Risk-Based Airport Selection for Runway Safety Assessments
Through the Development and Application of
Systems-Driven Prioritization Methodologies

Kyle M. Gallé, John C. Ale, Jr., Mohammad M. Hossain, Michael J. Moliterno, Maureen K. Rowell, Nikita V. Revenko, Ellen C. Rogerson, Scott F. Tucker, Kenneth G. Crowther, Member, IEEE, James H. Lambert, Member, IEEE, and Yacov Y. Haimes, Fellow, IEEE

Abstract — A runway incursion is the erroneous presence of an aircraft or other object on the runway. Runway incursions are rare precursors to aviation accidents and result from a variety of complex factors. Current quantitative methods are inadequate for analysis, forecasting, and understanding the risk profiles of airports. The Federal Aviation Administration (FAA) biennially employs runway safety action teams (RSATs) to evaluate airports with the aim of reducing runway incursions. This manuscript describes the development of a decision-making tool that combines systems methodologies to help the FAA select airports at which employing RSATs would be the most beneficial for the reduction of collision risk. These methodologies include hierarchical Bayesian modeling (HBM) and analytic hierarchy process (AHP).

HBM leverages the similarities between airports through conditioning incursion rate estimates on both historical data and the data sampled from hyper-distributions fitted using data from similar airports. AHP decomposes these RSAT placement decisions into a hierarchy of decisions to be examined independently; as the system evolves, the method allows the FAA to reflect changes in its knowledge and preferences in the weights of corresponding risk factors.

The key deliverable of this project, a workbook decision tool, places HBM within the AHP hierarchy. The final methodology output presents an ordered list of airports according to potential risk-reduction from RSATs. This combination further incorporates risk-based and decision-based approaches in order to better allocate resources and reduce runway incursions. The tool meets project goals by utilizing the FAA’s data to provide an analytically justifiable prioritized list, while building on current best practices for the RSAT selection process.

I. INTRODUCTION

Aviation has proven itself one of the most significant transportation innovations of the past century. Any airplane collision hurts society, as it usually results in loss of life, property damage, and a decrease in confidence in air travel [1].

The Federal Aviation Administration (FAA) monitors and evaluates the safety of all towered airports in the United States. High-risk situations for airplanes occur in the crowded air space around airports and on runways. The FAA defines a runway incursion as an incident in which an aircraft, vehicle, person or object on the ground poses a collision hazard with an aircraft that is taking off or landing at an airport under the supervision of the Air Traffic Control (ATC) tower [2]. In attempts to reduce the risk of runway incursions, the FAA established runway safety action teams (RSATs), multidisciplinary teams of experts (e.g., air traffic controllers, pilots, and other aviation experts, officials, and employees) that discuss and evaluate the safety of an airport’s runway operations [3]. The FAA’s Runway Safety office claims runway incursions are “precursors to aviation accidents” and have increased in rate since fiscal year 2007 [4]. After approximately two weeks of meetings, the RSAT produces a list of action items geared at improving the safety of the airport it evaluates. The FAA desires to send RSATs to each towered airport in the US every two years, but limited resources with which to execute RSATs prevent the achievement of this goal.

This project focuses specifically on the FAA’s use of RSATs to reduce the risk of accidents and incursions. To encourage safety through risk mitigation, this project focuses on a risk-based methodology developed for selecting airports for RSATs.

This paper summarizes work by the University of Virginia Systems and Information Engineering Capstone Team (Capstone), along with the University of Virginia Center for Risk Management of Engineering Systems (CRMES). The manuscript aids the FAA in fulfilling their goal of reducing runway incursions and runway incursion rates. The project’s contribution, a Microsoft Excel-based tool, will prioritize airports according to various contributing factors affecting the risk of a runway incursion. The FAA can use these results to better justify RSAT allocation and reduce the rate of runway incursions. Achieving the incursion reduction goal will benefit travelers, commercial airlines, general aviation, towered airports, the US Department of Transportation (USDOT) and the field of risk analysis.

II. BACKGROUND

Identification of risk factors, changes that negatively affect the state of the airport system, can improve mitigation
techniques, and reduce the likelihood of risky events. Risk analysis includes the creation of metrics that act as both a collective measurement and a balance between the competing objectives of any complex system. Risk assessment determines which events can cause failure and the likelihood and consequences of these events. Risk management explores mitigation strategies, possible tradeoffs, and how current decisions affect future options. This project aims to persuade the FAA to leverage all available information for the purposes of aviation safety improvement. Developing a tool that the FAA can use to apply its resources more effectively is the first step in achieving this goal.

The FAA’s prioritization tool builds on a benchmark prioritization tool employed by the Great Lakes region. The original workbook includes major factors that contribute to the occurrence of a runway incursion, including basic airport information, runway geometry, historical runway incursions and historical operations [3]. The team’s workbook modifies the original to meet the FAA’s requirements of producing an expert decision tool for the RSAT selection process. The resulting product includes a workbook that is presentable, understandable, and comprehensive. The analysis employed within this product utilizes scholarly research of runway incursions and risk-based methodologies in a multi-pronged approach to meet the challenges posed by addressing runway incursions. The final tool will directly contribute to the increased efficiency of RSAT selections.

III. METHODOLOGY AND RESULTS

This section details the methodologies integrated into the FAA’s prioritization system. Section A describes the hierarchical Bayesian modeling (HBM) process for leveraging the similarity between airports to derive confidence intervals for runway incursion rates. Section B describes the development and application of analytical hierarchy process (AHP) for decomposing the complex runway safety problem and integrating expert judgment regarding the effectiveness of RSATs at reducing the complex risk factors. Section C describes how these processes integrate with cost-benefit analytics to create an expert decision tool.

_A. Hierarchical Bayesian Modeling_

The data gathered by the FAA contains measurements for all eighty airports in the FAA Great Lakes region across eleven unique quantitative variables with varying scales. The lack of consistent structure amongst the variables, as well as the rare nature of incursion events and the diversity of airport operations, relegates standard data analysis. As an example, some airports have multiple crossing runways while others do not. The samples are not truly homogenous with several factors causing separations within the data. Certain airports, like Chicago's O'Hare International Airport, experience more runway incursions than smaller airport like Flying Cloud in Minnesota, most likely influenced by the increased magnitude of flights occurring at these airports [4].

The analysis requires addressing the data's inherent complications. One advantage of clustering groups of data is the ability to create more homogenous data sets, which increase the reliability of conclusions from data analysis techniques. Clustering uses measurements from each datum’s distance between variables and groups the data that are most alike [5]. Figure 1 presents an example by showing the relative distance between airports in the Great Lakes Region based on runway geometry.

Clustering algorithms can interpret and analyze data that contains several different scales and/or qualitative data. While clustering can create more homogeneity amongst the data points, it also reduces the amount of information within these groups; the groups contain a smaller number of data points. Addressing this issue requires capturing both the need for homogeneity and the need for maintaining information. Despite this tradeoff, clustering approaches contain limitations. Clustering cannot reduce the role that airports’ unique and individual variables play in runway incursion rates. Each airport’s inherent uniqueness, measured by the multiple factors, constrains the effectiveness of airport clustering. In spite of these limitations, clustering is appropriate for data visualization and as an analysis tool because it illuminates similarity and thereby increases the effectiveness of quantitative analysis.

Clustering on several uniquely scaled factors requires a flexible clustering methodology. For example, the Gower dissimilarity matrix is able to derive repeatable measures describing the similarity between groups [6]. The resulting

![Hierarchically Clustered Dendogram](image-url) The separation of airports into five clusters increases homogeneity within groups.
hierarchical clustered tree provides a visual representation of the likeness of each airport's geometric characteristics. The results, shown in Figure 1, indicate five distinct groups that arise from the clustering algorithm used.

Two approaches, classical statistics and Bayesian analysis, provide a thorough assessment of data in a logical, justifiable manner. The classical statistics approach to quantitative analysis can synthesize data into appropriate correlations between variables. Furthermore, the approach has built-in methods to validate assumptions. For standard linear regression, the resulting residuals must demonstrate homoscedasticity, independence, and a normal distribution. Classical statistics also has limitations—primarily that failure to meet these assumptions renders the investigation weak and/or meaningless. The FAA’s data demonstrate non-normal residuals and thus do not meet the required assumptions. Variable transformations (i.e., Box-Cox), which provide further explanation of the data and remove the constraint of homoscedasticity, do not yield residuals that meet the required assumptions.

Bayesian statistics imposes fewer constraints surrounding a belief that the data points are similar [7]. Clustering or logical demonstration justifies this assumption. Building on the clustering analysis, Bayesian analysis can demonstrate a posterior probability using the information from prior distributions. Hierarchical Bayesian modeling (HBM) uses this approach to manage complex systems where the underlying distribution is unknown [8]. Instead, the data selects and conditions an uninformative prior. Since the factors contributing to each runway incursion are complex and often unknown, this method leverages all available information. The scope of Bayesian modeling provides a more reasonable estimation of the runway incursion rate at airports [9]. The resulting confidence intervals receive conditional information from the groups of airports, not just the previous year's data. This grouping maximizes usage of information to provide confidence intervals. These intervals emphasize the inherent probabilistic nature of runway incursions over point estimates based on historic data. The Bayesian modeling approach provides the FAA with the means of updating and analyzing particular airports given the scarcity of airport data. Figure 2 shows the confidence intervals of runway incursion rates for select airports in the Great Lakes region. It represents the range in confidence intervals for a particular group of airports.

Creating a valid model requires identifying the probability distribution of the runway incursion rate for each cluster. A histogram of runway incursion rates from each cluster shows an exponential distribution. Combining this lower level distribution with the generalized hyper-distribution, an uninformative prior, provides the overall HBM.

Gibbs sampling provides a means to extracting the confidence intervals [10], [11]. The WinBUGS software developed by the BUGS project provides a convenient implementation of Gibbs sampling for Bayesian models. WinBUGS provides simulated HBM model results in appropriate confidence intervals. Results from this process demonstrate a similar median runway incursion rate across geometry groups, indicating that, independent of other factors, runway geometry is not a significant predictor of runway incursions. The results for one clustered group, shown in Figure 2, also contain outliers. Outlier airports are those that have incursion rate intervals that do not overlap with their respective group. These results further highlight extenuating causal factors unrelated to airport geometry.

The results cannot support conclusions implying a significant causal relationship between airport geometry and runway incursions. When grouped by geometry, some airports do not have overlapping incursion rate confidence intervals, implying that geometric factors do not explain the effective variance in runway incursion rates. Instead, several explanations may exist. Primarily, there may be a need to record more measurable factors regarding geometry. Extrapolation from more extensive data sets over longer time periods may provide a more conclusive answer. Sole reliance on data analysis inadequately captures the complexity of runway incursions. As such, the team must employ process techniques to extract expert judgment.

B. Analytical Hierarchical Process

While the process of selecting which airports should receive RSATs requires subjective decision-making, the decisions should maintain consistency. This consistency enables improved justification and continuous improvement in the program. To increase consistency, the decision-making should be as simple as possible. The analytic hierarchy process (AHP) addresses both consistency and complexity [12]. AHP’s goal is to produce a dynamic weighting scheme for system elements that exhibits the FAA’s knowledge, intuition and preferences, which ultimately influence the ordered list of airports. The model integrates judgments from experts regarding the ability of RSATs to influence operations intended to overcome the various risk factors. This represents a reinterpretation of the standard application of AHP, which incorporates the judgment of experts and not just their preferences.

When applied to the airport data, the weighting scheme generates a prioritized list of airports. AHP decomposes one complex decision or judgment about a system into many smaller decisions or judgments. This approach increases expert objectivity by allowing experts to analyze two elements at a time. The selection process arranges this approach into a hierarchy of system elements. The hierarchy consists of three levels, each addressed independently. Within each level, users make pair-wise comparisons of the airport risk factors with regard to their relative likelihood to be influenced and/or mitigated by an RSAT (or more specifically by the operations/recommendations resulting from the RSAT). Comparing the elements of one level at a time reduces the complexity of the decision-making. AHP requires that the FAA record the comparisons numerically.
This approach transforms personal knowledge into a mathematical framework, which then derives information from subjective expert input. AHP synthesizes experts’ numerical comparisons by first computing the geometric mean for each pair-wise comparison and then by entering the means into an overall preference matrix. Each level of the hierarchy generates separate preference matrices for its internal subgroups; the normalization of each matrix allows for the calculation of corresponding eigenvalues. The eigenvector corresponding to the largest eigenvalue, called the principal eigenvalue, is the priority vector. This mathematically derived vector, which assigns each system element a relative weight, is the main AHP output [12]. The generated weights apply to each level of the hierarchy and result in a prioritized list of airports. AHP permits the weighting scheme to react to modified preferences by integrating these changes in the expert judgment input step. The priority vector reflects changes in the numeric comparison of the elements and exhibits the modifications instantaneously. AHP checks the consistency of comparisons at each level of the hierarchy and requires each expert to maintain at least 90% consistency among his decisions [12]. The result of AHP applied to the airport selection process is a prioritized list of airports resulting from consistent and more transparent decisions.

The FAA system is large and the resulting pair-wise comparisons are too numerous to be effectively addressed by the FAA and is thus simplified by assuming a degree of consistency and removing certain comparisons of system elements. The system requires a modified AHP approach that consists of making assumptions about certain system elements and removing those comparisons from the expert judgment input step of AHP—essentially requiring a certain degree of consistency. Communication with FAA experts and a literature review support these assumptions. This approach uses the incomplete pair-wise comparisons (IPC) algorithm [13] to reduce the number of pair-wise comparisons needed from 80 to 23. In a Monte Carlo simulation study, the team concluded that without assumptions, 50% of the pair-wise comparisons could be randomly deleted with no significant reduction in consistency. Literature supports the conclusion that using these assumptions allows for the reduction of more comparisons while attaining higher consistency [13].

The AHP application to the FAA system begins with the construction of a system hierarchy. The hierarchy decomposes the selection of airports to receive RSATs into three levels: objectives, criteria, and airport data. The objectives consist of airport information, airport geometry, airport operations, events/severity, and RSAT history. Airport information describes the status of each airport with regard to certification and university affiliation. Airport geometry summarizes airfield layout, and airport operations explains the nature of activity at each airport. Events/severity describes the number and type of runway incursions that have occurred, and RSAT history reports the days since the last RSAT at each airport. Each objective has a second level consisting of criteria. For example, airport geometry has the criteria multiple crossing runways, short taxi routes/time compression, possible confusion while taxiing, number of intersections, and parallel or closely aligned runways. The FAA collects data for each criterion; these data constitute the third level and require no subjective decision-making. The following AHP steps handle each level of the hierarchy as independent problems.

The next step in AHP distributes a survey that allows experts to apply their judgment concerning which system elements an RSAT would most likely mitigate. A comprehensive survey, in which the expert completes every pair-wise comparison, causes information overload and potentially elicits invalid data. Two assumptions, transitivity and equivalence, provide an appropriate means to decrease expert fatigue. Transitivity implies that if \( x \) describes the relationship between \( A \) and \( B \), and \( y \) describes the relationship between \( B \) and \( C \), then \( x/y \) describes the relationship between \( A \) and \( C \). Equivalence implies that an expert ranks the importance of similar risk factors the same. Risk factors that address identical risk scenarios exist within the survey. As an example, an expert evaluating the risks associated with parallel and closely aligned runways might visualize the same risk scenario of a pilot taking off or landing on the wrong runway. The combined effect of these assumptions results in a survey that requires only 30% of the

![Figure 2. Confidence Intervals of Runway Incursion Rate](image-url) Some airports fail to have intervals within the median value of the group.
original comparisons. The condensed survey completes a condensed survey of pair-wise comparisons and records them as a 1, 5, or 9 in the following manner. Taking criteria \( A \) and \( B \), for example shown in Figure 3, an expert can judge which criteria is more significantly adjusted for by actions following an RSAT, followed by its corresponding score. If the expert chooses criteria \( A \) as more significant, a score of 1 identifies \( A \) as equally as significant as \( B \), 5 identifies \( A \) as more significant than \( B \), and 9 identifies \( A \) as much more significant than \( B \).

The preference matrix synthesizes these comparisons and derives judgments for the remaining criteria. The matrix consists of 1s, 5s, and 9s in the appropriate positions and the reciprocals \((1, 1/5, 1/9)\) in the inverted comparison positions. The completed preference matrices, once normalized, generate the priority vectors. These priority vectors translate into weights for each criterion. AHP yields weighting schemes for all the objectives and their corresponding criteria and forms a weighted sum. The application of the weighting scheme to each airport results in a prioritized list of airports. The reported list incorporates the order and intensity of airport rankings.

The integrated expert system does not use these approaches exclusively but instead synthesizes the methodologies into a process that will help to inform the prioritization process. It also allows for the acquisition and integration of additional data into its decision process. As the tool evolves, it will continue to adapt more quantitative methods, and it will integrate the information as it relates to runway safety and mature decision processes. This evolution, combined with other elements of the project, provides a framework to the FAA that is holistic, adaptive, and will support incremental changes as well as the current basic needs of risk-based, systems driven prioritization.

Figure 3. Expert Judgment Survey Using AHP This survey elicits expert knowledge with regard to critical runway factors.

Figure 4. Cost-Benefit Analysis This graph contains the airports bound by costs defined by the FAA.

C. Integrated Expert System

Each set of runway incursion data gathered by the FAA provides an important means of understanding the believed risk factors affecting the airports. Employing several methodologies to evaluate the data meets the necessary multiple perspectives of risk analysis. The addition of cost-benefit analysis to HBM, which leverages data across similar groups, and AHP, which aggregates various risk factors, creates a holistic assessment tool. This integrated expert system provides the FAA with the ability to better maximize RSAT effectiveness as it pertains to reducing the number and/or rate of runway incursions.

Figure 4 shows an example of a cost-benefit analysis that identifies which airports should receive RSATs according to adjustable cost-effective bounds. These bounds include the monetary cost of sending an RSAT team, the cost of a runway incursion incident, and the desired reduction of the runway incursion rate. Adjustment of these values shifts the bounds on the graph. The cost effectiveness of RSATs increases by selecting RSATs at airports within the cost constraints that maximize the overall impact of the RSAT. Overall, cost-benefit analysis helps integrate the various factors, but certain cost bounds require adjustments specified by the FAA.

The current tool framework integrates continuous improvement within its designs. Improvements to the current system could include migration to a web-based version of the workbook. The web-based version would support back-end data processing and maintenance by the central FAA authorities and decision analytic support to the regions and airports. This could increase the effectiveness of the RSAT selection process by adding the ability of real-time processing and collaborative working. While Excel provides simplicity, a web-based interface can extend the ease of use after the finalization of the workbook functionality.

Another possible improvement is the extension of the workbook tool to other FAA regions. This feature requires modification to the workbook tool to include specific scalability properties. The use of HBM does limit scalability by requiring model and data updates between years. AHP, however, incorporates the necessary flexibility. Overall,
further research and expansion of the tool become possible through analyses of more extensive data, objectively assessing the recommendations, and analyses on the impacts of RSATs. Combined with standardization of data collection techniques across regions could allow expansion of the tool for use by the multiple FAA regions. All of these possible additions could be part of a continual improvement process to provide the FAA with a streamlined system expandable beyond the Great Lakes region. Future versions of the analysis tool could include risk management techniques developed by the CRMES. These additional tools can allow federal management to better understand and utilize the resources of the RSATs.

V. CONCLUSION

The FAA and the CRMES are working together to address the multiple contributing factors associated with runway incursions. The team’s contribution directly benefits the FAA by enabling the agency to better prioritize which airports need an RSAT assessment. The RSAT selection tool synthesizes historical data and expert opinions, which allows the FAA to leverage the work of RSATs to bring significant change to aviation safety. HBM analysis yields no evidence to support or reject the significance of geometry as a factor related to runway incursions, while AHP evaluates the potential impact of an RSAT on all risk factors. Both methodologies combine to create a prioritization of RSATs that will subsequently maximize RSAT effectiveness by placing the teams at airports that need a runway safety assessment. This will promote airport safety across the US, which, in turn, will result in preventing future runway incursions nationwide.

ACKNOWLEDGMENTS AND DISCLAIMER

We thank the following individuals: Larry Cunningham, contractor for FAA Great Lakes Region; John Lott, Director of FAA Great Lakes Region; Greg Won, FAA Runway Safety Office. This paper is partially supported by the FAA, however, the authors take full responsibility for the ideas, concepts, and opinions presented here, which do not necessarily reflect those held by the FAA.

REFERENCES


