only flights where *XXXXXX* occurs at above median value are chosen. Given this situation, what is the probability that instances of *WEATHER FACTOR* will be above half the 50th percentile estimate from Question 17? Given these conditions, what portion of these 50,000 will experience more than the median number of *WEATHER FACTORS* in Question 17?	
Probability	**Portion (Count)**

Organizational/Regulatory Factor Influence			
Variable	**Rank**	**Influence as a % of highest ranked variable (Does not need to add up to 100%)**	**Direction of correlation (positive/negative)**
Darkness/Twilight			
Sun Glare			
Visibility Restriction			
Precipitation			



RQ15	Visibility Rank Correlation (I)
If 50,000 operations from the sample for *RESTRICTIONS TO VISIBILITY* are randomly chosen, then the number of those operations where a *RESTRICTED VISIBILITY* occurs should be approximately half the median value. Instead of randomly selecting these operations, suppose that only flights where *SUN GLARE* occurs at above median value are chosen (Question 16). Given this situation, what is the probability that instances of *RESTRICTED VISIBILITY* will be above half the 50th percentile value? Given these conditions, what portion of these 50,000 will experience more than the median *RESTRICTED VISIBILITY*?	
Probability	**Portion (Count)**

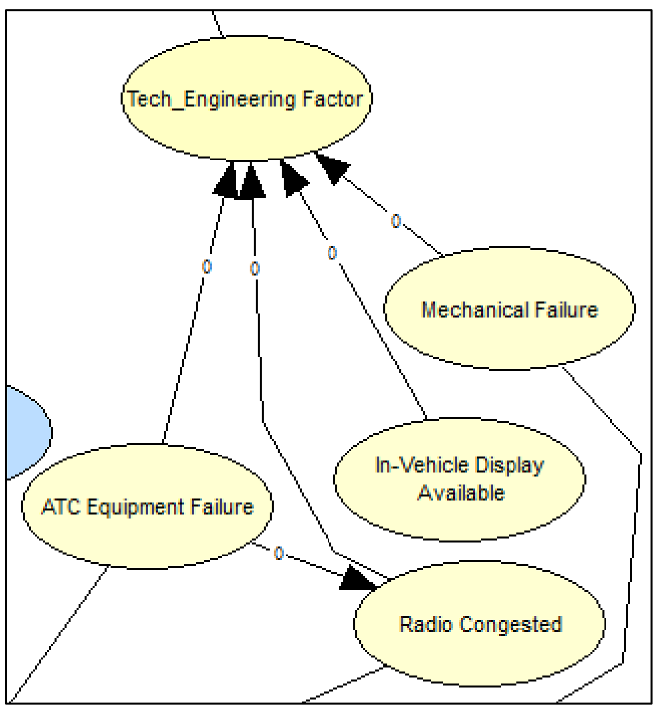
RQ16	Visibility Rank Correlation (II)
If 50,000 operations from the sample for *RESTRICTIONS TO VISIBILITY* are randomly chosen, then the number of those operations where a *RESTRICTED VISIBILITY* occurs should be approximately half the median value. Instead of randomly selecting these operations, suppose that only flights where *PRECIPITATION* occurs at above median value are chosen (Question 16). Given this situation, what is the probability that instances of *RESTRICTED VISIBILITY* will be above half the 50th percentile value? Given these conditions, what portion of these 50,000 will experience more than the median *RESTRICTED VISIBILITY*?	
Probability	**Portion (Count)**

RQ17	Radio Congestion Rank Correlation
If 50,000 operations from the sample for *RADIO CONGESTION* are randomly chosen, then the number of those operations where *RADIO CONGESTION* occurs should be approximately half the median value. Instead of randomly selecting these operations, suppose that only flights where an *ATC EQUIPMENT FAILURE* occurs at above median value are chosen (Question 18). Given this situation, what is the probability that instances of *RADIO CONGESTION* will be above half the 50th percentile value? Given these conditions, what portion of these 50,000 will experience more than the median *RADIO CONGESTION*?	
Probability	**Portion (Count)**

Technical/Engineering Factor Rank	
Variable	**Relative Rank**
ATC Equipment Failure	
Radio Congestion	
In-Vehicle Display	
Mechanical Failure	

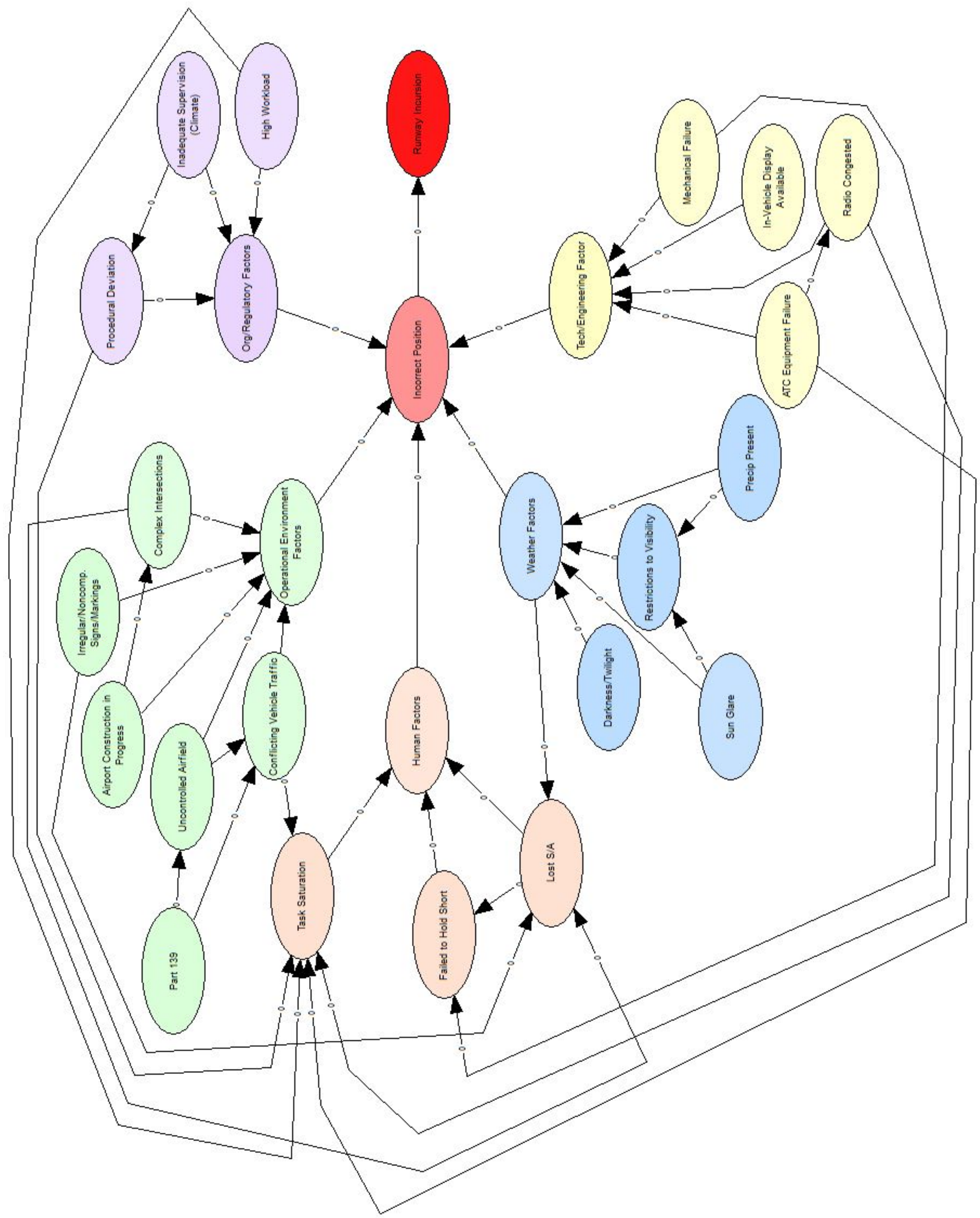
RQ18	Technical/Engineering Rank Correlation
If 50,000 flight operations from the sample in Question 21 are randomly chosen, then the number of those operations where a *TECHNICAL/ENGINEERING FACTOR* occurs should be approximately half the median value from Question 21. Instead of randomly selecting these operations, suppose that only flights where *XXXXX* is above its median value are chosen. Given this situation, what is the probability that *TECHNICAL/ENGINEERING FACTORS* will be above half the 50th percentile estimate from Question 21? Given these conditions, what portion of these 50,000 will experience more than the median number of *TECHNICAL/ ENGINEERING FACTORS* in Question 21?	
Probability	**Portion (Count)**

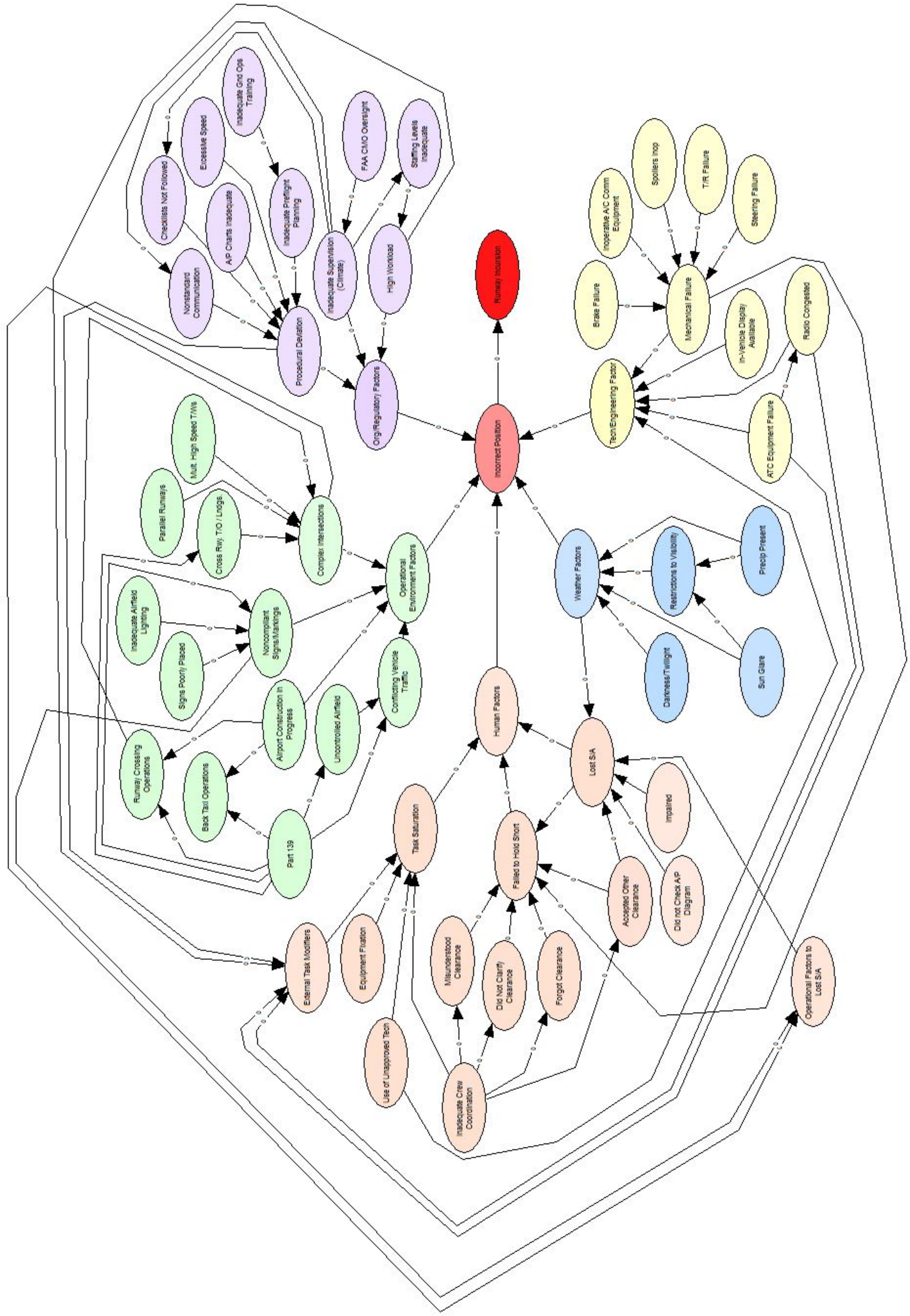
Technical/Engineering Factor Influence			
Variable	**Rank**	**Influence as a % of highest ranked variable (Does not need to add up to 100%)**	**Direction of correlation (positive/negative)**
ATC Equipment Failure			
Radio Congestion			
In-Vehicle Display			
Mechanical Failure			



Domain	Causal/Contributory Factors	Description
Organizational/Regulatory	Procedural Deviation	Describes deviation from established corporate, organizational, or regulatory procedures by an operator
	Inadequate Supervision (Climate)	Describes absence of supervisory input or pressure from the supervisory or organizational level to perform tasks motivated by factors inconsistent with a climate of compliance and safety
	High Workload	Describes workload levels that affect individual performance negatively by amplifying inattention or lack of focus
Operational Environment	Complex Intersections	Describes the probability that an airport has complex intersections - those taxiway or runway intersections where number of intersecting surfaces, signage, lighting, or aircraft geometry may combine to create unusually high confusion for operators
	Irregular Signs/Markings	Describes irregular or non ICAO-compliant signs and markings on an airport surface
	Construction	Describes the presence of construction in the airport movement area
	Conflicting Vehicle Traffic	Describes the presence of conflicting vehicle traffic in the airport movement area
Human Factors	Lost S/A	Describes a loss of flight crew or vehicle operator situational awareness, meaning that inappropriate mental representations are activated in spite of real world evidence. People then act “in the wrong scene,” and seek cues confirming their expectations, a behavior known as confirmation bias.

	Failed to Hold Short	Describes an aircraft or vehicle failure to remain in place behind a hold short line despite instructions by a controlling authority or regulatory requirement to do so
	Task Saturation	Describes an operational condition of no awareness of input from various sources, so decisions might be made with incomplete information and the possibility of error increases
Weather	Sun Glare	Describes the presence of sun glare that interferes with vision of ATC, flight crew, or ground vehicle operator
Technical	Radio Congestion	Describes occurrence of radio congestion that requires operators to initiate multiple calls or wait to call or respond to ATC instruction
	In-Cockpit Technology (Moving Map w/ Ownship. Etc.)	Describes the presence of technology in a flight deck or vehicle that displays the airport layout or combination of layout and vehicle or aircraft position in real time


Inadequate Supervision (Climate)
High Workload
Runway Incursion
Procedural Deviation
Org/Regulatory Factors
Incorrect Position
Tech/Engineering Factor
Mechanical Failure
In-Vehicle Display Available
Radio Congested
ATC Equipment Failure
Precip Present
Complex Intersections
Operational Environment Factors
Weather Factors
Restrictions to Visibility
Irregular/Noncomp. Signs/Markings
Darkness/Twilight
Sun Glare
Conflicting Vehicle Traffic
Human Factors
Airport Construction in Progress
Uncontrolled Airfield
Lost S/A
Task Saturation
Failed to Hold Short
Part 139



Runway Incursion
Incorrect Position
Org/Regulatory Factors
Procedural Deviation
Nonstandard Communication
Checklists Not Followed
Excessive Speed
A/P Charts Inadequate
Inadequate Gnd Ops Training
Inadequate Preflight Planning
Inadequate Supervision (Climate)
FAA CMO Oversight
Staffing Levels Inadequate
High Workload
Tech/Engineering Factor
Mechanical Failure
Inoperative A/C Comm Equipment
Spoilers Inop
T/R Failure
Steering Failure
Brake Failure
In-Vehicle Display Available
Radio Congested
ATC Equipment Failure
Weather Factors
Restrictions to Visibility
Precip Present
Darkness/Twilight
Sun Glare
Human Factors
Lost S/A
Impaired
Did not Check A/P Diagram
Accepted Other Clearance
Failed to Hold Short
Task Saturation
Misunderstood Clearance
Did Not Clarify Clearance
Forgot Clearance
Inadequate Crew Coordination
Use of Unapproved Tech
Equipment Fixation
External Task Modifiers
Operational Factors to Lost S/A
Operational Environment Factors
Complex Intersections
Mult. High Speed T/Ws
Parallel Runways
Cross Rwy. T/O / Lndgs
Noncompliant Signs/Markings
Inadequate Airfield Lighting
Signs Poorly Placed
Airport Construction in Progress
Conflicting Vehicle Traffic
Uncontrolled Airfield
Runway Crossing Operations
Back Taxi Operations
Part 139

APPENDIX F

Causal Codes Available for SME Review of ASRS Narratives

2.1 ***Communications***

2.1.1 Transmission was completely blocked

2.1.2 Transmission was partially blocked ("stepped-on")

2.1.3 Accepted a similar aircraft's clearance:

- with similar call signs
- without similar call signs

2.1.4 Deviation from established ICAO standard phraseologies

2.1.5 Used other than ICAO language requirements for air-ground radiotelephony communications (language normally used by the station on the ground or the English language) in a situation not covered by ICAO standard phraseology

2.1.6 Used language not in accordance with ICAO language requirements for air-ground radiotelephony communications (language normally used by the station on the ground or the English language)

2.1.7 Speech quality:

- not proficient in ICAO language requirements for air-ground radiotelephony communications (language normally used by the station on the ground or the English language)
- poorly enunciated or heavily accented
- spoken rapidly
- spoken with an inconsistent volume

2.1.8 Did not use headsets

2.1.9 Received clearance or instructions during periods of high cockpit workload

2.1.10 Did not advise ATC of a delay on the runway prior to take-off

2.1.11 Other (please specify).

2.2 ***Situational awareness***

2.2.1 Crew conducting checklists while taxiing

2.2.2 Crew member programming flight management system or other flight deck system while taxiing

2.2.3 Crew member was on another radio frequency

2.2.4 Competing radio communications

2.2.5 Unfamiliar with the aerodrome layout

2.2.6 Crew mistook their position on the aerodrome (thought they were in a different location)

2.2.7 Fatigue

2.2.8 Reported incorrect location to ATC

2.2.9 Taxied fast

2.2.10 Did not refer to the aerodrome diagram

2.2.11 Did not listen to the automatic terminal information service (ATIS)

2.2.12 Works on the manoeuvring area were not previously advised by NOTAM

2.2.13 Used out-of-date or inaccurate publications or charts

2.2.14 Failed to apply or correctly observe sterile cockpit procedures

2.2.15 Other (please specify).

Appendix F (Continued)

Causal Codes Available for SME Review of ASRS Narratives

2.3 ***Markings, signs and lighting***
2.3.1 Not ICAO-compliant
2.3.2 Not provided
2.3.3 Irregularly spaced
2.3.4 Ambiguous and difficult to follow
2.3.5 Poorly sized
2.3.6 Poorly situated
2.3.7 Poorly maintained
2.3.8 Other (please specify).

2.4 ***Clearances and instructions***
2.4.1 Misunderstood clearance:
- conditional
- follow
- other

2.4.2 Flight crew did not ask for clarification when they did not understand a clearance or instruction
2.4.3 Did not inform ATC when could not comply with a clearance
2.4.4 Forgot part of the clearance or instruction
2.4.5 Entered the runway after being instructed to "hold short"
2.4.6 Lined up on the runway after instruction to taxi to the runway-holding position (point)
2.4.7 Took off without a clearance after being instructed to "line up and wait"
2.4.8 Took off without a clearance after being instructed to taxi to the runway-holding position (point)
2.4.9 Landed or departed on the wrong runway
2.4.10 Landed or departed on the taxiway
2.4.11 Other (please specify).

2.5 Expert Rater-Developed Codes

2.5.1 Crew Coordination
2.5.2 Failure to Readback Clearance

APPENDIX G

SQL Code for Interrater Computations

Note to readers: The following code does not represent the complete SQL input for the interrater reliability, union, and intersection functions. Rather, based on space limitations and the repetitive nature of the code, it shows the input for one of the five iterations necessary to complete the function.

```
        set nocount on

        declare @icao table (id int identity(1,1),rating varchar(50))

        declare @BG table(asr varchar(30), apt varchar(30),r1 varchar(30),r2 varchar(30),
                r3 varchar(30), r4 varchar(30), r5 varchar(30), dq varchar(255), comment
varchar(255),
                exc varchar(30), sev varchar(20))

        declare @GJ table(asr varchar(30), apt varchar(30),r1 varchar(30),r2 varchar(30),
                r3 varchar(30), r4 varchar(30), r5 varchar(30), dq varchar(255), comment
varchar(255),
                exc varchar(30), sev varchar(20))

        declare @JT table(asr varchar(30), apt varchar(30),r1 varchar(30),r2 varchar(30),
                r3 varchar(30), r4 varchar(30), r5 varchar(30), dq varchar(255), comment
varchar(255),
                exc varchar(30), sev varchar(20))

        declare @asrs table (asr varchar(30))

        declare @interR table (asr varchar(30), who varchar(30), rating varchar(30),
codedRating int)


        insert @BG(asr,r1,r2,r3,r4,r5,exc) select '836163','2.3.6','2.3.3',",",",'No'
        insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '836163','2.3.3','2.3.4','2.3.6',","
        insert @JT(asr,r1,r2,r3,r4,r5,exc) select '836163','2.3.4',",",",",'No'
        insert @BG(asr,r1,r2,r3,r4,r5,exc) select '789540',",",",",",'Yes'
        insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '789540','2.5.1',",",","
        insert @JT(asr,r1,r2,r3,r4,r5,exc) select '789540','2.2.15','2.5.1',",",",'No'
        insert @BG(asr,r1,r2,r3,r4,r5,exc) select '817153','2.4.5','2.4.1',",",",'No'
        insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '817153','2.1.4','2.2.5','2.3.2',","
        insert @JT(asr,r1,r2,r3,r4,r5,exc) select '817153','2.3.4','2.2.5','2.4.1',",",'Yes'
        insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'824311','2.4.5','2.4.2','2.2.7','2.2.1','2.4.1','No'
        insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '824311','2.2.1','2.1.9','2.2.7',","
```

```
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '824311','2.2.7','2.1.9','2.4.5',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '856792','2.4.5','2.4.1',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '856792','2.4.1','2.2.7','2.4.5',",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '856792','2.4.1','2.4.5',",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'812538','2.4.5','2.2.5','2.2.2','2.2.1','2.1.9','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '812538','2.2.2','2.2.15','2.5.1',",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '812538','2.2.2','2.2.15',",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'942802','2.4.5','2.4.1','2.4.2','2.1.7','2.1.4','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '942802','2.1.7','2.4.1','2.4.5',",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '942802','2.4.1','2.1.4','2.1.7',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '979222','2.4.2','2.4.6','2.4.1',",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '979222','2.1.9','2.2.1','2.4.1','2.4.2',","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '979222','2.4.2',",",",",'Yes'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '785382','2.4.5','2.3.7','2.2.1','2.2.5',",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '785382','2.2.5','2.2.10','2.3.3','2.5.1',","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '785382','2.2.5','2.3.4','2.3.7','2.2.15',",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '792763','2.4.5','2.2.1','2.2.2','2.2.7',",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '792763','2.1.9','2.2.1',",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '792763','2.1.9','2.2.2','2.2.7','2.4.5',",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '784979','2.4.5','2.2.2',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '784979','2.2.15',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '784979','2.2.15',",",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '856457','2.3.4','2.3.6','2.2.10',",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '856457','2.2.6','2.3.4','2.3.6',",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '856457','2.3.4',",",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '776226','2.2.5','2.3.3',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '776226','2.2.1','2.2.13','2.3.4','2.4.2',","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '776226','2.3.4','2.2.1','2.2.13','2.4.2',",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '884378','2.2.5','2.3.7','2.3.1',",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '884378','2.2.13','2.3.8',",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '884378','2.3.4','2.3.7',",",",'Yes'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'867483','2.2.6','2.2.1','2.2.2','2.2.10','2.2.15','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '867483','2.2.1','2.2.9','2.4.5','2.5.1',","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '867483','2.2.1','2.2.9','2.2.15',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '792259','2.4.11',",",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '792259','2.2.11','2.5.1',",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '792259','2.2.11','2.4.11','2.5.1',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '837821','2.4.11','2.2.15',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '837821','2.4.11',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '837821','2.1.9','2.1.11',",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '793916','2.4.5','2.4.3',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '793916','2.2.9','2.2.15',",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '793916','2.2.9','2.2.15',",",",'No'
```

```
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '831760','2.3.2','2.2.10','2.2.5','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '831760','2.3.2','2.3.6','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '831760','2.3.4','2.3.8','','','','Yes'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '840502','2.2.15','2.4.11','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '840502','2.2.15','','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '840502','2.2.15','','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '858253','2.3.1','2.2.10','2.2.2','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '858253','2.3.6','2.2.5','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '858253','2.3.6','2.2.2','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '859637','','','','','','Yes'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '859637','2.2.10','2.2.5','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '859637','2.2.6','2.2.1','2.2.10','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '829659','','','','','','Yes'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '829659','2.2.13','','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '829659','2.4.1','2.3.4','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '969670','2.4.5','2.2.5','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '969670','2.2.5','','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '969670','2.1.9','2.2.5','2.3.4','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '978488','2.4.5','2.2.5','2.4.1','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '978488','2.3.4','2.2.5','2.1.9','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '978488','2.2.5','2.3.4','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '790954','2.2.15','2.1.11','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '790954','2.1.11','2.2.3','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '790954','2.1.11','2.2.15','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '974660','2.4.1','2.4.2','2.4.4','2.1.9','2.2.1','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '974660','2.2.1','2.2.4','2.4.1','2.5.1','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '974660','2.2.4','2.2.1','2.2.15','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '882759','2.4.1','2.4.5','2.4.6','2.2.1','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '882759','2.2.1','2.4.6','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '882759','2.2.1','2.4.6','2.2.15','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '847101','2.4.11','','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '847101','2.2.15','','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '847101','','','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '939675','2.4.5','2.4.1','2.2.10','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '939675','2.2.5','2.2.10','2.4.2','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '939675','2.2.5','2.2.15','2.4.5','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '782334','2.2.15','2.3.2','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '782334','2.2.15','','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '782334','2.2.15','2.3.1','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '839671','2.4.1','2.1.3','2.1.11','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '839671','2.1.3','','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '839671','2.4.1','','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '790028','2.2.10','2.2.15','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '790028','2.2.10','2.4.11','2.5.1','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '790028','2.4.4','2.5.1','','','','No'
```

```
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '895524','2.4.5','','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '895524','2.4.5','','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '895524','','','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '979457','2.1.2','2.2.15','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '979457','2.4.11','','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '979457','2.4.11','','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '808374','2.2.15','2.1.11','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '808374','2.2.4','','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '808374','2.2.4','','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '796225','2.4.5','2.3.6','2.2.15','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '796225','2.2.13','2.3.7','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '796225','2.3.4','2.3.8','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '964667','2.3.4','2.4.11','2.4.5','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '964667','2.3.2','2.4.5','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '964667','2.3.2','2.2.15','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '796451','2.4.9','2.2.6','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '796451','2.4.1','','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '796451','2.4.9','2.2.15','2.4.11','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '885530','2.4.1','2.4.6','2.4.5','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '885530','2.4.5','2.4.1','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '885530','2.4.1','2.4.5','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '924416','2.4.5','2.4.1','2.2.14','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '924416','2.4.1','2.4.6','2.1.11','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '924416','2.4.5','2.5.1','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'963731','2.4.4','2.4.5','2.2.10','2.2.14','2.2.5','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '963731','2.2.5','2.2.10','2.2.14','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select
'963731','2.2.14','2.2.10','2.4.4','2.4.5','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '785071','2.3.6','2.3.5','2.4.5','2.2.5','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '785071','2.3.3','2.3.4','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '785071','2.3.4','2.2.5','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'792103','2.4.1','2.4.5','2.2.13','2.3.4','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '792103','2.4.5','2.2.6','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '792103','2.2.5','2.2.15','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '773565','2.2.6','2.2.10','2.3.8','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '773565','2.2.5','2.3.7','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '773565','2.3.8','','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '902776','2.3.4','2.3.7','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '902776','2.3.6','2.3.4','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '902776','2.3.4','2.3.7','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '848283','2.1.11','2.2.15','','','','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '848283','2.1.11','','','','',''
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '848283','2.1.11','2.2.15','','','','No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '809047','2.2.6','2.2.14','2.2.5','','','No'
```

```
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '809047','2.2.14',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '809047','2.5.1','2.2.1',",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'781972','2.4.5','2.4.1','2.2.9','2.2.10','2.2.5','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '781972','2.4.5',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '781972','2.2.6','2.4.2','2.5.1',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '808596','2.2.6','2.2.5','2.4.2','2.5.1',",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '808596','2.2.6','2.5.1',",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '808596','2.5.1','2.2.15','2.2.6','2.4.2',",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'848426','2.2.6','2.2.5','2.4.2','2.4.5','2.2.15','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '848426','2.2.6','2.2.15',",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '848426','2.2.15','2.2.5',",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '926377','2.4.11',",",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '926377','2.4.11',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '926377','2.2.15','2.4.3',",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'891040','2.1.3','2.4.5','2.2.7','2.2.15',",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '891040','2.4.1',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '891040','2.1.3','2.2.15',",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'917147','2.3.6','2.2.10','2.2.2','2.2.15',",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '917147','2.2.10',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '917147','2.2.2','2.1.9','2.3.8',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '846943','2.4.1','2.1.11','2.2.15',",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '846943','2.4.11',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '846943','2.4.11',",",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'873641','2.3.6','2.2.10','2.4.5','2.2.5',",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '873641','2.3.7','2.2.10',",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '873641','2.3.7','2.2.5','2.2.10',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '1001233','2.4.10','2.2.6','2.2.10',",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '1001233','2.2.10',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '1001233','2.4.10','2.2.5','2.2.10',",",'Yes'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '840082','2.2.15','2.2.13',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '840082','2.2.15',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '840082','2.2.15','2.3.1','2.1.11',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '998522','2.4.9','2.2.6','2.2.5','2.3.4',",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '998522','2.3.4',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '998522','2.3.4','2.2.6','2.2.10',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '823433','2.4.11','2.4.2',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '823433','2.4.11',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '823433','2.4.11',",",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '840535','2.1.11','2.2.13',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '840535','2.2.3',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '840535','2.2.15','2.2.13',",",",'No'
```

```
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '949123','2.2.15','2.1.11',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '949123','2.1.11',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '949123','2.1.11','2.2.15',",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '813384','2.1.11','2.2.4',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '813384','2.4.1',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '813384',",",",",",'Yes'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '845126','2.3.4','2.4.5',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '845126','2.3.4','2.3.6',",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '845126','2.3.4',",",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'818774','2.4.5','2.2.10','2.3.7','2.3.3',",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '818774','2.2.9','2.2.10',",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '818774','2.2.5','2.2.2','2.3.7',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '906346','2.3.3','2.4.1',",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '906346','2.1.4',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '906346','2.3.4','2.4.1',",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '955078','2.2.6','2.2.10','2.4.4',",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '955078','2.1.7',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '955078','2.2.10','2.2.15',",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '971495','2.4.5','2.2.6','2.4.5','2.3.7',",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '971495','2.3.7','2.2.5','2.4.5',",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '971495','2.3.7','2.2.10','2.2.6',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'844690','2.4.9','2.2.10','2.2.5','2.2.15',",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '844690','2.2.10',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '844690','2.2.6','2.4.9',",",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select
'1019890','2.2.1','2.2.5','2.2.7','2.2.10','2.5.1','No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '1019890','2.2.10','2.4.1',",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '1019890','2.2.1','2.2.2','2.5.1',",",'No'
insert @BG(asr,r1,r2,r3,r4,r5,exc) select '838570','2.1.3',",",",",'No'
insert @GJ(asr,r1,r2,r3,r4,r5,exc) select '838570','2.1.3',",",",","
insert @JT(asr,r1,r2,r3,r4,r5,exc) select '838570','2.1.3',",",",",'No'


insert @icao(rating) select '2.1.1'
insert @icao(rating) select '2.1.10'
insert @icao(rating) select '2.1.11'
insert @icao(rating) select '2.1.2'
insert @icao(rating) select '2.1.3'
insert @icao(rating) select '2.1.4'
insert @icao(rating) select '2.1.5'
insert @icao(rating) select '2.1.6'
insert @icao(rating) select '2.1.7'
insert @icao(rating) select '2.1.8'
insert @icao(rating) select '2.1.9'
```

```
insert @icao(rating) select '2.2.1'
insert @icao(rating) select '2.2.10'
insert @icao(rating) select '2.2.11'
insert @icao(rating) select '2.2.12'
insert @icao(rating) select '2.2.13'
insert @icao(rating) select '2.2.14'
insert @icao(rating) select '2.2.15'
insert @icao(rating) select '2.2.2'
insert @icao(rating) select '2.2.3'
insert @icao(rating) select '2.2.4'
insert @icao(rating) select '2.2.5'
insert @icao(rating) select '2.2.6'
insert @icao(rating) select '2.2.7'
insert @icao(rating) select '2.2.8'
insert @icao(rating) select '2.2.9'
insert @icao(rating) select '2.3.1'
insert @icao(rating) select '2.3.2'
insert @icao(rating) select '2.3.3'
insert @icao(rating) select '2.3.4'
insert @icao(rating) select '2.3.5'
insert @icao(rating) select '2.3.6'
insert @icao(rating) select '2.3.7'
insert @icao(rating) select '2.3.8'
insert @icao(rating) select '2.4.1'
insert @icao(rating) select '2.4.10'
insert @icao(rating) select '2.4.11'
insert @icao(rating) select '2.4.2'
insert @icao(rating) select '2.4.3'
insert @icao(rating) select '2.4.4'
insert @icao(rating) select '2.4.5'
insert @icao(rating) select '2.4.6'
insert @icao(rating) select '2.4.7'
insert @icao(rating) select '2.4.8'
insert @icao(rating) select '2.4.9'
insert @icao(rating) select '2.5.1'
insert @icao(rating) select '2.5.2'

--cleanup
set nocount off
delete @BG where rtrim(ltrim(asr))=''
delete @JT where rtrim(ltrim(asr))=''
delete @GJ where rtrim(ltrim(asr))=''
set nocount on

update @BG set exc='No' where exc=''
update @JT set exc='No' where exc=''
```

```
update @GJ set exc='No' where exc=''



insert @asrs(asr) select distinct asr from @BG




insert @interR (who, asr, rating) select 'GJ',asr,r1 from @GJ
insert @interR (who, asr, rating) select 'GJ',asr,r2 from @GJ
insert @interR (who, asr, rating) select 'GJ',asr,r3 from @GJ
insert @interR (who, asr, rating) select 'GJ',asr,r4 from @GJ
insert @interR (who, asr, rating) select 'GJ',asr,r5 from @GJ

insert @interR (who, asr, rating) select 'JT',asr,r1 from @JT
insert @interR (who, asr, rating) select 'JT',asr,r2 from @JT
insert @interR (who, asr, rating) select 'JT',asr,r3 from @JT
insert @interR (who, asr, rating) select 'JT',asr,r4 from @JT
insert @interR (who, asr, rating) select 'JT',asr,r5 from @JT

insert @interR (who, asr, rating) select 'BG',asr,r1 from @BG
insert @interR (who, asr, rating) select 'BG',asr,r2 from @BG
insert @interR (who, asr, rating) select 'BG',asr,r3 from @BG
insert @interR (who, asr, rating) select 'BG',asr,r4 from @BG
insert @interR (who, asr, rating) select 'BG',asr,r5 from @BG

update @interR set codedRating=i.ID from
        @interR r inner join @icao i
                on r.rating=i.rating


--looking at rating congruence

--select * from @interR order by asr,who

--unique rating counts
/*
select unqratings=count(distinct rating),totratings=count(*),asr
        from @interR where rating!=''
                group by asr
                order by count(distinct rating) desc
*/

/*
```

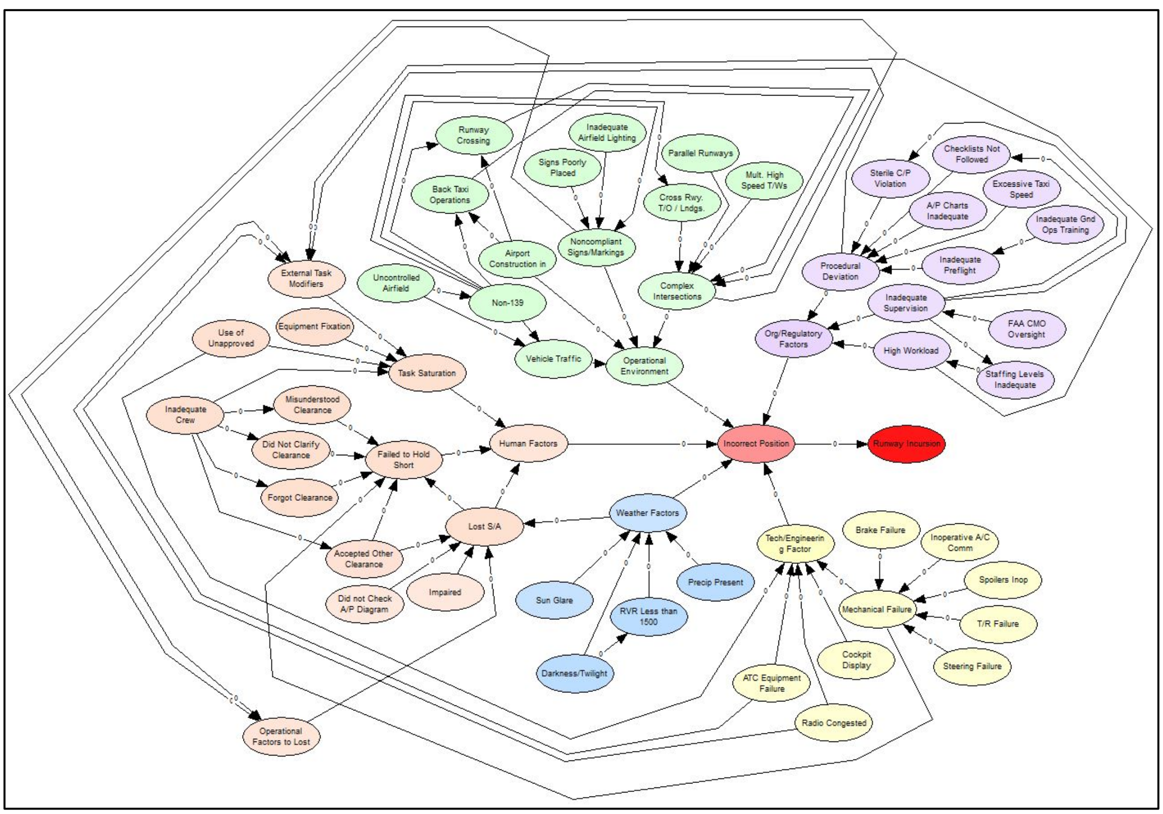
APPENDIX H

SME Structural Model Review Protocol

Your (SMEs) input, data from ASRS and RSO records, and a systematic review of the literature shaped the development of the model shown here. While data exist for some of the nodes, others will be populated with conditional, probabilistic estimates based on input from SMEs.

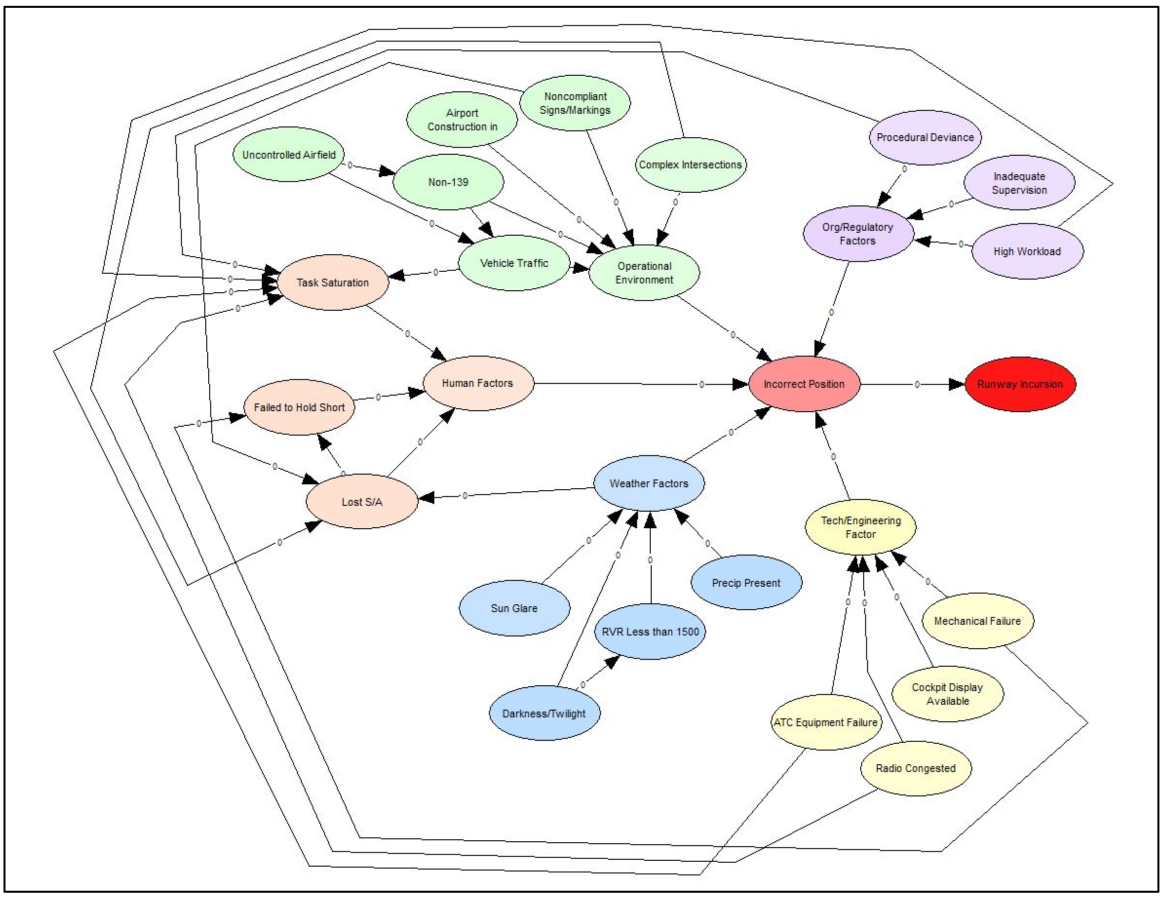
For this phase of the research, I will ask you to review the *structure* of the models below. This structure is important because it attempts to capture how and why RIs occur in a dynamic setting. The arrows between nodes, called *directed edges*, indicate a causal connection between variables called *parent* and *child nodes*. In subsequent sessions, you may be asked to identify the strength of association between nodes or the probability of a certain node state; however, in this phase, you need only focus on the direction of causality and the connections between nodes both within and across domains (identified by color coding).

Please take a few moments to review the model below, asking questions as they arise. Remember, this model is a mode detailed "landscape view" of the problem space. Once you have familiarized yourself with this model, we will move on to a more parsimonious version for your review.



Note that the model below is not fundamentally different from the first model. Instead, it is a more compact version. On this basis, we will begin the exercise of evaluating structure here, expanding our review to the more granular model next. As we work through this exercise, keep in mind the definition of a runway incursion (RI) is: "Any occurrence at an aerodrome involving the incorrect presence of an aircraft, vehicle, or person on the protected area of a surface designated for the landing and take-off of aircraft" (FAA, 2007, para. 2).

Let's review the model below beginning with the node labeled "Incorrect Position". This node is intended to capture the requirement that an aircraft or vehicle arrives at a position on an airfield, specifically in the runway protected zone, in an incorrect or unintended manner. The radiating nodes are essentially conditions that necessarily lead to this incorrect presence, but may not be sufficient to create the incorrect presence.



Now that we have discussed the interaction between domain-specific causal factors and the instance of an improper position combining to create an RI event, let's look at the causal relationships indicated by directed edges, starting with the unconditioned nodes (nodes without *parent nodes*) in the Organizational/Regulatory domain, and then moving inward before moving counter-clockwise to the next domain, and so on.

In this process, it may be helpful to refer back to the more granular model to gain a better understanding of the factors leading to the more general nodes.

Questions for SMEs:

• In your opinion/experience, does the node capture the relevant construct?
• What are the possible states of the node?
• Yes/No
• Correct/Incorrect
• Is the node sequenced properly?
• Is the node necessary to a complete causal sequence?
• Does the node have true, causal influence on its *child nodes*?
• Are the data to populate the node available?
• If so, what is your assessment of the quality of the data?
• What causal factors have not been accounted for?
• Why do you believe their inclusion is supported?
• Are the inter-domain links (*directed edges*) appropriate and necessary?
• Do you have any other comments/input?

APPENDIX I

Range Graphs by Question/Expert with Equal/Global Weights

```
                         Range graph of input data
Item no.:   1 Item name: CQ1 Scale: UNI
Experts
  1 [-------------------*-----------------------]
  2                     [--------*------------]
  3 [-----------*---------------]
  4             [---------*-------]
  5            [-------------------------------*------------------------------]
  6            [----------------*----------------]
Global                  [=======*============]
Equa [=======================*======================================]
Real:::::::::::::::::::::::#::::::::::::::::::::::::::::::::::::::::::::::::::
                           171
   25.00025                                                          500.005

Item no.:   2 Item name: CQ2 Scale: UNI
Experts
  1           [------------*---------------------]
  2                              [---------------*-----------------------------]
  3        [---------*---------------]
  4                [--------*------]
  5 *]
  6                     [---------------*---------------------]
Global                           [==============*============================]
Equ [========================*=====================================]
Real::::::::::::::::::::::::::::::::::::::::#::::::::::::::::::::::::::::::::::
                                           267
   0.1                                                               500.005

Item no.:   3 Item name: CQ3 Scale: UNI
Experts
  1 [*--]
  2 [-----*------]
  3          [---*----------------]
  4    [-*]
  5                              [---*----]
  6               [-------------------------------*-----------------------------]
Glo [=====*======]
Equ [===========*=================================================]
Real:::::::::::#::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
             8796000
   1000010                                                          50000500

Item no.:   4 Item name: CQ4 Scale: UNI
Experts
  1      [------------------------------*-------------------------------------]
  2                [-------------------*------------------]
  3 [-*----]
  4  [----*------]
  5          [----------*----]
  6      [-----------*------------------]
Global             [==================*==================]
Equ [================*============================================]
Real:::::::::::::::::::::::#::::::::::::::::::::::::::::::::::::::::::::::::::
                        6.5100002
```

0.5

```
20.0002
Item no.:    5 Item name: CQ5 Scale: UNI
Experts
  1 *--------------]
  2                [-----------------------------*------------------------------]
  3 *]
  4 [----*--------]
  5 *]
  6  [----*-------]
Global             [=============================*==============================]
Equ [==*============================================================]
Real::#::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
     0.175
    9.9999997E-005                                                     5.0000501

Item no.:    6 Item name: CQ6 Scale: UNI
Experts
  1                                 [--------------*----------------------------]
  2        [--*----------]
  3     [-----------*---------------]
  4               [-----*--------]
  5  [--*----]
  6 *--------]
Global     [==*=========]
Equ [============*====================================================]
Real:::::::::::#:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
               601
    50.0005                                                            3500.0349

Item no.:    7 Item name: CQ7 Scale: UNI
Experts
  1 [*-]
  2  [-*-]
  3                    [---------------------------------*-----------------------]
  4              [-----*----------------------]
  5             [*-]
  6 *]
Glob [=*=]
Equ [==========*=======================================================]
Real::::::::::::::#::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
                  569
    30.000299                                                          2750.0276

Item no.:    8 Item name: CQ8 Scale: UNI
Experts
  1                                  [---------------*----------------]
  2                                       [---------------------------*---------]
  3                                     [------*-----]
  4                                  [---------------*----------------]
  5                                     [-*------]
  6 [---------*-----]
Global                                    [===========================*=========]
Equal   [======================================*=============================]
Real::::::::::::::::::::::::::::::::::::::::#::::::::::::::::::::::::::::::::::::
                                            62
    25.00025                                                           95.000954
```

```
Item no.:   9 Item name: CQ9 Scale: UNI
Experts
  1 [-------------*------------------------------]
  2               [---------------*--------------]
  3                      [-----------------*--------------]
  4 [-------------*---------------------]
  5                                  [-----*-----------]
  6                               [--------------*------------------------------]
Global            [===============*==============]
Equal [================================*===============================]
Real:::::::::::::::::::::::::::::::::::::::::::::#:::::::::::::::::::::::::::::::
                                                 20
    5.0000501                                                         30.000299

Item no.:  10 Item name: CQ10 Scale: UNI
Experts
  1                                   [------------------------*----------]
  2                          [------------------------*-------------------------]
  3                          [---------------------*----------------------]
  4                                   [---------------*-------------------------]
  5                          [------------------------*-------------------------]
  6 [-----------------------------------------*------------------------]
Global                       [========================*=========================]
Equal          [======================================*========================]
Real::::::::::::::::::::::::::::::::::#::::::::::::::::::::::::::::::::::::::::::
                                      70
    50.0005                                                           95.000954

Item no.:  11 Item name: Q1 Scale: UNI
Experts
  1 [----*---]
  2                                          [----------------*------------]
  3                                          [--------------------*-------------]
  4                        [------------*---------------------]
  5                                  [-----------------------------*--------]
  6 [----*-----------------------------------]
Global                                       [================*============]
Equa [=========================================*============================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    4000.04                                                           90000.898

Item no.:  12 Item name: Q2 Scale: UNI
Experts
  1  [---*-----]
  2 *-]
  3                                         [---------------------*-------------]
  4                [------------*---------------]
  5                     [-------*-------------------]
  6  [-*-]
Glo *=]
Equ [==============*======================================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    300.00299                                                         95000.953
```

```
Item no.:  13 Item name: Q3 Scale: UNI
Experts
  1        [---*-----------]
  2                         [-----------------------*-----------]
  3 [----------*------------------------]
  4               [---------*----------]
  5                                                         [---*---------------]
  6                 [-------------------*-----------------------]
Global                      [=======================*===========]
Equal  [=========================*========================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    5000.0498                                                         99000.992

Item no.:  14 Item name: Q4 Scale: UNI
Experts
  1               [------*------]
  2                                                            [-------*--------]
  3                  [----------------------------*-------------------------]
  4                              [----------------*------------]
  5       [------*------------------------------------]
  6 *--]
Global                                                         [=======*========]
Equ [==============================*=========================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    500.005                                                           90000.898

Item no.:  15 Item name: Q5 Scale: UNI
Experts
  1 [------*---------]
  2        [------------------*------------]
  3   [--------*---------------------------]
  4   [-------------*-------------]
  5 [---------------*----------------------]
  6                                        [----------------------*-------------]
Global     [==================*============]
Equa [================*===================================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    7000.0698                                                         90000.898

Item no.:  16 Item name: Q6 Scale: UNI
Experts
  1 [----*----------]
  2      [---------*-------------------]
  3      [--------------------*-------------------------------------------------]
  4        [------------*-----------]
  5 [--*------]
  6 [----*----------]
Global   [=========*===================]
Equ [=========*================================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    2000.02                                                           40000.398
```

```
Item no.:  17 Item name: Q7 Scale: UNI
Experts
  1  [-----*--------]
  2 [--*------]
  3    [-------------------*--------------------------------------]
  4   [----*----------]
  5                        [------------*--------------------------------------]
  6 [--*------]
Glo [==*======]
Equ [=======*=====================================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    2000.02                                                          60000.602

Item no.:  18 Item name: Q8 Scale: UNI
Experts
  1                                                     [--------*-------------]
  2              [-------------------------*---------------]
  3                                [-----------------*-----------------]
  4                                         [--------*---------------------]
  5                  [-----------------*-----------------]
  6 [---*-----------------]
Global           [=========================*===============]
Equa [=========================================*============================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    5000.0498                                                        90000.898

Item no.:  19 Item name: Q9 Scale: UNI
Experts
  1     [-------------*------]
  2                       [--------*------------------]
  3 [--------*------------------]
  4                                               [------------*---------------]
  5                                [-----------------*---------]
  6 [--------*---------]
Global                    [========*==================]
Equa [======================*============================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    5000.0498                                                        88000.883

Item no.:  20 Item name: Q10 Scale: UNI
Experts
  1 |
  2      [-----------------------*------------------------------]
  3   [-----*---------------------------]
  4               [--------------*------------------------]
  5                                     [-----------------*--------------------]
  6 |
Global   [=======================*==============================]
Equ [================*==================================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    2.00002                                                           25000.25
```

```
Item no.:  21 Item name: Q11 Scale: UNI
Experts
  1                                            [----------*---------]
  2                                       [------------------*--------------------]
  3     [------*----------]
  4                                   [-------------*-------------------]
  5  |
  6         [--*-----]
Global                                    [==================*===================]
Equ  [=======================*========================================]
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     50.0005                                                         70000.703

Item no.:  22 Item name: Q12 Scale: UNI
Experts
  1  [---*---------]
  2            [----------------*-----------------------]
  3     [-------*----------------]
  4                                [-----------------------*-------------------]
  5  [----------*----------------]
  6                   [----------------*----------------------]
Global         [===============*=======================]
Equ  [=====================*==========================================]
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     3000.03                                                          48000.48

Item no.:  23 Item name: Q13 Scale: UNI
Experts
  1  |
  2       [--------------*--------------]
  3 [*-----------]
  4          [--------------------*-------------------------------------------]
  5  []
  6  |
Global    [==============*==============]
Equ  [*========================================================]
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     5.0000501                                                        25000.25

Item no.:  24 Item name: Q14 Scale: UNI
Experts
  1     [------*----------]
  2 [----*------]
  3  [---------*--------------]
  4              [-----------------*-------------------]
  5        [----------------------------*---------------------------------------]
  6        [------*----------------------]
Glob  [====*======]
Equa  [============*============================================]
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     1000.01                                                           50000.5
```

```
Item no.:  25 Item name: Q15 Scale: UNI
Experts
  1                                                  [------------*----------]
  2 [---*-]
  3 [----*----]
  4                  [------------------*------------------]
  5                                 [----------------*------------------------]
  6       [----*-------]
Glo [===*=]
Equa [=================*===================================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    2000.02                                                         90000.898

Item no.:  26 Item name: Q16 Scale: UNI
Experts
  1 |
  2            [-----*-----------]
  3 *]
  4       [------*-------]
  5          [------------------*---------------------------------------------]
  6                                    [--------------*----------------------]
Global         [=====*===========]
Equ [==============*=====================================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    250.0025                                                           50000.5

Item no.:  27 Item name: Q17 Scale: UNI
Experts
  1 [-------*-------]
  2         [-------*----------------]
  3 [---------------*----------------]
  4       [------------*-----------------------------]
  5 [---------------*------------------------]
  6                 [------------------------*--------------------------------]
Global      [=======*================]
Equ [================*==============================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    5000.0498                                                          50000.5

Item no.:  28 Item name: Q18 Scale: UNI
Experts
  1 |
  2 [-*-----]
  3 |
  4                 [--------------------*------------------------------------]
  5 *-]
  6 |
Glo [=*=====]
Equ *===========================================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    0.050000001                                                       14000.14
```

```
Item no.:  29 Item name: Q19 Scale: UNI
Experts
  1 |
  2               [----------------*---------]
  3               [--------------------------*---------------------]
  4        [----*-------------]
  5           [---------------*-----------------------------------------------]
  6      [----*----]
Global            [================*=========]
Equ [================*=============================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    150.0015                                                         70000.703

Item no.:  30 Item name: Q20 Scale: UNI
Experts
  1           [-----------*------------]
  2     [---*----]
  3 [-------*-----------------]
  4                               [-----------------*---------------------]
  5                                                 [------------*------------]
  6                       [----------------*---------------------]
Global  [===*====]
Equal [==========================*========================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    5000.0498                                                        90000.898

Item no.:  31 Item name: Q21 Scale: UNI
Experts
  1                                 [--------------------*--------------------]
  2    [----*--------]
  3 |
  4              [---------------*-----------------]
  5 |
  6 |
Global [====*========]
Equ [===*=============================================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    9.9999997E-005                                                   42000.422

Item no.:  32 Item name: Q22 Scale: UNI
Experts
  1 |
  2   [--*--------]
  3 |
  4                 [-----------------------------*--------------------------]
  5 |
  6 *-]
Globa [==*========]
Equ *==============================================================]
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
    0.40000001                                                       35000.352
```

APPENDIX J

Variable Names and Definitions

Causal/Contributory Factors	Description
Procedural Deviation	Describes deviation from corporate, organizational, or regulatory procedures by an operator
Inadequate Supervision (Climate)	Describes absence of supervisory input or pressure from the supervisory or organizational level to perform tasks motivated by factors inconsistent with a climate of compliance or safety.
High Workload	Describes workload levels that affect individual performance negatively by amplifying inattention or lack of focus
Organizational/Regulatory Factors	Describes abnormal factors in the organizational or regulatory domain such as those included in the model that are of sufficient influence so as not to be discounted as a potential contributing factor if an accident or incident were to occur
Complex Intersections	Describes the probability that an airport has complex intersections - those taxiway or runway intersections where number of intersecting surfaces, signage, lighting, or aircraft geometry may combine to create unusually high confusion for operators
Irregular Signs/Markings	Describes irregular or non ICAO-compliant signs and markings on an airport surface
Construction	Describes the presence of construction in the airport movement area
Non-Towered Airport	Describes whether an airport has an operational ground control or not
Part 139	Describes whether an airport is subject to compliance with 14CFR Part 139 regulatory compliance or not; 14 CFR Part 139 requires FAA to issue airport operating certificates to airports that--- Serve scheduled and unscheduled air carrier aircraft with more than 30 seats; Server scheduled air carrier operations in aircraft with more than 9 seats but less than 31 seats; and The FAA Administrator requires to have a certificate.
Conflicting Vehicle Traffic	Describes the presence of conflicting vehicle traffic in the airport movement area
Operational Environment Factors	Describes abnormal factors in the operational environment domain such as those included in the model that are of sufficient influence so as not to be discounted as a potential contributing factor if an

	accident or incident were to occur
Lost S/A	Describes a loss of flight crew or vehicle operator situational awareness, meaning that inappropriate mental representations are activated in spite of real world evidence. People then act "in the wrong scene," and seek cues confirming their expectations, a behavior known as confirmation bias.
Failed to Hold Short	Describes an aircraft or vehicle failure to remain in place behind a hold short line despite instructions by a controlling authority or regulatory requirement to do so
Task Saturation	Describes an operational condition of no awareness of input from various sources, so decisions might be made with incomplete information and the possibility of error increases
Human Factors	Describes abnormal factors in the human factors domain such as those included in the model that are of sufficient influence so as not to be discounted as a potential contributing factor if an accident or incident were to occur
Precipitation	Describes presence of precipitation at the airport surface
Sun Glare	Describes the presence of sun glare that interferes with vision of ATC, flight crew, or ground vehicle operator
Darkness/Twilight	Describes probability of operation during darkness or twilight
Restrictions to Visibility	Describes reduction in visibility at the airport surface by smoke, haze, fog, mist, or other phenomena that reduces visibility to under 1500 RVR
Weather Factors	Describes abnormal factors in the weather domain such as those included in the model that are of sufficient influence so as not to be discounted as a potential contributing factor if an accident or incident were to occur
Radio Congestion	Describes occurrence of radio congestion that requires operators to initiate multiple calls or wait to call or respond to ATC instruction
In-Cockpit Technology (Moving Map w/ Ownship. Etc.)	Describes the presence of technology in a flight deck or vehicle that displays the airport layout or combination of layout and vehicle or aircraft position in real time

ATC Equipment Failure	Describes the operational failure of ATC equipment such that normal communications or direction is interrupted
Mechanical Failure	Describes aircraft mechanical systems failure
Engineering/Technological Factors	Describes abnormal factors in the engineering or technological domain such as those included in the model that are of sufficient influence so as not to be discounted as a potential contributing factor if an accident or incident were to occur

APPENDIX K

BBN Model Details

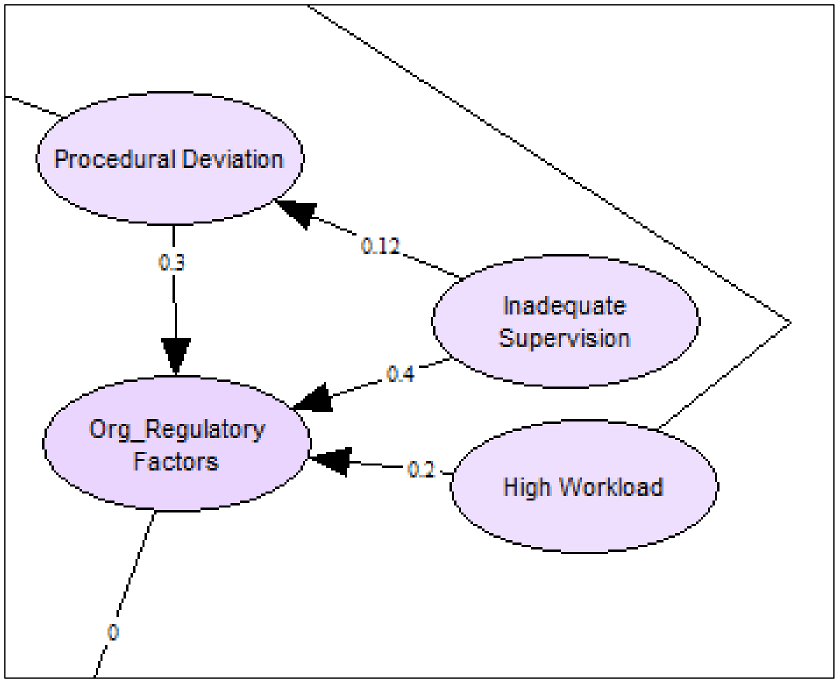


Figure 40. Compact model, organizational and regulatory domain.

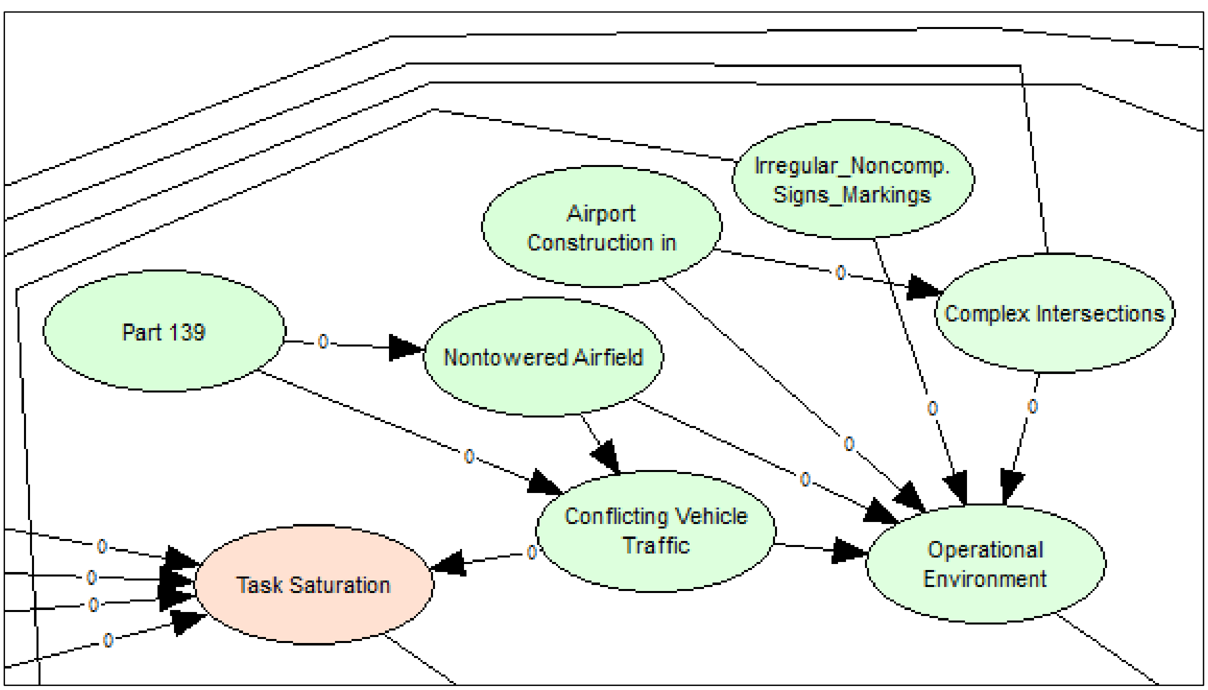


Figure 41. Compact model, operational environment domain.

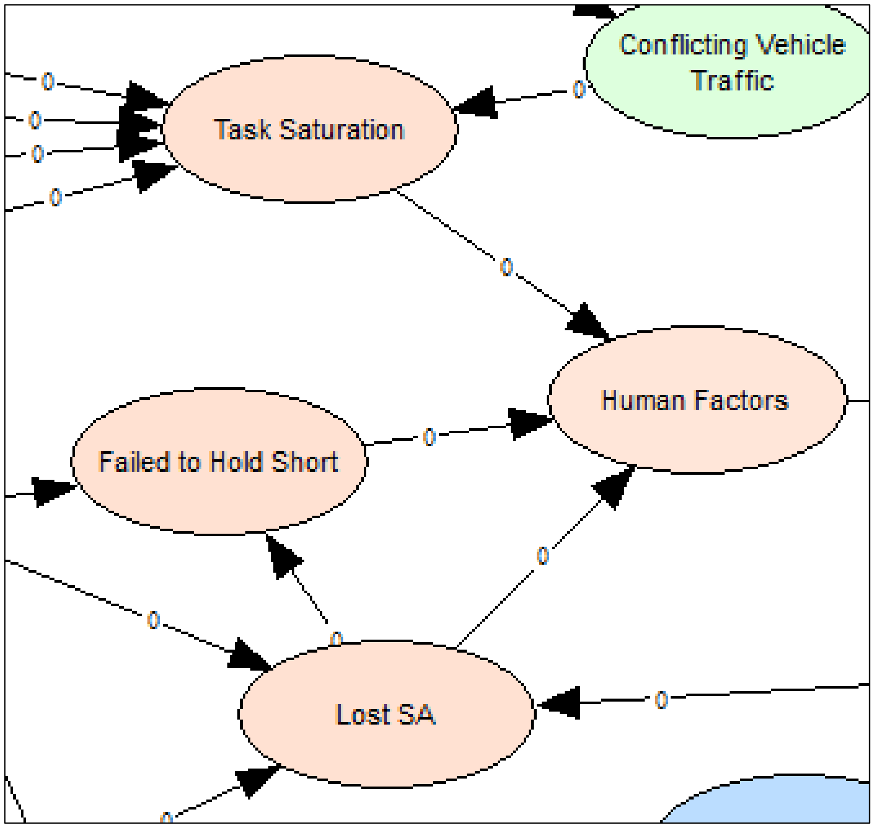


Figure 42. Compact model, human factors domain.

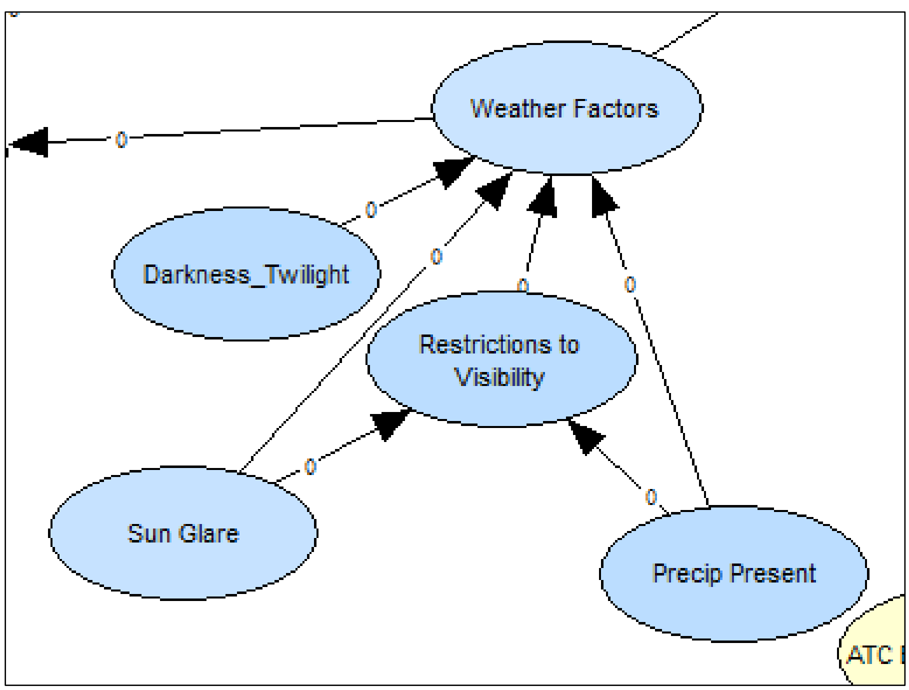


Figure 43. Compact model, weather domain.

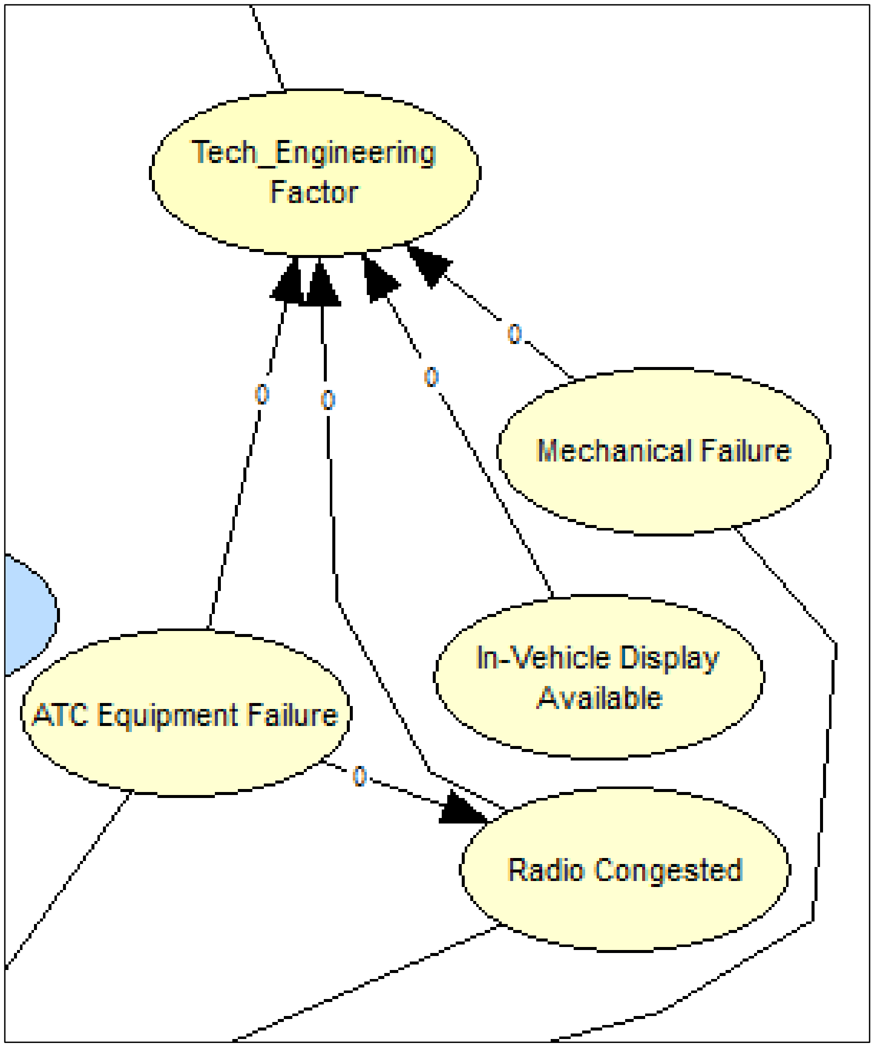


Figure 44. Compact model, technological and engineering domain.

APPENDIX L

BBN Rank Correlation Matrix

	RI	Org/ Regulatory	Tech/ Engineering	Weather	Human	Ops. Environment	Radio Congestion	Irregular Signs	A/P Construction	Darkness	Visibility	Precip	Part 139
RI	1.000	0.181	0.209	0.577	0.705	0.225	0.038	0.076	0.193	0.129	0.467	-0.017	-0.021
Org/Regulatory	0.181	1.000	0	0	0.187	0	0	0	0	0	0	0	0
Tech/Engineering	0.209	0	1.000	0	0.389	0	0.382	0	0	0	0	0	0
Weather	0.577	0	0	1.000	0.349	0	0	0	0	0.218	0.800	-0.029	0
Human	0.705	0.187	0.389	0.349	1.000	0.167	0.069	0.096	0.128	0.079	0.283	-0.011	-0.017
Ops. Environment	0.225	0	0	0	0.167	1.000	0	0.177	0.910	0	0	0	-0.086
Radio Congestion	0.038	0	0.382	0	0.069	0	1.000	0	0	0	0	0	0
Irregular Signs	0.076	0	0	0	0.096	0.177	0	1.000	0	0	0	0	0
A/P Construction	0.193	0	0	0	0.128	0.910	0	0	1.000	0	0	0	0
Darkness	0.129	0	0	0.218	0.079	0	0	0	0	1.000	0	0	0
Visibility	0.467	0	0	0.800	0.283	0	0	0	0	0	1.000	0.191	0
Precip	-0.017	0	0	-0.029	-0.011	0	0	0	0	0	0.191	1.000	0
Part 139	-0.021	0	0	0	-0.017	-0.086	0	0	0	0	0	0	1.000
Nontowered	0.021	0	0	0	0.015	0.093	0	0	0	0	0	0	-0.850
Complex Int.	0.210	0	0	0	0.157	0.924	0	0	0.800	0	0	0	0.000
Conf. Vehicle Traffic	0.031	0	0	0	0.036	0.084	0	0	0	0	0	0	-0.462
Mechanical Failure	0.249	0	0.750	0	0.467	0	0	0	0	0	0	0	0
In-Vehicle Display	-0.001	0	-0.190	0	0	0	0	0	0	0	0	0	0
ATC Equip Failure	0.037	0	0.465	0	0.066	0	0.800	0	0	0	0	0	0
Sun Glare	0.520	0	0	0.896	0.315	0	0	0	0	0	0.950	0	0
Inad. Supervision	0.131	0.840	0	0	0.115	0	0	0	0	0	0	0	0
Procedural Deviation	0.141	0.927	0	0	0.121	0	0	0	0	0	0	0	0
High Workload	0.195	0.213	0	0	0.333	0	0	0	0	0	0	0	0
Lost S/A	0.585	0	0	0.650	0.527	0.302	0	0.177	0.235	0.145	0.525	-0.020	0
Fail to Hold Short	0.488	0	0.605	0.248	0.700	0.117	0	0.069	0.091	0.056	0.202	-0.008	0
Task Saturation	0.247	0.468	0.080	0	0.388	0.008	0.171	0	0	0	0	0	-0.043
Incorrect Position	0.900	0.200	0.231	0.638	0.780	0.248	0.042	0.084	0.212	0.143	0.515	-0.019	-0.024

Parent-Child Corr.	
5 Highest Non Parent-Child Correlations	

	Nontowered	Complex Int.	Conflicting Vehicle Traffic	Mechanical Failure	In-Vehicle Display	ATC Equip Failure	Sun Glare	Inadequate Supervision	Procedural Deviation	High Workload	Lost S/A	Fail to Hold Short	Task Sat.	Incorr. Pos.
RI	0.021	0.210	0.031	0.249	-0.001	0.037	0.520	0.131	0.141	0.195	0.585	0.488	0.247	0.900
Org/Regulatory	0	0	0	0	0	0	0	0.840	0.927	0.213	0	0	0.468	0.200
Tech/Engineering	0	0	0	0.750	-0.190	0.465	0	0	0	0	0	0.605	0.080	0.231
Weather	0	0	0	0	0	0	0.896	0	0	0	0.650	0.248	0	0.638
Human	0.015	0.157	0.036	0.467	0	0.066	0.315	0.115	0.121	0.333	0.527	0.700	0.388	0.780
Ops. Environment	0.093	0.924	0.084	0	0	0	0	0	0	0	0.302	0.117	0.008	0.248
Radio Congestion	0	0	0	0	0	0.800	0	0	0	0	0.000	0	0.171	0.042
Irregular Signs	0	0	0	0	0	0	0	0	0	0	0.177	0.069	0	0.084
A/P Construction	0	0.800	0	0	0	0	0	0	0	0	0.235	0.091	0	0.212
Darkness	0	0	0	0	0	0	0	0	0	0	0.145	0.056	0	0.143
Visibility	0	0	0	0	0	0	0.950	0	0	0	0.525	0.202	0	0.515
Precip	0	0	0	0	0	0	0	0	0	0	-0.020	-0.008	0	-0.019
Part 139	-0.850	0	-0.462	0	0	0	0	0	0	0	0	0	-0.043	-0.024
Nontowered	1.000	0	0.400	0	0	0	0	0	0	0	0	0	0.037	0.023
Complex Int.	0	1.000	0	0	0	0	0	0	0	0	0.289	0.112	0	0.232
Conf. Vehicle Traffic	0.400	0	1.000	0	0	0	0	0	0	0	0	0	0.089	0.034
Mechanical Failure	0	0	0	1.000	0	0	0	0	0	0	0	0.800	0	0.275
In-Vehicle Display	0	0	0	0	1.000	0	0	0	0	0	0	0	0	-0.001
ATC Equip Failure	0	0	0	0	0	1.000	0	0	0	0	0	0	0.165	0.041
Sun Glare	0	0	0	0	0	0	1.000	0	0	0	0.586	0.224	0	0.575
Inad. Supervision	0	0	0	0	0	0	0	1.000	0.945	0	0	0	0.286	0.144
Procedural Deviation	0	0	0	0	0	0	0	0.945	1.000	0	0	0	0.301	0.155
High Workload	0	0	0	0	0	0	0	0	0	1.000	0	0	0.850	0.215
Lost S/A	0	0.289	0	0	0	0	0.586	0	0	0	1.000	0.373	0	0.646
Fail to Hold Short	0	0.112	0	0.800	0	0	0.224	0	0	0	0.373	1.000	0	0.538
Task Saturation	0.037	0	0.089	0	0	0.165	0	0.286	0.301	0.850	0	0	1.000	0.272
Incorrect Position	0.023	0.232	0.034	0.275	-0.001	0.041	0.575	0.144	0.155	0.215	0.646	0.538	0.272	1.000

Parent-Child Corr.	
5 Highest Non Parent-Child Correlations	

APPENDIX M

ASRS Reviewer Selection Criteria

Expert Initials: ______ Date: _____________

Resume or CV: ☐

Contact Information: ☐

Signed Consent Form: ☐

1. Independence

1 2 3 4 5 6 7 8 9 10

Notes:____

____

____

2. Diversity of Experience

1 2 3 4 5 6 7 8 9 10

Notes:____

____

____

3. Interest

1 2 3 4 5 6 7 8 9 10

Notes:____

____

____

4. Flight Experience

1 2 3 4 5 6 7 8 9 10

Notes:____

____

____

5. Familiarity with Current RI Mitigations

1 2 3 4 5 6 7 8 9 10

Notes:____

____

____