Eleventh USA/Europe Air Traffic Management Research and Development Seminar (ATM2015)

Risk Analysis Process Tool for Surface Loss of Separation Events

Eric B. Chang
The MITRE Corporation
Center for Advanced Aviation System Development (CAASD)
McLean, VA
Contact email: echang@mitre.org

ABSTRACT

The MITRE Corporation's Center for Advanced Aviation System Development (MITRE/CAASD) made improvements to the Risk Analysis Process (RAP) Tool scoring methods used in quantifying the risk level for loss of separation events. These enhancements are designed to evolve the present scoring methods by using operational data for trend analysis and promoting increased safety through risk mitigation and management. The new changes aim to simplify the tool's use and eliminate any potential biases associated with it. A newly modified RAP Tool has been developed for future evaluation of events and it closely aligns with the current Federal Aviation Administration (FAA) Safety Management System (SMS) Risk Matrix. The tool will be used by the FAA to closer examine the risk involved in loss of separation events in order to better prioritize their mitigations.

I. INTRODUCTION

The Risk Analysis Process (RAP) tool was originally developed by the European Organization for the Safety of Air Navigation (EUROCONTROL) and used extensively by regulatory authorities in the European Union (EU). The United States (U.S.) Federal Aviation Administration's (FAA's) Air Traffic Organization (ATO) Safety and Technical Training (AJI) adapted the tool for use in the National Airspace System (NAS) and presently oversees the development of the RAP tool. The tool estimates the relationship between actions and consequences and quantifies these relationships. Two versions of the tool presently exist: 1) an Airborne RAP tool used to evaluate loss of separation events involving airborne aircraft with a Measure of Compliance (MOC) less than 66% and 2) a Surface RAP tool used to evaluate surface events with less than 6,000 feet separation. These events are collectively known as Risk Analysis Events (RAE's) and are subject to further analysis by the respective RAP tool. Both versions of the tool align closely and map to the FAA's Safety Management System (SMS) Risk Matrix.

RAE panels have used the RAP tools to score loss of separation events based on interviews, surveillance data, and narrative documentation associated with the events. The enhancements aim to simplify the tool's use and eliminate potential biases associated with it. Both versions follow a similar framework, but this paper describes the scoring format of the RAP tool for surface events.

II. METHODS: SEVERITY

A. *Separation and Closure Rate*

Under the current score format, the severity of a RAE is evaluated using three factors: separation, rate of closure, controllability, and weather.

Empirical data from the 167 reviewed RAE's was used to adjust the scoring of the separation loss and rate of closure sections. Due to low data availability of events evaluated by the RAP Tool, additional data was added from the AJI Office of Runway Safety's Runway Incursion (RI) database. Data from each event included approximate separation loss distances. Closure rate values were calculated using Airport Surface Detection Equipment, Model X (ASDE-X) surveillance data. An additional total of 120 Class A-C categorized runway incursion events, found over the same time span as the RAP events were contributed from the RI database. The events included for analysis from the RAP database and runway incursion database all had less than 6,000 feet of separation.

Distributions for 287 events of separation distance and closure rates were constructed and fit to a best fit model. Most of the events included in the distributions (97.5%) were classified as Class C runway incursions. Separation and closure rate values was constructed using the slant range. The sample mean for the separation distance was 2,774 feet with a standard deviation of 1,826 feet and a skewness of 0.11.

The Pearson Type III distribution was chosen to model the historical separation values due to its ability to model several distribution types while accounting for the skewness of the distribution. The application of the Pearson Type III is to establish percentiles for the separation and closure rate distributions based on the mean, standard deviation, and skewness for each sample. Figure 1 show the distribution of separation and the goodness of fit of the curves as measured by the CDF values.

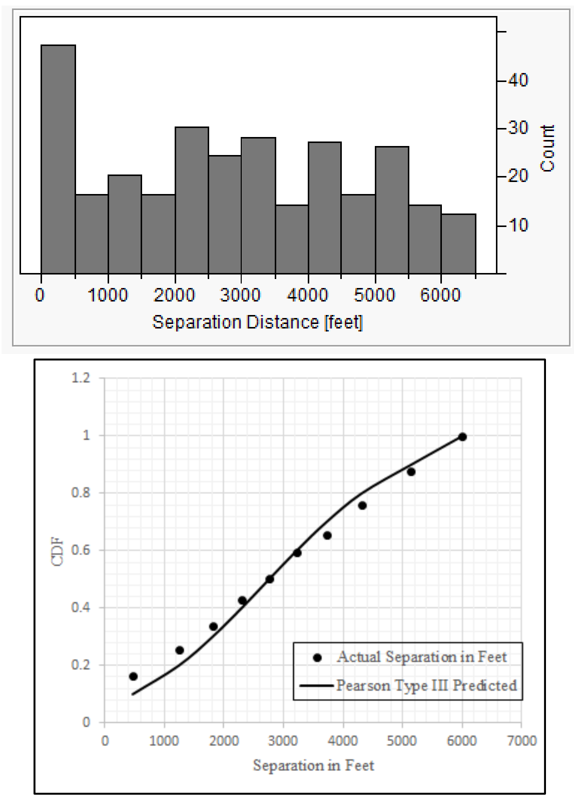

Figure 1: Separation Distance Distribution and CDF Comparison

A similar process was followed for analysis of the closure rate scores by fitting a Pearson Type III distribution to the closure rate values and mapping them to a 10 point scale. When analyzing the closure rate, both the lateral and vertical components were considered. The sample mean for the closure rate was 132.03 feet per second with a standard deviation of 93.09 feet per second and a skewness of 0.05.

When creating the closure rate distribution, a distinction was made between events involving aircraft with vehicles or pedestrians, and aircraft with other aircraft. In the databases used, 9.5 percent (26 out of 287) involved aircraft with vehicles or pedestrians.

Figure 2 shows the closure rate distributions for the aircraft-aircraft events and the goodness of fit curve as reflected by actual and predicted CDF values.

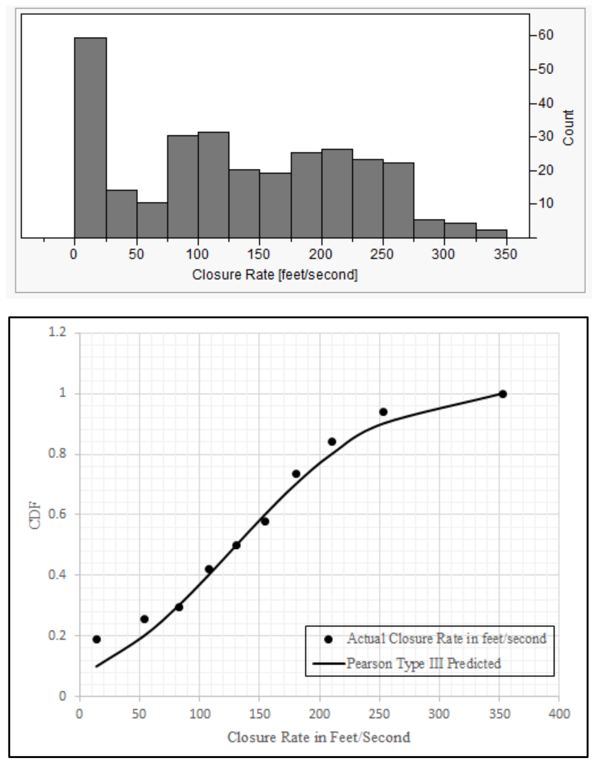

Figure 2: Closure Rate Distribution and CDF Comparison

Each of the separation and closure rate percentile scores mapped to the corresponding separation percentage. The scale was mapped linearly depending on the percentile rank. The separation section of the severity score carried a maximum score of 10. Table 1 and 2 displays the mapping of the scores for the two different separation bins.

TABLE 1: SEPARATION SCORES BY PERCENTILE RANK

Score	Percentile	Separation Distance Range [feet]
1	90 – 100	5139.9 - 6000
2	80 – 90	4309.5 - 5139.9
3	70 – 80	3718.1 - 4309.5
4	60 – 70	3217.6 - 3718.1
5	50 – 60	2754 - 3217.6
6	40 – 50	2294.1 - 2754
7	30 – 40	1806.5 - 2294.1
8	20 – 30	1241.4 - 1806.5
9	10 – 20	467.8 - 1241.4
10	0 – 10	0 - 467.8

TABLE 2: AIRCRAFT-AIRCRAFT EVENT CLOSURE RATE SCORES BY PERCENTILE RANK

Score	Percentile	Closure Rate [feet per second] Aircraft-Aircraft Events
0	0	Diverging
1	0 – 10	0 - 13.7
2	10 – 20	13.7 - 53.3
3	20 – 30	53.3 - 82.1
4	30 – 40	82.1 - 107.0
5	40 – 50	107.0 - 130.5
6	50 – 60	130.5 - 154.2
7	60 – 70	154.2 - 179.7
8	70 – 80	179.7 - 209.9
9	80 – 90	209.9 - 252.3
10	90 - 100	> 252.3

Based on the results from the databases for the two runway incursion event types, the closure rate values for aircraft with vehicle or pedestrian events were higher than multiple aircraft events for the same closure rate score. Using the same method of mapping the scores for both event types would not necessarily result in an elevated severity score for aircraft with vehicle or pedestrian events as compared to an event with the same closure rate involving multiple aircraft. Having a higher score for the same closure rate value for aircraft with vehicle or pedestrian events is due the higher level of risk associated with the event. In an events involving multiple aircraft, both parties have greater control to avoid a collision, compared with vehicle and pedestrian events, which can be more catastrophic in result at the same closure rate.

The higher closure rate values for the same score for aircraft with vehicle or pedestrian events compared to multiple aircraft events result from not having a sufficient amount of events. As an interim solution for aircraft with vehicle or pedestrian events, five percentile groups were used instead of ten, and the percentiles were rounded to ensure that the closure rates for the same score would be lower. As more data becomes available and if the mean closure rate values for aircraft with vehicle or pedestrian events becomes lower than multiple aircraft events, the same type of mapping method can be used. If the mean closure rate remains higher, the scoring scales will need to be adjusted in order to accurately evaluate the severity of the event. Table 3 shows the rounded score values, which only has five possible scores adjusted for vehicle-pedestrian events.

TABLE 3: AIRCRAFT-VEHICLE/PEDESTRIAN EVENT CLOSURE RATE SCORES BY PERCENTILE RANK

Score	Percentile	Closure Rate [feet per second] Aircraft – Vehicle/Pedestrian Events
0	0	Diverging
2	0 - 15	0 - 47.1
3	15 - 25	47.11 - 79.9
5	25 -50	79.9 - 137.9
8	50 - 75	137.9 - 194.6
10	75 – 100	> 194.6

B. *Controllability*

Controllability indicates the various actions or barriers in place in the NAS that either prevent or mitigate an event on the surface. Human-in-the-loop (HITL) simulations performed in [2] evaluated the importance of various controls in preventing potential events. The study simulated known possible events and determined how often the control was able to prevent the impending event from occurring. The simulations determined the following controls as having a role in calculating the risk depending on the type of event:

ATC Controls:

- Conflict Detection (On Time, Late, Not Detected)
- Plan (Correct, Inadequate, No Plan)
- Execution of Plan (Correct, Inadequate, No Execution)
- Recovery (Correct, Inadequate, No Recovery)
- See & Avoid Pilot Decision (Triggered, Not triggered)
- Crossing Traffic Location

Pilot Controls:

- Execution of Plan (Correct, Inadequate, No Execution)
- Recovery (Correct, Inadequate, No Recovery)
- Actions following TCAS or See & Avoid (Correct, Insufficient, Incorrect)

NAS Controls:

- Ground Safety Net Detection (Correct, Malfunction, Not Present)
- Runway Guard/Safety Lights Function (Correct, Malfunction, Not Present)
- Airport Signage/Markings Function (Correct, Malfunction, Not Present)

When scoring the controllability of the event, the RAP panels decide the level to which the controls were broken. Each control carries a point designation depending if the control was completely or partially broken. Complete control breakdowns occur when the control is not used or not followed. For example, no action taken or no recovery used. Partial control breakdowns are scored for controls that are only partially correct when executed. For example, inadequate action taken or inadequate recovery.

The number complete and partial control breakdowns are totaled and given an assigned point value. Complete control breakdowns are given a point value of 1 and partial control breakdowns are given a point value of 0.5. The sum of the score is divided by the total number of controls that are in place or available during an event, as shown by the following equation:

$$\frac{\sum Complete\ Breakdowns + \frac{1}{2}\sum Partial\ Breakdowns}{\sum Controls\ in\ Place} \quad (1)$$

The result of the equation provides a percentage of the available controls that are broken. The percent of control breakdowns is then mapped to a 0-50 risk score consistent giving more weight to the controllability section than the separation and closure rate sections. This is because the controls have human influence over the situation and are designed to prevent and mitigate loss of separation events, therefore they have more influence on determining the risk of the situation. The mapping of percentages to risk score is linear, weighting each control breakdown equally. Table 4 displays the controllability percentage mapping to its respective risk score.

TABLE 4: CONTROLLABILITY SCORING SCHEME

Control Breakdown Percentage	**Score**
0%	0
1 – 10%	5
>10 – 20%	10
>20 – 30%	15
>30 – 40%	20
>40 – 50%	25
>50 – 60%	30
>60 – 70%	35
>70 – 80%	40
>80 – 90%	45
>90 – 100%	50

C. Weather

Weather was often cited as a factor in the event narratives of past events. It was determined that weather was best suited to be included in the severity section using actual data to determine a risk score.

The International Civil Aviation Organization (ICAO) Runway Incursion Severity Classification (RISC) Version 2.0 created a model in [3] for evaluating the risk contributions of visibility to runway incursions. Each category weights the Runway Visual Range (RVR) or visibility by day or night. Tables 5 and 6 show the scores for each ceiling and RVR or visibility combination during day and nighttime, respectively. The scores assigned are on a 0-10 scale consistent with the other severity sections.

TABLE 5: DAYTIME RVR/VISIBILITY CEILING FACTOR

RVR or Visibility	Ceiling >1000 feet	Ceiling 500 feet	Ceiling 200 feet	Ceiling 100 feet
>3 Miles	0	1	3	10
1 Mile = 6000 feet	5	5	5	10
¾ Mile or 4500 feet	8	8	8	10
½ Mile or 3000 feet	10	10	10	10
¼ Mile or 1500 feet	10	10	10	10

TABLE 6: NIGHTTIME RVR/VISIBILITY CEILING FACTOR

RVR or Visibility	Ceiling >1000 feet	Ceiling 500 feet	Ceiling 200 feet	Ceiling 100 feet
>3 Miles	3	4	5	10
1 Mile = 6000 feet	8	8	8	10
¾ Mile or 4500 feet	10	10	10	10
½ Mile or 3000 feet	10	10	10	10
¼ Mile or 1500 feet	10	10	10	10

For each ceiling and RVR/visibility combination, the scores between day and night differ by a factor of 3 with a maximum possible risk score of 10. During both day and night, any RVR/visibility combined with a ceiling less than 100 feet is given the maximum score. During the day any situation where the RVR/visibility is less than ½ mile or 3,000 feet, the maximum score is given. This is the case for any condition during nighttime less than ¾ mile or 4,500 feet RVR/visibility. In the other combinations of RVR/visibility and ceiling, the scores exponentially increase as the RVR or visibility conditions worsen.

A similar scoring scheme was adopted for evaluating the weather conditions in the RAP Tool. In order to give the user flexibility with the input, the above tables were fit to logarithmic functions creating a dynamic risk scale that is adjustable depending on the weather conditions. Since the risk score remained the same for ceiling values less than 100 feet, curves were only fit for ceilings of 200 feet, 500 feet, and 1,000 feet categories. Equation 2, 3, and 4 shows the equations for each weather severity risk score by the logarithmic curve. The same curve equations are used to calculate the night weather risk score, by adding three to the equation result, similar to the scores developed by ICAO. An additional 3 points is added to the weather severity score during nighttime conditions.

Ceiling>1000 feet

$$Weather\ Severity = 5.915 - 5.609 \cdot \log(RVR) \qquad (2)$$

500 feet < Ceiling ≤ 1000 feet

$$Weather\ Severity = 6.147 - 5.004\ \cdot \log(RVR) \qquad (3)$$

Ceiling ≤ 500 feet

$$Weather\ Severity = 6.612 - 3.793 \cdot \log(RVR) \qquad (4)$$

D. Severity Calculation

The total severity score is calculated as a summation of the separation, closure rate, controllability, and weather if it is a factor. Depending on the total severity score, the score is mapped to the SMS Risk matrix, which determine risk levels.

III. METHODS: REPEATABILITY

A review of the repeatability sections and descriptive narratives of the 167 RAE's in the dataset revealed that not all causal factors contributing to loss of separations were being addressed or scored in the final risk calculation.

A review was conducted of the event summaries from 167 RAE's and 369 narratives from the AJI Runway Safety Runway Incursion Database. The intent was to capture contributing factors to an event and add new factors to the RAP Tool.

Based on the review of both data sets, 36 causal factors were identified in addition to the four existing repeatability sections from the RAP Tool. Based on the analysis of the RAE and Runway Incursion data sources, the original list of causal factors was reduced to 10. This eliminated any overlapping factors that were similar to each other. Table 7 provides the final list of factors and their number of occurrences during the time period analyzed.

TABLE 7: LIST OF RAP TOOL REPEATABILITY FACTORS

Factor	Data Source	Rate of Occurrence
ATC Procedures	RAE	**0.59%**
Pilot Procedures	RAE	**0.00%**
ATC Equipment	RAE	**1.19%**
Pilot Equipment	RAE	**0.00%**
ATC Human Resources	RAE	**0.00%**
Pilot Human Resources	RAE	**0.00%**
ATC Human Involvement	RAE	**67.26%**
Pilot Human Involvement	RAE	**44.64%**
Airport Construction	RI	**0.81%**
Runway/Taxiway Closure	RI	**0.54%**

Since the rates derived by the analysis of the RAE and RI datasets were determined by the event narratives, the rates of occurrences may not be accurate because certain details regarding the event may have been omitted from the event narrative.

The sample event rate includes several factors that had an occurrence rate of 0 percent, but does not necessarily mean they are not likely to occur in the future. Therefore, instead of using rates to determine repeatability, a similar scale to the original RAP Tool repeatable model is used, keeping the weighting of the degree of the causal factor consistent. Table 8 displays the list of repeatability factors that were inherited from the EUROCONTROL version of the tool and developed from causal factor analyses along with their associated point value.

TABLE 8: RAP TOOL REPEATABILITY FACTOR POINT VALUE

Factor	Point Value
Procedure Design Issue	12
Procedure Implementation Issue or Lack of	8
No Procedure Issue	0
Equipment Design Issue	12
Equipment Implementation Issue or Lack of	8
No Equipment Issue	0
Human Resources Management Design	12
Human Resources Management Implementation or Lack of	8
No Human Resources Management Issue	0
Airport Geometry Deficiency	12
Airport Geometry Insufficient	8
No Airport Geometry Issues	0
Non-Systemic/Human Involvement Issues with Contextual Conditions	12
Non-Systemic/Human Involvement Issues without Contextual Conditions	8
No Non-Systemic/Human Involvement Issues	0
Airport Construction with Contextual Conditions	12
Airport Construction without Contextual Conditions	8
No Airport Construction Issues	0
Runway Closure with Contextual Conditions	12
Runway Closure without Contextual Conditions	8
No Runway Closure Issues	0

The final section of the repeatability section is the window of opportunity. In this section, the user selects a combination of selections describing the current state of the traffic and the amount of workload that was present for the controllers.

EUROCONTROL studies have shown that repeatable errors are most likely to occur during normal operations and as a part of daily routine. As a result, a higher risk is associated with traffic and workload conditions that occur most often. Table 9 shows the window of opportunity scoring.

TABLE 9: RAP TOOL WINDOW OF OPPORTUNITY SCORING

Window of Opportunity	**Daily Routine**	**Workload Peak**	**Emergency Situations**
Normal	7	5	3
Degraded	6	4	2
Exceptional	3	2	1

IV. RISK CALCULATION

Using the total point values tabulated from the severity and repeatability section, the scores are fit to the matrices shown in Figures 3 & 4, which is based on the SMS Risk matrix as presented in [4]. Depending on if weather was a factor in the RAE, dictates which SMS matrix scale is used. The repeatability values were equally scaled up from the original repeatability section to accommodate the addition of new factors. The severity section scores are fit so that the catastrophic column is only used in the scenario of an actual collision. A combined risk score consisting of a severity and repeatability is given for ATC and ATC-Pilot to represent the NAS risk. The color categorization maps to the level of risk of the event (Green-Low, Yellow-Medium, and High-Red).

Repeatability						
	>31	1/5	2/5	3/5	4/5	5/5
	24-31	1/4	2/4	3/4	4/4	5/4
	17-23	1/3	2/3	3/3	4/3	5/3
	12-16	1/2	2/2	3/2	4/2	5/2
	0-11	1/1	2/1	3/1	4/1	5/1
		0-11	>11- 27	>27 - 46	>46	Collision
		Severity				

Figure 3: Surface RAP to SMS Matrix Map without Weather as a Factor

Repeatability						
	>31	1/5	2/5	3/5	4/5	5/5
	24-31	1/4	2/4	3/4	4/4	5/4
	17-23	1/3	2/3	3/3	4/3	5/3
	12-16	1/2	2/2	3/2	4/2	5/2
	0-11	1/1	2/1	3/1	4/1	5/1
		0-13	>13- 29	>29 - 48	>48	Collision
		Severity				

Figure 4: Surface RAP to SMS Matrix Map with Weather as a Factor

When scoring each RAE, the panel must reach a consensus regarding each input selection. When a consensus cannot be reached or there is not enough information to score the section accurately, the selection is left unscored. The Reliability Factor (RF) of the tool scores the number of sections that are scored completely and not left unscored due to a lack of consensus or information. The RF score dictates whether an event is accepted to the

Each severity section RF is weighted equally. There are 9 severity sections, each with a severity RF weight of 11.11%, equally divided across 100%. For every section that is scored or not left blank, the severity RF increases by the RF weight. When a section is not scored or a section is left blank, the severity RF does not increase. In sections that have both an ATC and Pilot option, the RF is distributed equally between the two columns. Example: the execution section has an ATC and Pilot option. If the ATC Execution section is not scored, but the Pilot Execution section is complete, the RF weight for this section would be only 5.56%, half of 11.11%.

The same method applies for computing the repeatability RF. This time, however, each section carries a weight of 12.5%, evenly distributed across 8 repeatability sections. As in the severity section, in repeatability sections with both an ATC and Pilot options, the RF is distributed equally.

A RF severity score and a RF repeatability score are computed along with the overall risk scores. The total RF score is computed by averaging the RF severity and RF repeatability scores. Events with a total RF score below a certain threshold are not accepted as part of the risk analysis process. The Quality Assurance (QA) staff will determine the RF threshold.

V. VALIDATION

All 167 RAE's were scored using the proposed severity method and compared to the scores from the original scheme produced by EUROCONTROL. The intent was to keep a consistent risk evaluation as the original scheme, while adding additional data and using empirical data. In order to maintain consistency the process of validation did not include the weather section and the new proposed controllability factors. In addition, full details regarding this information were not always available to provide an accurate score. Likewise, the new proposed repeatability factors were not included for the validation, resulting in the repeatability section being unchanged. The SMS matrix scale for severity was remapped to accommodate the severity section having different maximum scores compared to the EUROCONTROL model, without the inclusion of the weather section.

Based on the results of the validation between the new proposed scoring structure and the original RAP Tool, the majority of RAE's did not differ in their final risk score. In the events where the difference was greater than one, the change was often due to the weighting format of the original scheme. For example, in the original scheme the separation scale assigned a score of three for distances between 400 feet and 6,000 feet and several of the severity scores increased. Under the proposal, the severity score changed for events by as much as seven for separation and closure rate. This in conjunction with portions of the controllability section that were weighted more heavily (e.g. Recovery) resulted in a change in risk score of greater than one.

Although 6 events (~6%) of events changed scores by more than two points, no events changed risk categorizations by more than two levels. In total, there were 26 NAS scores out of 167 total events (15%) that changed levels of risk categorization. Each of the events were reviewed individually and the differences were accepted by the RAP panel.

The proposed scoring method for the RAP Tool enhances the calculation of a risk value for surface loss of separation events based on empirical data. The changes provide a technique of scoring that is easily adjustable as historical data changes. Biases in the original RAP Tool version that were introduced by the weighting of the point scale are reduced in the proposed version by equally weighting each of the different sections. Future additions to the tool should address the need to weight individual sections based on collected data. This could be accomplished by incorporating HITL simulation results that estimate the importance level of each control. Incorporation of weighted controls would better fit surface events in the NAS and provide a more accurate risk evaluation.

VI. USE CASES AND FUTURE WORK

Upon scoring completion of each RAE event, the panel are able to select more in depth causal factors associated with the event. These factors and associated risk score are used by the FAA to prioritize risk factors in the NAS. Among these prioritizations is the FAA Top 5 Hazards, which are a quantifiable list of hazards that contribute to the highest risk events in the NAS. These hazards are a high priority performance goals, which are mitigated through a series of corrective action plans. The FAA Top 5 Hazards are often selected using data generated from the RAP Tool in order to identify factors associated with the highest risk.

Going forward the RAP Tool will continually evolve as the main risk quantification tool for AJI. Currently, the airborne and surface based tools are the only models currently in use, however there are additional plans to create a tool that evaluate terrain based loss of separations. As the current tools modernize, there are plans to continually update the data distributions that are used to drive the tool as more data becomes available. Furthermore, there are plans to investigate weighting the various sections of the tool as it contributes to the overall risk score. By giving different weights to the sections, it is anticipated a more accurate risk calculation will be obtained since different severity and repeatability components have more influence on affecting the overall risk level. It is anticipated these changes will make the tool more robust and accurate in depicting the risk of an event.

REFERENCES

[1] E. Chang. Risk Analysis Process (RAP) Tool Enhancements," MITRE Technical Report, MTR 120487, 2012

[2] P. Moertl. "Human-in-the-Loop Simulation of an Integrated Ground Movement Safety System," MITRE Technical Report, MTR 05W0000074, 2007

[3] ICAO "Runway Incursion Severity Classification (RISC)," International Civil Aviation Organization., unpublished.

[4] FAA Air Traffic Organization. "Safety Management System Manual v3.0," Federal Aviation Administration., unpublished.

[5] A. Sheffield. "Standard Operating Procedures for the Quality Assurance Risk Analysis Process," Federal Aviation Administration., unpublished.

ACKNOWLEDGMENTS

The author wishes to thank Dr. Pradip Som, Art Sheffield, and Bryan Rife of the Federal Aviation Administration and Mike McFadyen of SENTEL

Corporation for their valuable input and oversight for this work. Appreciation goes to EUROCONTROL as well for their collaboration with the tool. The author also thanks Shweta Mulcare and Chris Devlin of the MITRE Corporation for all their help throughout.

AUTHOR BIOGRAPHY

Eric B. Chang received a M.S. in Aerospace Engineering from Georgia Institute of Technology, Atlanta, GA, U.S in 2010 and a B.S. in Mechanical Engineering from Carnegie Mellon University, Pittsburgh, PA, U.S in 2008. He is currently a Senior Aerospace Engineer for the MITRE Corporation, McLean, VA, U.S where he works on Air Traffic Control related safety analyses.

NOTICE