DEVELOPMENT OF A BAYESIAN BELIEF NETWORK RUNWAY INCURSION AND EXCURSION MODEL

Lawrence L. Green
Vehicle Analysis Branch (E401)
1 North Dryden Street (Mail Stop 451)
NASA Langley Research Center, Hampton VA 23681

Abstract
In a previous work, a statistical analysis of runway incursion (RI) event data was conducted to ascertain the relevance of this data to the top ten Technical Challenges (TC) of the National Aeronautics and Space Administration (NASA) Aviation Safety Program (AvSP). The study revealed connections to several of the AvSP top ten TC and identified numerous primary causes and contributing factors of RI events. The statistical analysis served as the basis for developing a system-level Bayesian Belief Network (BBN) model for RI events, also previously reported. Through literature searches and data analysis, this RI event network has now been extended to also model runway excursion (RE) events. These RI and RE event networks have been further modified and vetted by a Subject Matter Expert (SME) panel. The combined system-level BBN model will allow NASA to generically model the causes of RI and RE events and to assess the effectiveness of technology products being developed under NASA funding. These products are intended to reduce the frequency of runway safety incidents/accidents, and to improve runway safety in general. The development and structure of the BBN for both RI and RE events are documented in this paper.

Keywords
Runway Safety, Runway Incursion (RI), Runway Excursion (RE), Bayesian Belief Network (BBN)

Introduction
One focus area of the National Aeronautics and Space Administration (NASA), enabled through the Aviation Safety Program (AvSP) of the NASA Aeronautics Research Mission Directorate and in cooperation with the Federal Aviation Administration (FAA), is to improve aviation safety. The AvSP (AvSP, 2014) seeks to provide increasing capabilities to:

• predict and prevent safety issues;
• monitor for safety issues in-flight and lessen their impact should they occur;
• analyze and design safety issues out of complex system behaviors;
• analyze designs and operational data for potential hazards.

The AvSP explores hardware and software systems (technologies or products) that will operate in the Next Generation Air Transportation System (NextGEN, FAA 2014a). Runway safety is one thrust of investigation and research. The two primary components of runway safety are runway incursion (RI) and runway excursion (RE) events. A runway incursion is 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 [as defined by the FAA Office of Runway Safety (FAA, 2014b)]. A runway excursion is an incident involving only a single aircraft defined as a veer-off or overrun off the runway surface (FAA, 2013). In short, RI and RE are adverse ground-based aviation events that endanger crew, passengers, aircraft and perhaps other nearby people or property. An overrun is a runway excursion in which the aircraft departs the end of a runway; a veer-off is a
runway excursion in which an aircraft departs the side of a runway. As an RE event may occur on landing or takeoff, four types of RE events are possible: landing overrun, landing veer-off, takeoff overrun and takeoff veer-off. Some sources also include within RE events an aircraft attempting a landing that touches down in the undershoot area of the designated landing runway within the aerodrome perimeter (SKYbrary, 2014).

Model Development
In a previous paper, a statistical analysis (Green, 2013) of RI event data (FAA, 2014b) was conducted to ascertain its relevance to the top ten technical challenges of AvSP (AvSP, 2014). This research then served as the basis for developing a Bayesian Belief Network (BBN) intended to model the primary causes and contributing factors associated with RI events (Green, 2014). Additional literature study, described subsequently, identified many similarities between RI and RE events. A separate BBN model for RE events has now been developed, similar to that for RI events.

The BBN models presented herein are generic, system-level representations intended to model RI and/or RE events, and to capture the multi-dependencies (interactions) of causal and contributing factors from various problem domains. In general, the modeling steps undertaken include: 1) determining the causalities and cause-to-effect relations based on the historical risks and anticipated future risks from safety data/database and literature reviews, 2) constructing a baseline risk-based causal model as a BBN, 3) conducting Subject Matter Experts (SME) Knowledge Elicitation (KE) sessions to review the baseline model structure and to elicit the Conditional Probability Table (CPT) values for the baseline model without product insertions, and 4) inserting the NASA safety technologies/products into the model and eliciting CPT values with products included. The expected modeling results include 1) a quantification of the likelihood/probability of concerned aviation risks, 2) an assessment of the direct risk mitigation effectiveness of the NASA safety technologies/products, 3) a portfolio gap analysis and 4) a sensitivity analysis for risk drivers.

Although both RI and RE events may involve numerous contributing factors (e.g., airport layout, airport operations, weather, training and mechanical failures), RI events are more complex than RE events. RI events are “people intensive”, possibly involving multiple pilots, controllers, airport employees or contractors and perhaps other participants. There are also organizations (FAA, Airport Management, Aircraft Operators, etc.) behind each of the people directly involved in the runway safety events. RI events are also “communication (Comm) intensive”; several instances of two-party communications must simultaneously function properly in order to avoid problems. Two-party communications involve both the content and hardware transmission of information. Instances of two-party communications exist between all the people involved in the event (e.g., pilot to pilot, pilot to controller, controller to controller, controller to airport personnel). All runway safety events are short in duration and timing is critical; typical landing or takeoff reaction times are about 20 seconds (or less) and the event severity can easily escalate with just slightly different timing.

The SME panel unanimously agreed that developing a model for runway safety was much more difficult than developing one for in-flight operations. Two members of this SME panel have also participated in prior similar model reviews hosted by this NASA team for different applications. However, in this case, it was quite challenging to even achieve agreement on the basic structure of the BBN model among the NASA team and the SME panel. Numerous alternative models have been developed, discussed and discarded by the NASA team, either because they did not provide a satisfactory causal path, or because they were deemed to be too complex for use within the SME elicitation process for the purpose of portfolio assessment. The model proposed and discussed during the November 2013 session (Green 2014) was significantly changed by the SME panel at that time and then significantly changed again by the same SME panel (April 2014). Yet another SME review by the same panel (July 2104) altered the model structure further. However, through all the discussions and modifications, the runway safety model has been steadily improved and simplified.

Literature Review
Prior papers fully documented the statistical analysis of RI data (Green 2013) and the initial development of a model for RI events (Green 2014). As noted previously, the generic RI event model has been implemented as a
BBN model (Hayhurst, 2003 and Luxhoj, 2003a / 2012). Other similar efforts have been recently documented within a group working at NASA LaRC (Hayhurst, 2003; Luxhoj, 2003a, 2003b, 2012; Ancel and Shih, 2012; Shih, Ancel, and Jones, 2012). Among the significant findings in the literature on this topic were that the number of RI events is increasing with time (NASA, 2003) and intersecting runways are noted as the highest contributing factor. A recent U.S. Department of Transportation, Volpe Center (Cardosi, 2010) report shows that crossing the hold short line, entering the runway, and crossing a runway as the most likely types of RI events. A recent journal article (Torres, 2011) illustrates a dramatic increase in the number of RI reported in 2008 compared to previous years, with pilot deviations always being the largest source of these events. One FAA report (FAA, 2007) described the strong correlation among airport geometry, complexity, and various communication tools (including signage and runway markings) with RI events.

This paper documents recent changes (July 2014) to the RI portion of the BBN model and the development of a similar model for RE events. The RE event rate is quoted in several references: about 1 to 2 per million flights for the period 1990 through 2006 (NRL, 2010), up to 16 accidents and incidents per year during the period 1978 to 2008 (TRB, 2008), and 30 runway excursion accidents per year (Hindsight, 2014). The approach and landing phases of flight as having shown little improvement in safety over the last decade (up to 2008) and RE events are the third greatest source of aircraft crashes behind in flight loss of control and controlled flight into terrain (ATSB, 2008a). According to one source, the frequency or runway incursions is about half that of runway excursions, which may amount to 10 to 20 overruns and veer-offs each year (ATSB, 2008a); the data used in this paper suggests that severe RI (Category A / B) are about as common RE.

According to one source (NLR 2010), landing RE events are the most common, representing about 77% of all RE events. Some contributing factors are shared across the various types of RE events, noted above. The most common contributing factor associated with landing overruns is wet/contaminated runways, with long landings being the second most common contributor. Several other contributing factors are also noted (incorrect decision to land, speed to high, late/incorrect use of brakes, late/incorrect use of thrust reversers, aquaplaning, tailwind and being too high on approach). For landing veer-offs, the most common contributing factors include crosswind, nose wheel steering problems, collapsed landing gear, hard landing, tire failures and asymmetric power. The most common contributing factors to takeoff overruns include late abort/reject (after the speed V1 is attained) and takeoff mass too high/incorrect. The most common contributing factors for takeoff veer-offs include inadequate supervision of the flight (NLR 2010).

Numerous other references cite additional cause and contributing factors for RE events, which are summarized here and noted again later in this paper where they are particularly relevant. For example, the direct role of air traffic control (ATC) personnel in runway excursions was relatively small (NLR, 2010). Several other contributing factors, including communication/coordination/planning, poor decision making processes about landing or takeoff under adverse circumstances and approach/takeoff procedures are important (TRB, 2008). One contributing factor not previously mentioned is the inconsistent reporting of runway conditions and braking action at airports across the world (ATSB, 2008a). Numerous challenges exist for improving runway safety for existing airports (ATSB, 2008b). While contaminated runways (ice, snow, slush, wet or flooded) are a significant contributor to RE events, almost 47% of RE events occurred on dry runways (van Es, 2005). Takeoff runway excursions were likewise predisposed by a number of factors (SMS, 2009).

Some other papers also discuss other model development and application efforts, also aimed at improving runway safety. For instance, one reference describes an analysis tool to quantify risk and support planning and engineering decisions when determining runway safety area (RSA) requirements for various types and sizes of airports (TRB, 2008). The FAA is developing integrated risk models to forecast the risk and assess the impact of additional control measures at specific airports based on traffic volumes, complexity, and environmental factors (USDOT, 2012). Another study is taking a more novel and holistic approach to make sure that resources spent by airports to improve runway safety are actually used to address the most common types of RE events (Binghamton Univ, 2012). Yet another study employed human-in-the-loop simulations to evaluate traffic capacity at the Los Angeles International Airport (Barnett, 2010). An automated risk rating model for RI events was presented in another report (Cardosi, 2005). Another paper by the same author
explores what is known about the human errors and other factors that have been identified as contributing to runway incursions, and offers some error mitigation strategies (Cardosi, 2010). An example of a BBN devoted to RI events is given in another source (Goodheart, 2013). The RI and RE models could be joined together through a set of common definitions for accident and incidents, based upon the level of aircraft damage and passenger injuries (FSF, 2009) but this was deemed out of scope for the current models.

**Model Description**

A typical BBN consists of the model structure and the model content. The model structure consists of a set of relevant definitions, as well as the node names, the node states, the ordering of the defined states for each node to facilitate SME comment, the connecting link topology and the connecting link priority as they enter specific nodes (again, to facilitate SME comment). The model content consists of the sets of marginal and Conditional Probability Table (CPT) values. During the first phase of a typical BBN development cycle, NASA researchers develop (based upon database and literature search) and propose a model structure to an SME panel; the development step may take months to complete. Then, the various elements of the proposed structure are reviewed, modified and vetted by the SME panel. Once the model structure has been agreed upon and vetted by the SME panel, a CPT elicitation process (model population) is conducted by a facilitator on behalf of NASA to determine the appropriate model probability content. Some portions of the model review, modification, vetting and population can be conducted in parallel. Once the model has been populated, it is executed to obtain a set of baseline results and a baseline sensitivity analysis.

An SME panel consisting of four consultants was assembled to review the model structure and to populate its content. The SME panel included two pilots and two other aviation expert consultants. The preliminary BBN RI event model discussed in the prior paper (Green 2014, describing the SME session of November 2013) was substantially modified and simplified during the second SME panel review (April 2014), as shown in Exhibit 1 (please see next page). The nodes in Exhibit 1 are color coded to indicate associations among the various groups. Generally, the flow of specific contributing factors through causal paths is from left to right in the Exhibit. Numerous contributing factors items funnel together into the black node identified as “Primary Error State”. This node can be thought of as the event initiation and the start of the active mitigation phase. The RI Event severity, judged after the fact by a review board, includes the effects of any active mitigation actions that were applied (e.g., go-arounds and aborted takeoffs, and other evasive maneuvers). These active mitigations occur within specified collision scenarios and are subject to the available reaction time.

The SME panel validated many of the proposed definitions and most of the proposed model structure. However, the SME panel also provided significant clarification of several essential definitions within the RI event model. It is intended that the node names are suggestive of the states of each node; hence, limited clarifying information is presented about each of the nodes. Most of the nodes are binary, meaning they have only two possible states: present as an issue in RI events or not; where more states are present in a node, this will be made clear from the explanation of the node given subsequently. The goal of the SME elicitation is to provide probabilities for each of the possible states; for example, for the node “Airport Layout”, the SME goal is to determine the probability of that the Airport Layout is an issue or not in the RI event.
The reader should first observe the numbers in Exhibit 1, representing the three possible primary participants in an RI event: 1 is the pilot or pilots (orange node), 2 is the air traffic controller (ATC) or controllers (cyan node) and 3 is a airport or contractor vehicle driver (pink node). Each of these nodes has a number of other color-coded nodes with links pointing into these three primary nodes. Likewise, each of these three primary nodes have links pointing into the black node (Primary Error State). Starting with the green nodes (middle top), and moving counter-clockwise, the node descriptions of the RI event model follow:

**Airport Issues**

**Airport Layout.** The airport layout is an issue (FAA, 2007). This may include potentially confusing elements such as parallel runways (with spacing of less than 1000 feet), intersecting runways, and taxiways parallel to and near runways, numerous taxiways crossing runways instead of perimeter taxiways.

**Signs, Markings and Lighting.** The signs, markings and/or fixed equipment (e.g., lights) at the airport are deficient. This problem may be exacerbated under severe weather conditions when signs, etc. may be obscured from view (Green 2013).

**Airport Construction or RW/TW Closure.** Airport construction or runway/taxiway closure is an issue.

**Contamination Control.** Contamination control, generally related to rain, snow or ice, is an issue (Green 2013, Chow and Gulding, 2013).

**Airport Issues.** One or more of the issues within this grouping are present.
ATC HFACS Issues
Next are the purple nodes (upper left) which represent some of the most (as determined by the SME panel) Human Factors Analysis and Classification System (HFACS, Wiegmann and Shappell, 2003) issues.

ATC Cert Training Issues. Certification training for the ATC involved is an issue.

ATC OTJ Training Issues. On the job training for the ATC involved is an issue as a distraction (Cardosi, 2010).

ATC Mental or Physical State. The current physical or mental state of the ATC involved is an issue.

ATC HFACS Issues. One or more of the HFACS issues within this grouping is an issue for the ATC involved (Wiegmann and Shappell, 2003).

ATC Operational Issues
The reader will note that a similar purple HFACS group accompanies both the pilot and the vehicle driver, though in the latter case, the HFACS group only includes training. The cyan nodes (middle left) are other contributing factors that may influence the performance of the ATC.

Automation Interaction Issues (ATC). Automation interaction is an issue for the ATC involved (Rudisill, 1995).

Abnormal Air Traffic Volume or Complexity. The traffic volume or complexity at this airport, at this time is an issue. For example, if the average traffic volume is high, it may cause a significantly increased work load for controllers and/or pilots; if low, it may result in extended periods of inactivity for controllers (ALPA, 2007).

Staffing or Procedures Issues. The staffing level and/or work load management not appropriate for the situation is an issue, or the use of ambiguous or non-standardized ATC procedures is an issue (ALPA, 2007).

ATC Operational Issues. One or more of the issues within this grouping are present.

Two Party Communications
The next group (blue nodes) describe the state of the system-wide two-party communications.

Communication Content Issues. The completeness, correctness timeliness or complexity of communicated information is an issue. This may include the lack of a required usage for a call sign. Information may not have been transmitted at the appropriate time, i.e., it was not delayed. (Green 2013).

Comm Hardware Error. Comm transmission is an issue. This may occur when the Comm system fails to operate as expected and may include blocked (“stepped on” communications where one party cuts off the communications of another), partially blocked (garbled or inaudible Comm transmission), hardware limitations / malfunctions and/or faulty headset jacks or connections (Green 2013).

Two Party Communication Issues. Comm Content Issues or a Communications Hardware Error has resulted in a Two Party Communications Error and is an issue.
Pilot HFACS Issues
Another group of HFACS contributing factors is present for the pilot. The grouping is the same as before, although the relevant probabilities of these factors being an issue in an RI event may be different than for the ATC.

Pilot Cert Training Issues. Certification training for the pilot(s) involved is an issue.

Pilot OTJ Training Issues. On the job training for the pilot(s) involved is an issue as a distraction (Cardosi 2010).

Pilot Mental or Physical State. The current physical or mental state of the pilot(s) involved is an issue.

Pilot HFACS Issues. One or more of the HFACS issues within this grouping is an issue for the pilot(s) involved (Wiegmann and Shappell, 2003).

Other Pilot Operational Issues
Other contributing factors that may influence the performance of the pilot(s) involved are show in the orange nodes.

Inappropriate Aircraft Operations: Pilot operations of the aircraft, outside of the flight operational manual guidelines, is an issue causing the RI event.

Automation Interaction Issues (Pilot): Automation interaction is an issue for the pilot(s) involved (Rudisill, 1995).

Pilot Operational Issues: One or more of the issues within this grouping are present.

Driver Operational Issues
These issues include the two nodes below. (purple and pink nodes, top right)

Driver Training: The training of airport vehicle drivers is an issue (Cardosi, 2010).

Driver Operational Issues: One or more of the HFACS issues within this grouping is an issue for the vehicle driver(s) involved (Wiegmann and Shappell, 2003).

The preceding discussion covers all the nodes on the periphery of the left hand side of Exhibit 1. These are all the issues potentially present that enable the RI event to occur. The nodes and states on the right hand side of the Exhibit generically define a specific RI event, of which numerous types and combinations may occur. NASA would hope to be in a position to broadly address many, if not all, of these specific RI event types with technology injections. Starting with the black node of Exhibit 1 (middle right), the node descriptions follow.

Primary Error State: The primary error source is either Controller Error (typically, loss of oversight), Pilot Error (typically, failure to hold short of a runway without authorization), or Other (includes mechanical failure and Driver Error, i.e., a failure to hold short of a runway without authorization). The SME panel excluded from consideration in this model non-airport authorized vehicles and all pedestrians on the runway.

Collision Scenarios. The SME panel identified the most common collision scenarios: crossing in front of a aircraft on departure, crossing in front of an aircraft on arrival, or intersection events (crossing active runways) and other (everything else that leads to an RI event) (Cardosi, 2010).
**Reaction time.** this node has two states defined by the SME panel, short (eight seconds or less) and long (nine seconds or more)

**Final RI Event Severity.** The RI event severity as would be reported by the FAA, including the impact of Contributing Factors and Mitigating Actions is established here. The states enumerated by the SME panel are accident / near miss, or other. Although less severe RI event categories have been defined by the FAA, these were deemed out of scope for this BBN model because the SME panel could not provide sufficient discrimination among these less severe RI events.

This concludes the presentation of the RI event model. Likewise, the SME panel vetted many of the proposed definitions and most of the proposed model structure of the RE event model, shown in Exhibit 2. The SME panel again provided significant clarification of several essential definitions within the RE event model. Moreover, the SME panel suggested several simplifying structural changes to the model. The remainder of this section describes the current preliminary RE event model. The node name for each is presented along with some clarifying comments. Again, most of the nodes are binary with only two possible states (present as in issue in RE Events or not); where more states are present in a node, this will be made clear from the explanation of the node given subsequently.

**Exhibit 2.** The Runway Excursion Bayesian Belief Network.

**Airport Issues**
Beginning with green nodes in the upper right corner, and moving counter-clockwise, the nodes and states are below.
Approach and Departure Constraints. The physical or regulatory constraints on approach or departure trajectories for the airport in question are an issue.

Contamination Control. Contamination control (e.g., rain, snow or ice) for the airport in question is an issue.

Runway Length. The runway length is an issue. This may be due to prevailing wind conditions, runway maintenance, or an aircraft landing on a runway that is too short for safe operations.

Airport Issues: One or more of the issues within this grouping are present.

ATC HFACS Issues
As in the RI event network, the ATC HFACS Issues group (purple nodes) is repeated in the RE event network. The SME panel rated the relative importance of these contributing factors for the ATC as being of much less consequence for RE events than for RI events. Next, the cyan nodes are described:

ATC Operational Issues

Runway Assignment. The runway assignment provided by ATC is an issue. This may be due to prevailing wind conditions, runway maintenance, or unusual airport operations.

Runway Collision Avoidance. An RI event (typically failure to hold short of an active runway) has precipitated an RE event. This was noted by the SME panel as being an extremely rare occurrence.

Contribution to Unstabilized Approach. The ATC has provided instructions that contribute to an unstabilized approach.

Lack of Current Weather Information. The ATC involved have provided non-current weather information that contributes to an RE event.

Pilot HFACS Issues

Pilot Operational Issues. One or more of the issues within this grouping are present.

Pilot Operational Issues and Remainder Nodes

Inappropriate Aircraft Operations. Pilot operations of the aircraft, outside of the flight operational manual guidelines, is an issue causing the RE event.

Unstabilized Approach. The pilot(s) involved have failed to perform a stabilized approach.

Pilot Operational Issues. One or more of the issues within this grouping are present.

Aircraft Automation Issues (blue node). Automation interaction is an issue for the pilot(s) involved Rudisill, 1995).

Pilot Error. A pilot error has initiated an RE event.
Weather Issues (yellow node). Weather issues have contributed to, or caused, an RE event.

Mechanical Failure (pink node). Mechanical failure has contributed to, or caused, an RE event.

RE Event Initiated: This node simply states whether an RE event has been initiated or not.

The most recent SME session (July 2014) resulted in a fully vetted BBN model for both RI and RE events. Moreover, the SME panel elicitation of marginal and conditional probabilities has also been completed. Thus, all the ingredients for a fully vetted and fully populated set of baseline models have been obtained. In order to demonstrate how a BBN model would function, random values for all of the conditional probability tables have been inserted into the final RI model so that model operations can be simulated as shown in Exhibit 3. The node probabilities are shown overlaying the model structure from Exhibit 1. In this case, the marginal (leaf node) probabilities for all of the contributing factors have been set to zero, and the reaction time has been set to “long” (i.e., more than 8 seconds) for the collision scenario “crossing in front of departure”. The model indicates that the probability of an accident or near miss (red node at middle right) is about 22% for this random scenario.

Exhibit 3. Sample RI Model Execution (baseline scenario).
Exhibit 4. Sample RI Model Execution (worst-case scenario).

In Exhibit 4, the same example has been shown again, but with all the leaf node marginal probabilities set to 1 (all the issue are a certainty) and the reaction time has been set to short for the same collision scenario. In this random, worst-case, the probability of an accident or near miss is increased to about 75%.

Conclusions
Preliminary Bayesian Belief Network (BBN) models for Runway Incursion (RI) and Runway Excursion (RE) events have been developed. Numerous considerations surrounding the process of developing the preliminary models have been documented in this and a previous paper. The proposed RI and RE event models have been thoroughly reviewed by a Subject Matter Expert (SME) panel. Numerous changes to the model structure (definitions, node names, node states and the connecting link topology) were suggested by the SME panel. The structural details of the resulting BBN models for RI and RE events have also been documented within this paper. Finally, sample executions of the final RI model (July 2014), using random conditional probability tables have been presented for the baseline and worst-case scenarios; the resulting probability of an accident or near miss increases substantially for the worst-case scenario, compared with the baseline scenario.

Recommendations
It is recommended that the model structures presented herein and the CPT values developed by the SME panel be validated by comparison to available data. It is recommended that model structures presented herein be expanded to include the injection of technology products intended to improve runway safety and that SME input is used to characterize the impact of these technology products. It is also recommended that the resulting BBN for RI and RE events be used by NASA to generically model the causes of RI and RE events and to assess the effectiveness of technology products being developed under NASA funding.
References


Green


