4-2021

# Development of a Safety Performance Decision-Making Tool for Flight Training Organization

Marisa D. Aguiar

# DEVELOPMENT OF A SAFETY PERFORMANCE DECISION-MAKING TOOL FOR FLIGHT TRAINING ORGANIZATIONS

By

Marisa D. Aguiar

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
April 2021

**DEVELOPMENT OF A SAFETY PERFORMANCE DECISION-MAKING TOOL FOR FLIGHT TRAINING ORGANIZATIONS**

Marisa D. Aguiar

This dissertation was prepared under the direction of the candidate's Dissertation Committee Chair, Dr. Carolina L. Anderson, and has been approved by the members of the dissertation committee. It was submitted to the College of Aviation and was accepted in partial fulfillment of the requirements for the Degree of
Doctor of Philosophy in Aviation.

Carolina L. Anderson, Ph.D.
Committee Chair

Dothang Truong, Ph.D.
Committee Member

Steven Hampton, Ed.D.
Associate Dean, School of Graduate Studies, College of Aviation

Kenneth P. Byrnes, Ph.D.
Committee Member

Alan J. Stolzer, Ph.D.
Dean, College of Aviation

Gregory S. Woo, Ph.D.
Committee Member (External)

Lon Moeller, J.D.
Senior Vice President for Academic Affairs & Provost

April 7, 2021

## ABSTRACT

Researcher: Marisa D. Aguiar

Title: DEVELOPMENT OF A SAFETY PERFORMANCE DECISION-MAKING TOOL FOR FLIGHT TRAINING ORGANIZATIONS

Institution: Embry-Riddle Aeronautical University

Degree: Doctor of Philosophy in Aviation

Year: 2021

Title 14 of the Code of Federal Regulations (CFR) Part 141 flight training organizations are actively pursuing ways to increase operational safety by introducing advanced risk assessment and decision-making techniques. The purpose of the dissertation was to create and validate a safety performance decision-making tool to transform a reactive safety model into a predictive, safety performance decision-making tool, specific to large, collegiate Title 14 CFR Part 141 flight training organizations, to increase safety and aid in operational decision-making. The validated safety decision-making tool uses what-if scenarios to assess how changes to the controllable input variables impact the overall level of operational risk within an organization's flight department.

Utilizing SPIs determined to be most indicative of flight risk within large, collegiate flight training organizations, a predictive, safety performance decision-making tool was developed utilizing Monte Carlo simulation. In a high-risk system beset with uncertainty, applying Monte Carlo simulation addresses the need to accommodate uncontrollable inputs into the model in a manner that enables the model to produce meaningful output data. This research utilizes the validated equations drawn from the non-statistical model developed by Anderson, Aguiar, Truong, Friend, Williams, &

Dickson (2020) for the mathematical inputs driving the computational nodes, including the SPIs, as the foundation to develop the safety performance decision-making tool.

The probability distributions of the uncontrollable inputs were drawn from a sample of operational data from September 2017 to September 2019 from a large, collegiate 14 CFR Part 141 flight training organization in the southeastern United States. The study conducted simulation runs based on true operational ranges to simulate the operating conditions possible within large, collegiate CFR Part 141 flight training organizations with varying levels of controllable resources including personnel (Aviation Maintenance Technicians and Instructor Pilots) and expenditures (active flight students and available aircraft).

The study compared the output from three different Verification Scenarios—each using a unique seed value to ensure a different sample of random numbers for the uncontrollable inputs. ANOVA testing indicated no significant differences appeared among the three different groups, indicating the results are statistically reliable.

Four What-if Scenarios were conducted by manipulating the controllable inputs. Mean probability was the key output and represents the forecasted level of operational risk on a standardized 0-5 risk scale for the Flight Score, Maintenance Score, Damage and Related Impact, and an Overall Risk Score. Results indicate the lowest Overall Risk Score occurred when the level of personnel was high yet expenditures were moderate.

Changes to the controllable inputs are reflected by variations to the outputs demonstrating the utility and potential for the safety performance decision-making tool. The outputs could be utilized by safety personnel and administrators to make more informed safety-related decisions without expending unnecessary resources. The model could be adapted for use in any CFR Part 141 flight training organization with data collection capabilities and an SMS by modifying the input value probability distributions to reflect the operating conditions of the selected 14 CFR Part 141 flight training organization.

## DEDICATION

I dedicate this dissertation to my father, Captain Orlando Aguiar (1965-1994), and the crew and passengers of American Eagle Flight 4184. May we never forget the lessons learned and continue passionately striving to keep our pilots, crew, and passengers safe.

To my little world-changers, Jamison Orlando and Dominic Samuel, may you always have the courage to pursue your wildest dreams; you are capable of more than you could ever imagine.

.

## ACKNOWLEDGEMENTS

I sincerely thank my dissertation committee for their time, thoughtful feedback, and encouraging support throughout the dissertation process. I especially thank my dissertation chair, Dr. Carolina Anderson, for her mentorship, friendship, gentle guidance, and allowing me the creative freedom to blossom as a researcher. Thank you to my external committee member, Dr. Gregory Woo, for his exemplary patience and for taking the time to walk me through challenging methodological concepts.

Thank you to my committee member and Program Coordinator, Dr. Dothang Truong, for his invaluable support, guidance, and mentorship throughout the dissertation process and during my time as a residential student within the program. His insightful feedback and constructive guidance provided in both the coursework and throughout the dissertation process have kept me encouraged and focused.

A simple acknowledgment cannot begin to describe the immense gratitude I have for the mentorship, guidance, and friendship I have received from Dr. Haydee Cuevas. I would not have made it through this program without her unwavering positivity, genuine compassion, and superior guidance. I am immensely grateful for everything she has taught me both academically and about life. The world needs more educators like her.

Thank you to Dr. Alan Stolzer for believing in me by allowing me to be the first Doctoral Graduate Research Assistant. This life-changing opportunity has grown me into a confident, competent researcher and passionate aviation safety advocate. I am eternally grateful.

Thank you to my colleagues in the program and fellow Doctoral Graduate Research Assistants for your support and humor when I needed it most. I will always smile fondly upon memories of working in the office together. Thank you to all the Ph.D. in Aviation faculty who have supported me and provided guidance both academically and professionally. I am thankful for every opportunity I have had to learn and grow.

Thank you to my family and friends who have supported me through all the ups and downs of this journey. I especially want to thank my mom for being the first one to believe in me. I would have never made it to where I am today without the unconditional support from my mother. Thank you to my incredible life partner, Jamison, for everything he has sacrificed for me to pursue this dream. Jamison's daily support, encouragement, and love saw me through so many challenges.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# CHAPTER I

## INTRODUCTION

A CFR Part 141 organization can be defined as a pilot training school certified under the specifications defined by the Federal Aviation Administration (FAA) under Title 14 of the Code of Federal Regulations Part 141 (Federal Aviation Administration, 2017). As defined in Advisory Circular 141-1B, academic institutions may offer aviation-related degrees and pilot training under CFR Part 141; CFR Part 141 flight schools have the option to utilize a wider variety of training tools; although, dedicated flight training facilities, qualified flight instructors, and FAA-approved course curricula are still required (FAA, 2017).

Per Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020), a *large* CFR Part 141 could be defined as a pilot training school operating under Title 14 CFR Part 141 with the following criteria:

- At least 500 student pilots
- A fleet of at least 50 aircraft with the integrated flight instrument system capabilities
- A Flight Data Monitoring system with data collection
- A scheduling system
- An active and robust Safety Management System (SMS)

The complexity of many aviation accidents and incidents combined with rapid technological progress has left traditional bottom-up and top-down system safety assessment techniques outdated and inadequate (Dakwat & Villani, 2018; Dekker, 2011; Stringfellow, Leveson, & Owens, 2010). A major limitation of traditional safety assessment techniques is the challenge of considering all potential risks that may arise from multiple variables interacting together (Dakwat & Villani, 2018). Mitigative actions based on the analyses of previous accidents and incidents are both reactive and insufficient to further the progress of proactive safety management (ICAO, 2013). Additionally, the absence of accidents and incidents within CFR Part 141 flight training organizations does not assume operations are functioning at the optimum level of safety (Adjekum, 2014; Cassens, 2010; Keller, 2015; Mendonca & Carney, 2017). A modern approach to safety management includes proactively addressing safety risks rather than relying on inspections and remedial actions.

The complex, high-risk nature of CFR Part 141 flight training organizations grants particular susceptibility to risk, potentially leading to a series of systematic failures. This drift into failure occurs through the slow normalization risk, occurring as an incremental deterioration of safe operating conditions propelled by organizational failures, misunderstood technology, and social influences (Dekker, 2011).

To avoid the process of drifting into failure, organizations are developing ways to increase their level of safety by incorporating advanced risk assessment and decision-making techniques to strengthen the risk management element of the organization's SMS (Ale, Bellamy, Cooke, Goossens, Hale, Roelen, & Smith, 2006). Simulation modeling techniques are becoming more widely utilized in complex, high-risk systems across various domains to optimize the safety assessment process (Blair, 2017; Chen & Jing, 2016; Gunduz, Birgonul, & Ozdemir, 2017; Hadjimichael, 2009; Stonesifer, Calkin, Thompson, & Kaiden, 2014). SPIs are useful for observing and monitoring known risks, as well as detecting future risks to elicit corrective action before an adverse event occurring. SPIs play a valuable role within an organization's SMS by enabling performance-based safety management while supporting the organization's unique safety objectives (Pierobon, 2016).

However, existing SPIs, although useful in measuring the effectiveness of the organization's SMS, are incapable of providing a true predictive approach to safety decision-making, as the data collected to feed into the SPIs are based on events, instances, and operations that have already occurred (Patriarca, Di Gravio, Cioponea, & Licu, 2019). Thus, any responses or corrective action made based on these data findings is a retroactive approach to safety. The use of what-if scenarios via simulation allows for an in-depth look at interactions within the system and assesses the impact of a change to the system before any changes take place, rather than retrospectively assessing the effects of a change. Further research is needed to transform SPIs into predictive safety decision-making tools capable of taking proactive safety one step further by modeling the potential of the system without compromising resources.

**Statement of the Problem**

Traditionally, the aviation industry has focused on the utilization of historical events, such as accident data, or those indicators of safety that are clearly measurable (Oswald, Zhang, Lingard, Payam, & Tiendung, 2018). However, safety monitoring based on relevant, operational SPIs is still a reactive approach to safety monitoring backed by linear reasoning; whereas the aviation risk assessment process must continually evolve and improve by considering new approaches to safety monitoring and decision-making that provides greater insight into *why* accidents occur and *how* safety is best achieved. Domain-specific SPIs provide a one-size-fits-all approach to safety monitoring, whereas forecasting models provide safety personnel with the ability to foresee how changes to various operating conditions impact the overall safety of the system pertinent to their particular operation (Hadjimichael, 2009).

Further research is needed within the industry to transform reactive safety models based on SPIs into safety decision-making tools, capable of handling the predictive uncertainty inherent to CFR Part 141 flight training organizations while incorporating the use of what-if scenarios, to evaluate how modifying the controllable input variables impact the safety and efficiency of the complex system as a whole. A safety decision-making tool, particular to large, collegiate CFR Part 141 flight training organizations yet adaptable to accommodate any flight training organization with data acquisition capabilities and an active SMS, would not only allow for a more proactive approach to safety but could also assist those in administrative roles with critical decision-making.

**Purpose Statement**

The purpose of the research was to create and validate a safety performance decision-making tool to transform a non-statistical model composed of 12 SPIs determined by Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020) to be most indicative of flight risk specific to 14 CFR Part 141 flight training organizations into a predictive, safety performance decision-making tool. The model uses what-if scenarios to evaluate how changing controllable input variables affect the level of operational risk within the system, portrayed within the model as the risk score outputs.

The validated model will utilize what-if scenarios to assess how changes to the controllable input variables influence the overall level of risk within the organization's flight department and various other departments.

The current research utilized the SPIs drawn from the non-statistical SPI model developed by Anderson et al. (2020), as these SPIs have been found to be most indicative of operational flight risk for a 14 CFR Part 141 flight training organization. Anderson et al. (2020) created and validated a non-statistical model encompassing SPIs from both flight and maintenance operations and their related formulae drawn from a two-year sample of operational flight and maintenance data. For the purpose of this dissertation, the SPIs from the non-statistical model developed by Anderson et al. (2020) were used as the foundation to develop a safety performance decision-making tool based on the input variables for the chosen SPIs. Monte Carlo simulation was conducted and run to enable the SPI model to handle uncertainty in some of the key, influential variables.

**Significance of the Study**

The extant literature indicated a deficit of predictive, safety performance decision-making tools specific to large, collegiate CFR Part 141 flight training organizations; therefore, this research fills an operational need within the industry. The study also extends the research conducted by Anderson et al. (2020) by expanding the non-statistical model into a safety performance decision-making tool utilizing Monte Carlo simulation to improve the accuracy and robustness of the flight training organization's SMS.

The research also improves the current understanding of the factors most substantially contributing to flight risk within large, collegiate CFR Part 141 flight training organizations. As a safety decision-making tool, the model could also be used by the administration within a large, collegiate CFR Part 141 flight training organization to rationalize new hires, technology acquisitions, and other safety-related initiatives by modeling the potential of modifying resources, or controllable inputs, without the risk associated with actually expending the organization’s resources. The model could also be modified for applicable use by any flight training organization with data acquisition capabilities and an operational SMS.

Theoretically significant, the model provides a mechanism for expanding the breadth of knowledge related to optimizing resources from both flight and maintenance operations to enhance operational safety for CFR Part 141 flight training organizations. Further, a thorough review of the extant literature indicated a gap in the process of going from traditionally reactive SPIs into safety performance decision-making tools with forecasting abilities for safety decision-making purposes specific to CFR Part 141 flight training operations. The research fills this gap by providing a validated safety decision-making tool, specific to CFR Part 141 operations, to further move the needle in the direction of proactive, rather than reactive, aviation safety assessment techniques.

The model created within this dissertation has a high level of generalizability, as the model could be adapted for use in any large CFR Part 141 flight training organization with data collection capabilities and an active SMS. This dissertation describes the process of transforming a reactive safety model composed of SPIs into a safety performance decision-making tool; thus, a 14 CFR flight training organization could utilize its own unique SPIs by determining the probability distributions of the uncontrollable input variables, further enhancing the generalizability of the safety performance decision-making tool. Providing large, CFR Part 141flight training organizations with a safety decision-making tool will enhance the risk management component of the organization’s SMS by taking an increasingly proactive approach to safety by providing insight into the impact changes to operating conditions may have on the safety of the overall operation. The ability to forecast operating conditions using Monte Carlo simulation will allow CFR Part 141 flight training organizations to make better informed safety-related decisions while optimizing efficiency without compromising safety.

**Research Question**

1. How can the SPI model developed by Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020) be transformed into a predictive, safety performance decision-making tool with the ability to run what-if scenarios?
2. How do changes to the controllable input variables impact the overall risk score?

**Delimitations**

The model is designed to measure the potential for increased or decreased flight risk for large, collegiate flight training programs within the United States. The displayed level of risk associated with monthly operating conditions can be used to make safety-related decisions by organizational safety personnel. The model does not measure occupational risks, such as injuries incurred in the maintenance hangar or personal slips, trips, and falls. The model does not measure cases of gross negligence, such as the willful disregard of standard operating procedures unless such occurrences are deemed to be systemic in nature. Security threats, including suicide and sabotage, are also not considered. Human performance state measurements were excluded from the analysis. Although there are some factors not covered in the study, these delimitations do not affect the rigor of the model as the SPIs utilized were chosen by Subject Matter Experts (SMEs) to be most appropriate in gauging flight risk for large, collegiate flight training operations within the United States (Anderson et al., 2020). The model is also highly adaptable and could be modified to include the delimitations not considered within the research, assuming the organization has the necessary data available.

**Limitations and Assumptions**

The research conducted for the purpose of this dissertation was limited to the creation and validation of a safety performance decision-making tool utilizing Monte Carlo simulation to transform a non-statistical model composed of the ten SPIs determined by Anderson et al. (2020) into a predictive, safety performance decision-making tool capable of running what-if scenarios to evaluate how changing controllable input variables within the system affect the overall level of operational risk, portrayed within the model as the overall risk score output. The variables used in this model are limited to those determined to be most useful in measuring flight risk in a large, collegiate CFR Part 141 flight training organization by SMEs in flight and maintenance operations (Anderson et al., 2020). Additionally, the model could easily be adapted to accommodate other flight training organizations with data collection capabilities and an operational SMS.

Per Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020), the model assumes a *large* CFR Part 141, defined as a pilot training school operating under Title 14 CFR Part 141, possesses the following operational criteria:

- At least 500 student pilots
- A fleet of at least 50 aircraft with integrated flight instrument system capabilities
- A Flight Data Monitoring system with data collection

- A scheduling system
- A robust and active Safety Management System

This assumption reflects the current state of most large CFR Part 141 flight training operations.

**Summary**

High-risk organizations, such as CFR Part 141 flight training organizations, are actively pursuing ways to increase their level of safety by incorporating improved risk assessment and decision-making techniques designed as fundamental parts within the system (Ale, Bellamy, Cooke, Goossens, Hale, Roelen, & Smith, 2006). Simulation modeling techniques are becoming more widely utilized in complex, high-risk systems across various domains to optimize the safety assessment process (Blair, 2017; Chen & Jing, 2016, Gunduz, Birgonul, & Ozdemir, 2017; Hadjimichael, 2009; Stonesifer, Calkin, Thompson, & Kaiden, 2014). However, existing SPIs, although useful in measuring the effectiveness of the organization's SMS, are incapable of providing a true predictive approach to safety decision-making (Patriarca, Di Gravio, Cioponea, & Licu, 2019). Aviation safety must continue to improve by considering new approaches to safety monitoring and decision-making that provide greater insight into *why* accidents occur and *how* safety is best achieved.

A safety decision-making tool, particular to large, collegiate CFR Part 141 flight training organizations yet adaptable to accommodate any flight training organization with data acquisition capabilities and an operational SMS, would allow for a more proactive approach to safety by assisting those in administrative roles with critical decision-making. Thus, the purpose of the research is to create and validate a safety performance decision-making tool based on a non-statistical risk assessment model, or SPI model, determined by Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020) to represent flight risk within large, collegiate CFR Part 141 flight training organizations. The validated model utilizes what-if scenarios to assess how modifying the controllable input variables impacts the overall level of risk within the organization's flight department and various other departments. Monte Carlo simulation was conducted and run to enable the SPI model to handle uncertainty in some of the key, influential variables.

In terms of significance, the extant literature indicated a deficit of predictive, safety performance decision-making tools specific to large, collegiate CFR Part 141 flight training organizations; therefore, this research fills an operational need within the industry. Additionally, the research enhances the depth of understanding of the factors most substantially contributing to flight risk within large, collegiate CFR Part 141 flight training organizations, thus advancing flight safety. As a safety decision-making tool, the model could also be used by the administration within a large, collegiate CFR Part 141 flight training organization to rationalize hiring, technology acquisition, and other safety-related enterprises by modeling the potential of modifying resources, or controllable inputs, without the risk associated with actually expending the organization's resources. The purpose of the research is to create and validate a safety performance decision-making tool to transform a nonstatistical model composed of domain-specific SPIs into a safety decision-making tool adaptable for use in any flight training organization with data gathering capabilities and an operational SMS. Additionally, the model provides a mechanism for expanding the breadth of knowledge related to optimizing resources from both flight and maintenance operations to enhance operational safety for CFR Part 141 flight training organizations. Providing large, CFR Part 141flight training organizations with a safety decision-making tool will enhance the risk management element of the organization’s SMS by taking an increasingly proactive approach to safety by providing insight into the impact changes to operating conditions may have on the safety of the overall operation.

**Definitions of Terms**

| | |
|---|---|
| 14 CFR Part 141 | This part prescribes the requirements for issuing pilot school certificates, provisional pilot school certificates, and associated ratings, and the general operating rules applicable to a holder of a certificate or rating issued under this part (Federal Aviation Administration, 2017). |
| Flight Data Monitoring | The analysis of flight data which allows safety managers to identify trends and fully investigate the circumstances behind events flagged (EASA, 2016). |
| Logistical Delay Time | The time from when a flight crew reports an aircraft as “down for maintenance” to the time the maintenance personnel opens a work order in order to address the discrepancy (Anderson et al., 2020). |
| Monte Carlo Simulation | A mathematical technique that uses randomly generated values for uncontrollable variables to |

| | |
|---|---|
| | model risk or uncertainty in a certain system (Dunn & Schultis, 2011). |
| Occurrences | Accidents or incidents. |
| Safety Culture | The attitudes, beliefs, perceptions, and values that employees share concerning safety in the workplace (Cox & Cox, 1991). |
| Safety Management System | SMS is the formal, top-down, organization-wide approach to managing safety risk and assuring the effectiveness of safety risk controls. It includes systematic procedures, practices, and policies for the management of safety risks (FAA Order 8000.369). |
| Safety Performance Indicator | A data-based parameter used for monitoring and assessing performance (ICAO, 2013b). |

**List of Acronyms**

| | |
|---|---|
| AHP | Analytic Hierarchy Process |
| ANOVA | Analysis of Variance |
| ASAP | Aviation Safety Action Program |
| ATC | Air Traffic Control |
| ICAO | International Civil Aviation Organization |
| ISO | International Organization for Standardization |
| FAA | Federal Aviation Administration |
| FLT | Flight |

| | |
|---|---|
| GSA | Generalized Sensitivity Analysis |
| MX | Maintenance |
| NAC | No Aircraft Available |
| NMAC | Near Mid-Air Collision |
| NTSB | National Transportation Safety Board |
| RPM | Revolutions Per Minute |
| SME | Subject Matter Expert |
| SMS | Safety Management Systems |
| SPI | Safety Performance Indicator |
| SSP | State Safety Program |

.

# CHAPTER II

# REVIEW OF THE RELEVANT LITERATURE

This chapter describes the present literature surrounding flight safety for CFR Part 141 flight training organizations; safety performance monitoring and measurement; justification surrounding the need for predictive rather than reactive safety monitoring; justification for the use of Monte Carlo simulation methods; and a detailed description of the theoretical foundation driving the research.

## Flight Safety for CFR Part 141 Flight Training Organizations

Within the United States, flight training is administered under the oversight of the Federal Aviation Administration (FAA) according to the federal regulations outlined in Title 14 of the Code of Federal Regulations (CFR) Parts 61, 141, or 142 (FAA, 2016). 14 CFR Part 141 flights schools are certified by the FAA and must meet strict standards to ensure optimal safety with requirements for personnel, aircraft, facilities, operational rules, and curriculum, allowing these organizations to train pilots more efficiently by reducing the flight hour requirements (Mendonca & Carney, 2017).

The International Civil Aviation Organization (ICAO) released the Safety Management Annex (Annex 19) in 2013, requiring participating ICAO member states to launch a State Safety Program (SSP) and implement an SMS (ICAO, 2013b). SMS provides CFR Part 141 flight training organizations with the ability to identify and mitigate safety risks before an accident occurring (Chen & Chen, 2014).

The Safety Risk Management element of the SMS is of particular importance, as flight training is inherently a high-risk activity (Cassens, 2015). Management and safety personnel within a CFR Part 141 are constantly making decisions on risk acceptability;

therefore, safety efforts must focus on the hazards posing the greatest risk to safe operations (Lu, 2016). The hazard identification process should encompass proactive, reactive, and predictive safety data collection techniques and approaches (ICAO, 2013b).

The severity of aircraft accidents is a particularly challenging variable to anticipate and predict (Bastos, 2005, Mendonca & Carney, 2017). Thus, risk analysis techniques, such as the use of risk matrices, for flight schools must take into consideration the pertinent safety attributes of the organization, including its safety culture, specific operational conditions, and the applicable safety standards (Mendonca & Carney, 2017). However, the safety effort of a 14 CFR Part 141 flight school will not succeed exclusively by adherence to standard operating procedures and company policy (ICAO, 2013b). Rather, SMS encourages taking a proactive approach to safety by continuing to develop and adapt to the safety risk management process. Introducing novel techniques to the safety assessment process beyond reactive risk matrices, such as a predictive safety decision-making tool, will transform the risk assessment process from reactive to predictive with very little risk involved.

**Safety Performance Monitoring**

Mitigative actions based on the analyses of previous accidents and incidents are both reactive and insufficient to further the progress of proactive safety management (ICAO, 2013). Additionally, the absence of accidents and incidents within CFR Part 141 flight training organizations does not assume operations are functioning at the optimum level of safety (Adjekum, 2014; Cassens, 2010; Keller, 2015; Mendonca & Carney, 2017). A modern approach to safety management includes proactively addressing safety risks rather than relying on inspections and remedial actions.

With the introduction and requirement of an SMS, the focus is shifting from archaic forms of reactive data collection and analysis toward approaches and techniques that bolster and improve the effectiveness of the organization's SMS. A vital portion of this process includes the development and implementation of safety performance indicators (SPIs). ICAO Doc 9859, Safety Management Manual, and ICAO Annex 19 define an SPI as a data-driven safety constraint used for observing and evaluating an organization's safety performance. SPIs are used to monitor and mitigate known safety risks to elicit corrective action before an adverse event occurring. Pierobon, 2016). (Pierobon, 2016).

**Safety performance indicators.** Quantitative performance indicators must be identified to achieve optimal safety within the organization. SPIs allow for the formation, execution, and review of safety policies within an organization (Reiman & Pietikäinen, 2010). Ensuring the SPIs meet the organization's predetermined safety goals has posed one of the greatest challenges to the development of a performance algorithm with risk prediction capabilities (Janicak, 2015). Thus, the particular safety requirements of the organization must be identified and prioritized throughout the process as pertinent SPIs are selected (Blair, 2017).

SPIs have been developed and utilized to improve the risk assessment process of various high-risk domains, including the aviation industry. Hadjimichael (2009) published a model founded on operational SPIs airlines can utilize to assess operational risk (Hadjimichael, 2009). Netjasov, Crnogorac, and Pavlović (2019) proposed a conflict risk assessment model composed of a set of seven SPIs specific to the Air Traffic Management system safety. Domain-specific SPIs have been useful in improving safety within the civil aviation domain as well (Chen and Li, 2016). Additionally, Panagopoulos, Atkin, and Sikora (2017) proposed a framework proposing how organizations could use SPIs for root-cause analysis of safety considerations. Findings exemplify the usefulness of SPIs in providing insight into the operating conditions of the organization as a whole.

**Existing methods for determining and measuring SPIs.** Effective safety management requires thoughtful consideration of the system and the processes driving the system; this cannot be achieved without some form of measurement (Safety Management International Collaboration Group, 2013). Rather than selecting SPIs based on convenience, SPIs must be selected with consideration given to the feedback required to ensure the organization's requirements for safety management can be effectively evaluated. The selection of SPIs can be determined through a systems analysis based on safety audit results (Jackman, 2018).

Focus groups utilizing SMEs are another approach to determining the most relevant SPIs (Anderson et al., 2020). The focus groups could also be used to develop the algorithms for each SPI and provide useful feedback on the selected SPIs. The use of focus groups over mathematical methods presents many advantages. For example, the use of focus groups is relatively inexpensive. The facilitated discussion process utilized to elicit information from focus groups allows expert participants to build upon each other's responses; this is useful for needs assessments and evaluation purposes (Leung & Savithiri, 2009). Focus groups also allow researchers to obtain more information from verbal, candid responses than may be obtained via survey methods. However, the focus group methodology is not without limitations. Focus groups rely heavily on facilitated discussion to produce results emphasizing the critical of the facilitator's skills as a moderator. Finally, the use of focus groups makes the findings more difficult to generalize to the larger population due to the inherent weakness of the focus group selection process (Leung & Savithiri, 2009).

**Forecasting to improve safety outcomes.** Aviation safety has been managed based on analyzing accidents and incidents after they have already occurred. Although this strategy has allowed the industry to make strides in improving safety, a major drawback is the reactive nature of this approach; as, safety analysis based on hindsight has restricted the process to primarily focusing on innately negative aspects, such as errors and failures within the system (Patriarca, Di Gravio, Cioponea, & Licu, 2019). Rather, the cyclical approach of measuring, analyzing, and providing feedback through a robust SMS has the potential to provide a more holistic, data-driven approach to safety monitoring. Thus, rather than focusing solely on historical events or reports monitoring should take a more proactive approach by assessing the various components of the system and how they contribute to the functioning of the system as a whole. This could be accomplished by incorporating forecasting techniques into the safety risk management element of an organization's SMS to aid in further understanding the performance variability that occurs within complex systems like aviation.

Traditionally, organized institutions have been relatively resistant to change (Jepperson, 1991; Verweijen & Lauche, 2019). As organizations adhere to institutional standards, such as those prescribed by the FAA under CFR Part 141 operating conditions, safety monitoring practices could become increasingly taken-for-granted leading to problems within the operation (Verweijen & Lauche, 2019). In high-hazard industries, this lack of adaptability can undermine organizational safety. Despite statistically high safety rates within the aviation domain, occurrences continue to take place. Researchers argue that a proactive, systematic analysis of safety risk management utilizing modern techniques, such as forecasting, could aid in evolving an industry that has been traditionally reactive to one that is proactive in its risk assessment process (Dyhrberg & Jensen, 2004; Insua, Alfaro, Gomez, Hernandez-Coronado, & Bernal, 2019; Verweijen & Lauche, 2019).

Although risk matrices have been able to provide qualitative assessments of risk on an ordinal scale, risk matrices provide little insight into the consequences of various choices made by the organization and how these consequences impact the system as a whole. Rather, forecasting models provide sophisticated methods to assess aviation safety occurrence outcomes to bolster an aviation organization's safety risk management practices (Insua et al., 2019). Further, forecasting models could be utilized for decision-making purposes by aviation authorities, insurance companies, aviation operators, aviation companies, and aviation training facilities (Insua et al., 2019).

## Monte Carlo Simulation

Monte Carlo simulation provides a useful methodology to propagate uncertainties further evolving reactive safety models and indices into innovative and predictive models useful for forecasting safety performance (Hacura, Jadamus-Hacura, & Kocot, 2001). Monte Carlo methods use repeated random sampling to estimate the many potential outcomes that cannot be determined with certainty. This is accomplished by modeling ranges of potential values where uncertainty exists by analyzing the combination of outputs produced by the model. Thus, the outputs provide a range of possible outcomes as well as a probability density curve used to determine outcome frequency. Monte Carlo simulation is particularly useful for modeling complex systems where uncertainty exists to assess the impact of risk. Monte Carlo methods have led to several innovative improvements in various fields such as physics, game theory, finance, maritime, nuclear, and aviation (Hacura, Jadamus-Hacura, & Kocot, 2001).

The selection of either analytical (e.g. point estimate methods) or simulation methods (e.g. Monte Carlo simulation) will be shaped by the following considerations (Safety and Reliability Society, n. d., p. 3):

- Complexity of the system
- Scope
- Accuracy
- Future development
- Application

**Advantages of Monte Carlo simulation.** According to Stolzer and Goglia (2015), Monte Carlo methods have many appealing characteristics over point estimate methods. Monte Carlo simulation methods provide researchers with more valuable information than point estimate methods; account for inherent uncertainties; and provide the location of any specific risk estimate allowing for a level of risk to be selected within the model that corresponds to the desired level of risk protection (Stolzer & Goglia, 2015). From a research perspective, the process of building the simulation can also enhance the depth of understanding of the true system. Monte Carlo methods can be used for sensitivity analysis and system optimization without impacting the real system (Spall, 2003). Using Monte Carlo methods allows for improved control over experimental conditions within the modeled system. Finally, researchers can either compress or expand time within the model, something not possible working within the limitations of the real system (Spall, 2003).

The usefulness of Monte Carlo methods is echoed by Faghih-Roohi, Xie, and Ng (2014), who used Monte Carlo simulation as an analytical approach to accident risk modeling in the maritime environment. Faghih-Roohi et al. (2014) support Monte Carlo simulation applications for risk modeling due to the probabilistic attributes associated with risk. Basic statistics, such as summary statistics or accident rates, are not adequate for long-term risk prediction, further testifying the usefulness of Monte Carlo methods to evaluate risk amidst extensive uncertainties (Faghih-Roohi et al., 2014).

**Disadvantages of Monte Carlo simulation.** However, simulating the real system using Monte Carlo methods does pose several disadvantages. For example, depending on the commercial simulation software packages used, it may be very costly and time-consuming to build a simulation. Further, Monte Carlo simulation relies on random number generation to solve deterministic problems; therefore, it is possible that a simulation could be stretched beyond the limits of credibility influencing the validity of the model when using commercially-sold software packages due to their lack of consideration of the underlying assumptions and limitations determined by the researcher (Spall, 2003). Another potential disadvantage is that Monte Carlo simulation provides several, perhaps millions, of runs at given input values, whereas analytical solutions provide exact values (Spall, 2003).

**Monte Carlo process and tools.** Monte Carlo simulations perform risk analysis in complex systems by creating a model of potential results by using probability distributions for any variable within the model that has inherent uncertainty. Evaluating the outputs of probability distributions allows for a much more realistic method of describing uncertainty. The fundamental steps of conducting a Monte Carlo simulation are as follows:

1. Define the problem and simulation features
2. Identify the key components and variables within the model
3. Define input parameters, including probability distributions and equations, for each variable
4. Define simulation scenarios
5. Select control values that will be manipulated
6. Run the simulation with a predetermined amount of trials (e.g. 1,000 trials)
7. Analyze the results of the output tables using both descriptive statistics and sensitivity analysis to test edge cases
8. Either return to Step 4 and redefine the next scenario or choose to complete the simulation at this point (Ayres, Schmutte, & Stanfield, 2017)

A computer is required to run Monte Carlo simulations. In addition to a basic PC spreadsheet, various probabilistic simulation platform software exists to run Monte Carlo simulations, such as Analytica by Lumina Decision Systems. Analytica is software for developing and evaluating quantitative decision models for modeling risk and uncertainty.

**Monte Carlo applications in aviation research.** Safety performance assessment, based on advanced risk assessment methodologies, is a pressing challenge within the air transport and training sector (Di Gravio, Mancini, Patriarca, & Costantino, 2015).

Historically, the aviation domain has used simple metrics such as accident rates to gauge safety performance; however, reactive metrics are not representative of the level of safety present across the various facets of the system (Di Gravio et al., 2015).

The FAA and the EUROCONTROL Performance Review Commission have identified shared performance indicators to monitor safety by proposing a standard occurrence reporting and assessment plan defined under ESARR 2 Appendix A and B (EUROCONTROL, 2009). ESARR 2 Appendix A and B outlines the process of collecting and recording the information elicited from safety occurrence reports. This plan was developed based on James Reason's Swiss Cheese Model of Accident Causation, which relates organizational failures to an alignment of metaphoric "holes" or weaknesses in the system so when these holes line up, a hazard slips through the holes of the various layers of defenses leading to drift into failure (Dekker, 2011; Reason, 1997).

However, safety assessment must consider the potential impact of any safety-related event. Minor, or less serious events, may happen more frequently testifying the importance of including occurrence statistics rather than solely accident statistics (Di Gravio et al., 2015). Using proactive safety indicators, Monte Carlo simulation has the potential to provide an analytical model, based on historic data distributions, allowing the decision-maker to model potential events and determine how these less serious events, or occurrences, impact the safety of the system.

Over the past decade, Monte Carlo simulation has been used for modeling and calculating aircraft collision risk both on the ground and in the air. Jacquemart and Morio (2013) created a Monte Carlo simulation to evaluate conflict probabilities between aircraft, demonstrating the utility of Monte Carlo simulation for air transportation safety.

Belkhouche (2013) utilized Monte Carlo simulation for collision risk modeling and assessment for autonomous air vehicles to calculate the probability of a mid-air collision occurring in the presence of uncertainties. According to Belkhouche (2013), Monte Carlo methods have an important advantage in aircraft collision risk modeling because it does not explicitly use speed and orientation information, such as collision cone angles, to calculate the probability of a collision occurring in the presence of uncertainties in non-linear systems with non-Gaussian, or non-normal, distributions; rather, collision risk is expressed as simple inequalities allowing for the estimation of probability under difficult and varying scenarios. In their text, Dunn and Shultis (2011) exemplify the application of Monte Carlo methods across domains and situations of varying complexity. Careddu, Costantino, and Di Gravio (2008) and Stroeve, Blom, and Bakker (2013) have used Monte Carlo methodologies to validate advancements made on runway incursion events. Di Gravio, Mancini, Patriarca, and Costantino (2015) conducted a study aimed at improving Air Traffic Management safety by creating a statistical model of safety events using Monte Carlo simulation to predict safety performance, further validating the utility of Monte Carlo simulation in improving air transportation safety. However, the extant literature indicates a deficit of Monte Carlo simulation models to be used as safety decision-making tools specific to CFR Part 141 flight training organizations.

Based upon a review of the relevant literature and due to the influx of uncertainties and daily variability in the air traffic system, Monte Carlo is an appropriate method for forecasting safety performance within the aviation industry. Further, for the purpose of the research, Monte Carlo is the most appropriate methodology due to a large majority of the input variables being subject to uncertainty.

## Gaps in the Literature

Although forecasting methods, such as Monte Carlo simulation, have grown in application in the aviation sector over the past decade, the industry still relies heavily on reactive processes, such as risk matrices alone and SPIs, within their SMS risk assessment process. SPIs, although useful in measuring the effectiveness of the SMS, are incapable of providing a true predictive approach to safety. A thorough review of the extant literature indicated a gap in the process of transitioning from traditionally reactive SPIs into safety decision-making tools with forecasting abilities for safety decision-making purposes specific to CFR Part 141 flight training operations. Further research is needed to transform SPIs within a non-statistical model into predictive safety decision-making tools capable of taking proactive safety one step further by modeling the potential of the system without compromising resources.

The extant literature also indicates a deficit of validated models capable of utilization as decision-making tools specific to CFR Part 141 flight training organizations. Two commercial airlines, legacy carrier, Southwest Airlines, and a large, low-cost carrier based out of Brazil, are currently in the process of developing safety performance tools based on risk assessment models composed of domain-specific SPIs; however, these models do not utilize simulation and are reactive in nature. The models developed by these air carriers apply to commercial operations and would be difficult to adapt to flight training operations. Thus, the research conducted for this dissertation fills an operational need within the industry.

This research fills these gaps by providing a validated safety decision-making tool, specific to CFR Part 141 flight training operations, bolstering the research in the area of proactive, rather than reactive, aviation risk assessment techniques. The model could also be adapted to accommodate the operational needs of any flight training organization with data procurement abilities and an operational SMS.

**Theoretical Framework**

The theoretical framework driving the research was founded upon a non-statistical model developed by Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020). Anderson et al. (2020) conducted a sequential, mixed-method design study including a qualitative data collection and analysis phase, followed by a quantitative data collection and analysis phase.

Subject Matter Experts (SMEs) in the area of CFR Part 141 maintenance and flight operations selected the appropriate Safety Performance Indicators (SPIs). Once the appropriate SPIs had been selected, formulas were developed to quantify each selected SPI, based on monthly, operational-performance data collected by a CFR Part 141 flight school in the Southeast region of the United States. The Risk Indicator Score Card was developed to compute a standardized risk score for each month of both flight and maintenance operations. Expert elicitation was used to establish inter-rater reliability for the assessment of SMEs' evaluations.

Twelve SPIs were selected for use within the model. SPIs 1-6 MX encased the maintenance side of operations; SPIs 1-6 FLT includes indicators relevant to flight operations. The SPIs, variables, and brief descriptions can be found in Table 1. Table 2 outlines the SPIs and their quantifiable formulae.

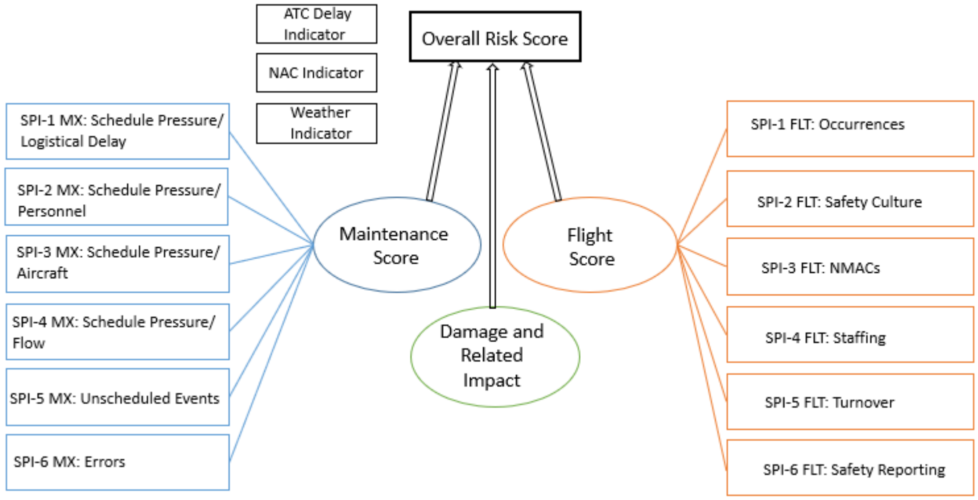



*Figure 1*. Diagram of the non-statistical model developed by Anderson et al. (2020) composed of SPIs and associated indicators.

Table 1

*Safety Performance Indicators and Attributing Variables*

| SPI | Variables | Description |
|---|---|---|
| SPI-1 MX: Schedule Pressure | Logistical Delay Time (minutes) | Used to measure the schedule pressure faced by personnel; provides insight into the efficiency of the operation; saturation indicator. |
| SPI-2 MX: Schedule Pressure/ Personnel | Aviation technicians available<br>Fleet flight time | Used to determine whether there are too few technicians available increasing the likelihood of an error occurring. |
| SPI-3 MX: Schedule Pressure/ Aircraft | Percentage of aircraft available<br>Total aircraft in fleet | Analyzes the schedule pressure technicians experience by assessing the number of aircraft down for maintenance relative to the total number of aircraft available in the fleet. |

| SPI | Variable | Descriptions |
|---|---|---|
| SPI-5 MX: Unscheduled Events | Unscheduled maintenance orders under $10k<br>FAA occurrences reports<br>Fleet flight time | Measures the oversights made by technicians; although rare, selected due to the catastrophic nature associated with errors committed by maintenance personnel. |
| SPI-6 MX: Errors | Number of aircraft dispatched with maintenance errors<br>Number of total work orders processed | Selected to capture the total volume of maintenance orders processed related to fleet flight time; indicates the overall health of the operation and insight into when a safe threshold of schedule pressure may have been exceeded. |
| SPI-1 FLT: Occurrences | Number of reported tail strikes<br>Number of hard landings<br>Number of unstabilized approaches<br>Number of RPM overspeeds<br>Number of over/under G exceedances<br>Number of flap overspeeds<br>Fleet flight time | Selected as a general assessment of how safely the aircraft are being flown. |
| SPI-2 FLT: Safety Culture | Safety culture survey criterion<br>Number of safety culture surveys received | Based upon the institution’s yearly safety culture survey designed to annually assess the state of the organization’s safety culture. |
| SPI-3 FLT: NMACs | Number of traffic conflicts<br>Fleet flight time | Chosen for tracking internal traffic conflicts. |
| SPI-4 FLT: Staffing | Number of full-time equivalent instructor pilots (average weekly)<br>Active flight students (average weekly) | Selected to assess the level of saturation within the flight department to ensure there are enough flight instructors staffed to meet flight student demands. |

| SPI | Variables | Description |
|---|---|---|
| SPI-5 FLT: Turnover | Number of months flight instructors are active at the institution | Selected to measure the average experience level of instructor pilots working at the institution; it was assumed a correlation exists between the level of experience and safety. |
| SPI-6 FLT: Safety Reporting | Number of events reported | Selected as an assessment of safety and the climate of the organization's reporting culture. |
| Damage and Related Impact | Number of NTSB accident reports<br>Number of FAA incident reports<br>Number of unscheduled maintenance reports > $10,000<br>Fleet flight time | Provides a comprehensive, external perception of the risk associated with the operation. |

Table 2

*Safety Performance Indicators and Quantifiable Formulae*

| SPI | Formulae |
|---|---|
| SPI-1 MX: Schedule Pressure/ Logistical Delay | $Logistical\ delay\ time\ (minutes)$ |
| SPI-2 MX: Schedule Pressure/ Personnel | $\frac{Aviation\ maintenance\ technicians\ aviailable}{Fleet\ flight\ time}$ |
| SPI-3 MX: Schedule Pressure/ Aircraft | $\frac{Percentage\ of\ aircraft\ avaialble}{Total\ aircraft\ in\ fleet}$ |

| SPI | Formulae |
|---|---|
| SPI-4 MX: Schedule Pressure/ Flow | $\frac{Number\ of\ total\ maintenance\ orders\ processed}{Fleet\ flight\ time}$ |
| SPI-5 MX: Unscheduled Events | $\frac{Unscheduled\ maintenance\ orders\ under\ \$10k + FAA\ occurrences\ repor}{Fleet\ flight\ time}$ |
| SPI-6 MX: Errors | $\frac{Number\ of\ aircraft\ dispatched\ with\ maintenance\ error}{Number\ of\ total\ work\ orders\ processed}$ |
| SPI-1 FLT: Occurrences | $\frac{(reported\ tail\ strikes) + (number\ of\ hard\ landings) + (number\ of\ unstabilized\ appr\text{c} \\ (Number\ of\ RPM\ overspeeds) + (number\ of over/underg) \\ +(number\ of\ flap\ overspeeds)}{Fleet\ flight\ time}$ |
| SPI-2 FLT: Safety Culture | $\frac{(0.039 * PI3) + (0.064 * SO3) + (0.079 * EFLS3) + (0.085 * EFLS8) + (0.092 * PS1) \\ +(0.081 * PS3) + (0.067 * PS7) + (0.043 * PRS1) + (0.07 * EC2) + \\ (0.072 * RS2) + (0.043 * QNH4) + \\ (0.032 * QNH5) + (0.018 * MO1)}{Number\ of\ surveys\ collected}$ |
| SPI-3 FLT: NMACs | $\frac{Number\ of\ traffic\ conflicts}{Fleet\ flight\ time}$ |
| SPI-4 FLT: Staffing | $\frac{Number\ of\ full\ time\ equivalent\ instructor\ pilots}{Active\ flight\ students}$ |
| SPI-5 FLT: Turnover | $Number\ of\ months\ as\ an\ instructor\ pilot\ at\ the\ institution$ |
| SPI-6 FLT: Safety Reporting | $Number\ of\ events\ reported\ (ASAP\ and\ internal)$ |
| Damage and Related Impact | $\frac{(\left(\begin{matrix} NTSB\ accident\ reports * Impact\ Value) + FAA\ incident\ reports + \\ Unscheduled\ MX\ reports > 10K \end{matrix}\right)}{Fleet\ flight\ time}$ |

**The Risk Indicator Score Card.** Ultimately, an individual standardized risk score, as well as an overall risk score, was developed to calculate the monthly level of risk associated with flight and maintenance operations. Figure 2 demonstrates the nonstatistical model output before its transformation into the Risk Indicator Score Card. Depicted in Figure 3, the model output transitions into the Risk Indicator Score Card, representing risk on a 0-5 risk scale. Figure 4 demonstrates the display format potential of the Risk Indicator Score Card for utility and ease of use.

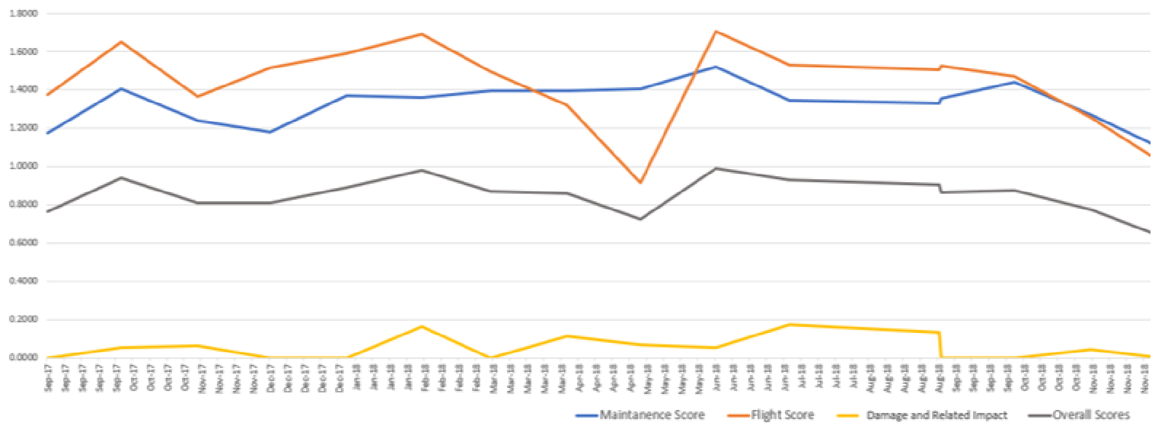
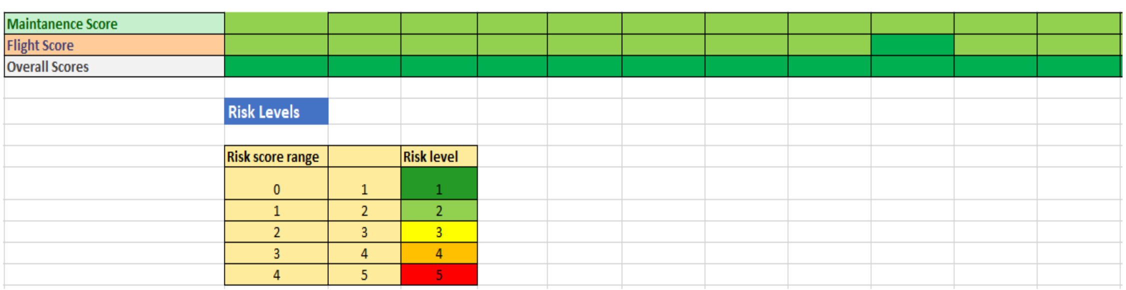



*Figure 2*. Nonstatistical model output prior to Risk Indicator Score Card transition.
*Note*. Notational data only. Not representative of actual data or performance.



| | | |
|---|---|---|
| Maintanence Score | | |
| Flight Score | | |
| Overall Scores | | |
| Risk Levels | | |
| Risk score range | | Risk level |
| 0 | 1 | 1 |
| 1 | 2 | 2 |
| 2 | 3 | 3 |
| 3 | 4 | 4 |
| 4 | 5 | 5 |

*Figure 3*. Output of the Risk Indicator Score Card.
*Note*. Notational data only. Not representative of actual data or performance.

| Risk score range | | Risk level |
|---|---|---|
| 0 | 1 | 1 |
| 1 | 2 | 2 |
| 2 | 3 | 3 |
| 3 | 4 | 4 |
| 4 | 5 | 5 |

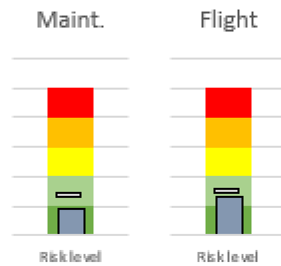



*Figure 4.* Display potential for the Risk Indicator Score Card.
*Note.* Notational data only. Not representative of actual data or performance.

**Similar efforts.** Southwest Airlines and a Brazilian low-cost carrier are conducting similar efforts relevant to commercial flight operations. Both airlines are in the process of developing or have developed an algorithm that provides a risk score for both the operation and individual safety scores for each department (Southwest Airlines, 2019). Using the foundations of ICAO Annex 19 and FAA guidance, Mendonca and Carney (2017) have also developed a model for CFR Part 141 operators; however, the model focuses specifically on using the four components of SMS and is intended to encourage a thriving safety culture among CFR Part 141 operatives. Additionally, the model developed by Mendonca and Carney (2017) has no predictive capabilities.

**From reactive to predictive safety monitoring.** The research conducted for this dissertation transformed the non-statistical model developed by Anderson et al. (2020) into a predictive, Monte Carlo simulation composed of real-time data input for the chosen SPIs. The Monte Carlo simulation is useful for safety decision-making to run what-if scenarios for the assessment of how variations to input variables impact the overall level of operational risk within the organization’s flight department.

## Summary

Under Annex 19, the International Civil Aviation Organization (ICAO) requires members to establish a State Safety Program, requiring certain services be provided to implement a Safety Management System (SMS). SMS provides CFR Part 141 flight training organizations with the ability to foresee and mitigate potential safety risks before an adverse event occurs (Chen & Chen, 2014). SMS encourages taking a proactive approach to safety by continuing to develop and adapt the safety risk management process.

Traditionally, aviation safety has been managed on the basis of analyzing accidents and incidents after they have already occurred. Although this strategy has allowed the industry to make strides in improving safety, a major drawback is the reactive nature of this approach; safety analysis based on hindsight has restricted the process to primarily focus on innately undesirable aspects within the system (Patriarca, Di Gravio, Cioponea, & Licu, 2019). Effective safety monitoring should take a proactive approach by assessing the various components of the system and how they contribute to the functioning of the system as a whole. SMS has traditionally utilized SPIs to supervise known safety risks and expose developing risks to elicit corrective action before an adverse event occurring. SPIs play a valuable role in SMS by enabling performance-based safety management. However, SPIs are reactive in nature; therefore, introducing novel techniques to safety assessment beyond reactive risk matrices, such as a predictive safety decision-making tool, will transform the risk assessment process from reactive to predictive with very little risk involved. This is a modern approach to safety management and includes safety risks being addressed proactively rather than relying on inspections and remedial actions.

Researchers also argue that a proactive, systematic analysis of safety risk management utilizing modern techniques, such as forecasting, could aid in evolving an industry that has traditionally been reactive to one that is proactive in its risk assessment process (Dyhrberg & Jensen, 2004; Insua, Alfaro, Gomez, Hernandez-Coronado, & Bernal, 2019; Verweijen & Lauche, 2019). Forecasting models provide sophisticated methods to assess the outcomes of aviation safety occurrences to bolster an aviation organization's safety risk management practices (Insua et al., 2019). Further, forecasting models could be utilized for decision-making purposes by aviation authorities, insurance companies, aviation operators, aviation companies, and aviation training facilities (Insua et al., 2019).

Monte Carlo simulation provides a useful methodology to account for uncertainties in the model's predictive algorithms and allows for the modeling of intricate systems where uncertainty exists, or random variables are involved, to assess the impact of risk without impacting the real system (Spall, 2003). Over the past decade, Monte Carlo simulation has been used for modeling and calculating aircraft collision risk both on the ground and in the air. However, the extant literature indicates a deficit of Monte Carlo simulation models to be used as safety decision-making tools specific to CFR Part 141 flight training organizations. Based on a review of the relevant literature and due to the influx of uncertainties and daily variability in the air traffic system, Monte Carlo is an appropriate method for forecasting risk within the aviation industry.

A thorough review of the extant literature indicated a gap in the process of transitioning from traditionally reactive SPIs into safety decision-making tools with forecasting abilities for safety decision-making purposes specific to CFR Part 141 flight training operations. The extant literature also indicates a deficit of validated models capable of utilization as decision-making tools specific to CFR Part 141 flight training organizations. This research fills these gaps by providing a validated safety decision-making tool, specific to CFR Part 141 flight training operations, bolstering the research in the area of proactive aviation safety assessment techniques. The model could also be adapted to accommodate the operational needs of any flight training organization with data procurement abilities and an operational SMS.

The theoretical framework driving the research is the non-statistical model developed by Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020). Anderson et al. (2020) built and validated, via expert elicitation, a non-statistical model composed of 12 Safety Performance Indicators (SPIs) encompassing both flight and maintenance operations. The data is based on a two-year sample of operational performance data from a 14 CFR Part 141 flight training facility in the southeastern United States. SPIs were used to develop a Risk Indicator Score Card depicting the level of monthly, operational risk for flight operations; maintenance operations; and overall, combined operations.

# CHAPTER III

# METHODOLOGY

The dissertation utilized the Monte Carlo simulation method to build a safety decision-making tool based on SPIs determined by Anderson et al. (2020) to represent flight risk within large, collegiate CFR Part 141 flight training organizations to evaluate predictive, what-if scenarios to evaluate how the variations to controllable input variables affect the risk score outputs indicating the level of risk posed to safe operating conditions. The study did not involve human subject testing or data collection from human subjects; thus, the research did not require Institution Review Board (IRB) approval.

## Research Method Selection

The study used the quantitative method to convert the non-statistical model developed by Anderson et al. (2020) using Monte Carlo simulation into a safety decision-making tool to run what-if scenarios to assess how modifications to the controllable input variables impact the level of operational risk within an organization's flight department. The use of Monte Carlo simulation is valuable in accommodating the uncertainty and variability of 22 uncontrollable input variables, as the only controllable input variables are the four listed below. The remaining variables were subject to uncertainty.

- The number of full-time instructor pilots,
- The number of aviation maintenance technicians available,
- The number of active flight students, and
- The total number of aircraft in the fleet.

Papadopoulos and Yeung (2001) describe many advantages for using Monte Carlo simulation to address uncertainty, including the ability of Monte Carlo simulation to handle large amounts of uncertainty within the input variables and the lack of concerns regarding the interactions between input variables. Faghih-Roohi et al. (2014) support the use of Monte Carlo simulation for risk modeling due to the probabilistic attributes associated with risk; whereas basic statistics, such as summary statistics or accident rates, are insufficient for long-term risk prediction. Monte Carlo methods also allow for sensitivity analyses and evaluation of the system without the need to operate the real system, leaving valuable resources uncompromised (Spall, 2003). Thus, the model created for the purpose of this dissertation can exemplify the effects of uncertainty in typical, collegiate flight operations by simulating many thousand potential outcomes to generate an accurate representation of the range of probable outcomes given the uncertainty of the uncontrollable input variables (Farrance & Frenkel, 2014).

The current research addresses the following research questions:

1. How can the SPI model developed by Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020) be transformed into a predictive, safety performance decision-making tool with the ability to run what-if scenarios?
2. How do changes to the controllable input variables impact the overall risk performance score?

To address Research Question 1, Monte Carlo simulation was utilized to transform the non-statistical risk assessment model composed of SPIs developed by Anderson et al. (2020) into a predictive, safety performance decision-making tool. Research Question 2 was answered by utilizing distributions and ranges of values to simulate the many thousands of potential outcomes within the what-if scenarios allowing for an assessment of how the variations to the controllable input variables influence the overall level of operational risk. After manipulating the controllable input variables, or resources with respect to personnel, students, and aircraft, the probability distribution output from the what-if scenarios then allows safety personnel and administration to make more informed safety-related decisions, based on the level of risk predicted by the what-if scenarios, without expending unnecessary resources.

## Population and Sample

**Population and sampling frame.** The target population to which the model generalizes is large, collegiate CFR Part 141 flight training organizations within the United States operating under the specifications defined by the FAA within Title 14 of the Code of Federal Regulations Part 141 (Federal Aviation Administration, 2017). The sampling frame consisted of two-years of operational data from both flight and maintenance operations dating from September 2017 to September 2019 for a large, collegiate CFR Part 141 flight training organization in the southeastern United States.

**Sample size.** The sample data used to determine the probability distributions of the uncontrollable input variables within the model was comprised of two years of operational flight and maintenance data from a large, collegiate 14 CFR Part 141 flight training organization in the southeastern United States. Monte Carlo simulation utilizes probability distributions drawn from raw operational data to simulate the vast range of operating conditions within large, CFR Part 141 operations.

**Sampling strategy.** To ensure simulation scenarios are representative of the target population, true operational ranges representative of a large, collegiate 14 CFR Part 141 flight training organizations in the United States were used to enhance the generalizability of the model. The study conducted simulation runs based on the true operational ranges specified below to simulate the range of operating conditions possible within a large, collegiate CFR Part 141 flight training organization with varying levels of resources with respect to personnel (Aviation Maintenance Technicians and Instructor Pilots), students, and aircraft:

- Aviation Maintenance Technicians available: 14-35
- Aircraft available: 50-82
- Full-time Instructor Pilots: 100-200
- Active Flight Students: 335-1300

These ranges were selected because they are reflective of the higher and lower operational limits of the sample data drawn from a large, collegiate CFR Part 141 flight training operation in the southeastern United States. The model could easily be adapted for use in both small and large CFR Part 141 flight training organizations and any flight training organization with data procurement abilities and an operational SMS.

## Data Collection Process

**Design and procedures.** This section describes the design and use of the mathematical model in detail. Figure 5 depicts the structural definition of the model in Analytica. The green-colored squares depict the four controllable input variables. The light blue-colored ovals represent the 22 uncontrollable input variables specified as probability distributions supplying an array of random values to the model based on probability distributions drawn from the raw data sample. The blue rounded rectangular boxes are SPIs from the non-statistical model developed by Anderson et al. (2020) and depict calculation nodes producing the results of the model. The equations driving these calculations can be found in Table 5 and will be described further later in this section.

The orange trapezoid represents a value that is input as a constant. The impact value was input into the model as a constant value as injuries and damage are challenging to predict due to their variability in nature. Thus, a constant value of 1 indicated no damage or injuries incurred was selected for the purpose of this dissertation. The pink hexagons represent the risk score output variables.

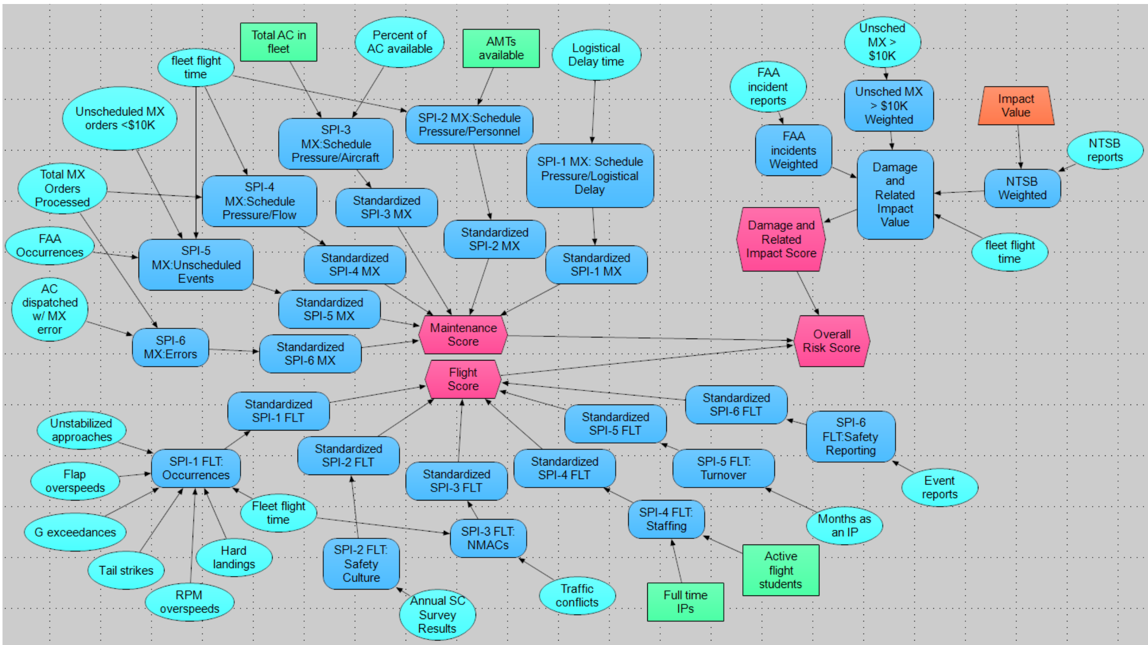



*Figure 5*. Structural definition of the model in Analytica.

**Monte Carlo process and steps**. The steps involved in preparing, creating, and running a Monte Carlo simulation can be found in Figure 6.

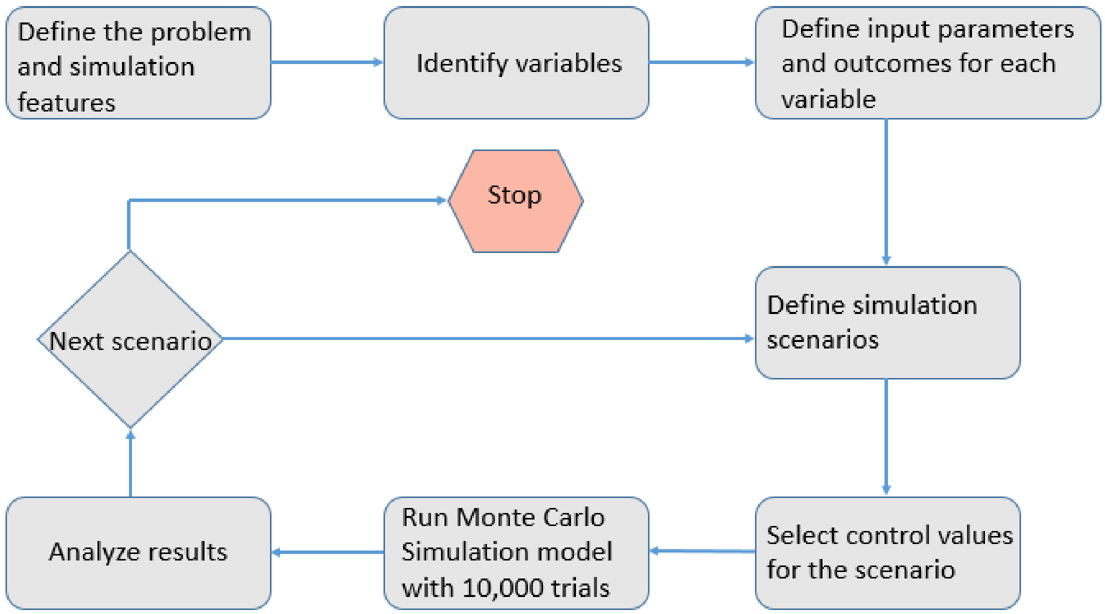



*Figure 6.* Monte Carlo steps and processes overview adapted from "Stats: data and models," by R.D. De Veaux, P. F. Velleman, and D. E. Bock, 2012. Copyright 2012 by Pearson Education, Inc.

*Step 1-Defining the problem and simulation features***.** The first step involves identifying the problem, scope, and research questions for driving model development. A detailed discussion of the problem, scope, and research questions driving the research can be found in Chapter 1 of this manuscript.

*Step 2- Variable identification.* The next step involves identifying the key components and variables within the model. The input variables for the model were selected based on the contributing variables relevant to each SPI within the non-statistical model developed by Anderson et al. (2020). The variables, relevant SPIs, and categorization as either uncontrollable or controllable input variables can be found in Table 3.

Table 3

*Input and Output Variables for the Model*

| Relevant SPI | Variables | Variable Type |
|---|---|---|
| | Fleet flight time (hobbs) | Input Uncontrollable |
| SPI-1 MX: Schedule Pressure | Logistical Delay Time (minutes) | Input Uncontrollable |
| SPI-2 MX: Schedule Pressure/ Personnel | Technicians available | Input Controllable |
| SPI-3 MX: Schedule Pressure/ Aircraft | Percentage of aircraft available | Input Uncontrollable |
| | Total aircraft in fleet | Input Controllable |
| SPI-4 MX: Schedule Pressure/ Flow | Number of total maintenance orders processed | Input Uncontrollable |
| SPI-5 MX: Unscheduled Events | Unscheduled maintenance orders under $10k | Input Uncontrollable |
| | FAA occurrences reports | Input Uncontrollable |
| SPI-6 MX: Errors | Number of aircraft dispatched with maintenance errors | Input Uncontrollable |
| SPI-1 FLT: Occurrences | Number of reported tail strikes | Input Uncontrollable |
| | Number of hard landings | Input Uncontrollable |
| | Number of unstable approaches | Input Uncontrollable |
| | Number of RPM overspeeds | Input Uncontrollable |
| | Number of G exceedances | Input Uncontrollable |
| | Number of flap overspeeds | Input Uncontrollable |
| SPI-2 FLT: Safety Culture | Number of surveys collected | Input Uncontrollable |
| | Factor Scores | Input Uncontrollable |

| Relevant SPI | Variables | Variable Type |
|---|---|---|
| SPI-3 FLT: NMACs | Number of traffic conflicts | Input Uncontrollable |
| SPI-4 FLT: Staffing | Number of full-time equivalent instructor pilots (Average weekly number) | Input Controllable |
| | Active flight students (Average weekly number) | Input Controllable |
| SPI-5 FLT: Turnover | Number of months flight instructors are active at institution (average) | Input Uncontrollable |
| SPI-6 FLT: Safety Reporting | Number of events reported (ASAP and event) | Input Uncontrollable |
| Damage and Related Impact | Number of NTSB accident reports | Input Uncontrollable |
| | Number of FAA incident reports | Input Uncontrollable |
| | Number of unscheduled maintenance reports > $10,000 | Input Uncontrollable |
| Outputs | Maintenance Score | Output |
| | Damage and Related Impact Score | Output |
| | Flight Score | Output |
| | Overall Risk Score | Output |

*Step 3- Defining parameters*. The third step involved defining the input parameters for each variable. This included defining the probability distribution of the data relevant to each variable and the associated equations for each SPI to the variable's parameters. To accomplish this, a two-year sample of operational data from a large, collegiate CFR Part 141 flight training organization was analyzed.

*Determining the distributions for uncontrollable inputs.* The distributions for the uncontrollable inputs were derived from a two-year sample of operational data from a large, collegiate CFR Part 141 flight training organization in the southeastern United States. Utilizing Minitab 19 statistical software, the sample of data for each uncontrollable input was run through Minitab 19 to identify the distributions of the data.

Each sample of data produced a Goodness of Fit Test table and probability plots to visually identify the distributions. The visual probability plots were used to initially determine distributions of the data, and the *p*-values from the Goodness of Fit Test table were used to validate the distributions. The distributions for the uncontrollable inputs can be found in Table 4.

Table 4

*Probability Distributions for Uncontrollable Input Variables*

| Uncontrollable Input Variable | Data Type | Probability Distribution |
|---|---|---|
| Fleet flight time (hobbs) | Continuous | Normal |
| Logistical Delay Time (minutes) | Continuous | Weibull |
| Percentage of aircraft available | Discrete | Uniform |
| Number of total maintenance orders processed | Discrete | Logistic |
| Unscheduled maintenance orders under $10k | Discrete | Binomial |
| FAA occurrences reports | Discrete | Geometric |
| Number of aircraft dispatched with maintenance errors | Discrete | Bernoulli |
| Number of reported tail strikes | Discrete | Poisson |
| Number of hard landings | Discrete | Poisson |
| Number of unstable approaches | Discrete | Lognormal |
| Number of RPM overspeeds | Discrete | Poisson |
| Number of G exceedances | Discrete | Poisson |
| Number of flap overspeeds | Discrete | Poisson |

| Uncontrollable Input Variable | Data Type | Probability Distribution |
|---|---|---|
| Number of traffic conflicts | Discrete | Binomial |
| Number of months flight instructors are active at institution (average) | Continuous | Certain |
| Number of events reported (ASAP and event) | Discrete | Negative Binomial |
| Number of NTSB accident reports | Discrete | Binomial |
| Impact value | Discrete | Certain |
| Number of FAA incident reports | Discrete | Binomial |
| Number of unscheduled maintenance reports > $10,000 | Discrete | Poisson |

*Defining outcome equations.* Once the distributions have been determined, the outcome for each component of the model was defined. This was accomplished by defining the associated equations for each SPI to the variable's parameters in Analytica. Table 5 delineates each SPI, the Damage and Related Impact variable, and their associated equations that were used within the model. The mathematical algorithms and concepts used for the simulations were derived from focus group participants and SMEs in the areas of flight and maintenance operations at a large, collegiate CFR Part 141 flight training organization and were externally validated utilizing an independent group of external SMEs in the area of commercial flight safety operations (Anderson et al., 2020).

Table 5

*Model Equations*

| SPI | Equation |
|---|---|
| SPI-1 MX: Schedule Pressure/ Logistical Delay | $Logistical\ delay\ time\ (minutes)$ |
| SPI-2 MX: Schedule Pressure/Personnel | $\frac{Aviation\ maintenance\ technicians\ aviailable}{Fleet\ flight\ time}$ |
| SPI-3 MX: Schedule Pressure/Aircraft | $\frac{Percentage\ of\ aircraft\ avaialble}{Total\ aircraft\ in\ fleet}$ |
| SPI-4 MX: Schedule Pressure/ Flow | $\frac{Number\ of\ total\ maintenance\ orders\ processed}{Fleet\ flight\ time}$ |
| SPI-5 MX: Unscheduled Events | $\frac{Unscheduled\ maintenance\ orders\ under\ \$10k+FAA\ occurrences\ reports}{Fleet\ flight\ time}$ |
| SPI-6 MX: Errors | $\frac{Number\ of\ aircraft\ dispatched\ with\ maintenance\ error}{Number\ of\ total\ work\ orders\ processed}$ |
| SPI-1 FLT: Occurrences and Close Calls | $\frac{(reported\ tail\ strikes)+(number\ of\ hard\ landings)+(number\ of\ unstabilized\ approaches)+(Number\ of\ RPM\ overspeeds)+(number\ of over/under g)+(number\ of\ flap\ overspeeds)}{Fleet\ flight\ time}$ |
| SPI-2 FLT: Safety Culture | $\frac{(0.039*PI3)+(0.064*SO3)+(0.079*EFLS3)+(0.085*EFLS8)+(0.092*PS1)+(0.081*PS3)+(0.067*PS7)+(0.043*PRS1)+(0.07*EC2)+(0.072*RS2)+(0.043*QNH4)+(0.032*QNH5)+(0.018*MO1)}{Number\ of\ surveys\ collected}$ |
| SPI-3 FLT: NMACs | $\frac{Number\ of\ traffic\ conflicts}{Fleet\ flight\ time}$ |

| SPI | Equation |
|---|---|
| SPI-4 FLT: Staffing | $\frac{Number\ of\ full\ time\ equivalent\ instructor\ pilots}{Active\ flight\ students}$ |
| SPI-5 FLT: Turnover | $Number\ of\ months\ as\ instructor\ pilot\ at\ institution$ |
| SPI-6 FLT: Safety Reporting | $Number\ of\ events\ reported\ (ASAP\ and\ internal)$ |
| Damage and Related Impact | $\frac{\left(\begin{matrix}(NTSB\ accident\ reports * Impact\ Value) +\ FAA\ incident\ reports + \\ Unscheduled\ MX\ reports > 10K\end{matrix}\right)}{Fleet\ flight\ time}$ |
| Damage and Related Impact Score | $5 * \frac{(Actual\ value - 0)}{(0.0025 - 0)}$ |
| Maintenance Score | $\sum \begin{bmatrix}(SPI\ 1\ MX * 0.10) + (SPI\ 2\ MX * 0.15) + (SPI\ 3\ MX * 0.10) + \\ (SPI\ 4\ MX * 0.10) + (SPI\ 5\ MX * 0.10) + (SPI\ 6\ MX * 0.25)\end{bmatrix}$ |
| Flight Score | $\sum \begin{bmatrix}(SPI\ 1\ FLT * 0.25) + (SPI\ 2\ FLT * 0.125) + (SPI\ 3\ FLT * 0.25) \\ +(SPI\ 4\ FLT * 0.125) + (SPI\ 5\ FLT * 0.125) + (SPI\ 6\ FLT * 0.125)\end{bmatrix}$ |
| Overall Risk Score | $\sum \begin{bmatrix}(Maintenance\ Score * 0.3) + (Flight\ Score * 0.3) \\ +(Damage\ and\ Related\ Impact\ Score * 0.4)\end{bmatrix}$ |

*Step 4- Define simulation scenarios.* Next, the simulation scenarios were defined. For the purpose of the research, the scenarios were based upon manipulation of the four controllable input variables: the number of aviation maintenance technicians (AMTs) available, the total number of aircraft in the operational fleet, the number of active flight students, and the number of full-time equivalent instructor pilots (IPs).

The selection of scenarios for the study was designed to reflect typical operating conditions within a large, collegiate CFR Part 141 training operation and included manipulation of the controllable input variables to simulate changes to typical operating conditions to determine how these changes impacted the level of risk within the system as a whole. The study conducted simulations runs with the following specifications to provide output data for a large, CFR Part 141 operation with varying levels of resources.

Table 6

*Ranges of Controllable Input Variables for Simulation Runs*

| Controllable Input | Range |
|---|---|
| AMTs available | 14-35 |
| Aircraft available | 50-82 |
| Full-time instructor pilots (Ips) | 100-200 |
| Active flight students | 335-1300 |

These ranges were selected because they are reflective of the higher and lower operational limits of the sample data drawn from a large, collegiate CFR Part 141 flight training operation in the southeastern United States. By conducting simulation runs that model a range of available resources with regard to personnel, students, and aircraft, decision-makers could then determine the optimal level of resources necessary to meet operational demands while staying above a predetermined level of acceptable risk, thereby maintaining safety. Data collected from the scenarios, defined by using different specifications for the controllable input variables, were compared for sensitivity effects and were organized in a graphical output depicting the relationship between the controllable inputs and resulting risk score outputs.

*Step 5- Select control values.* To demonstrate the utility of the safety performance decision-making tool for real-world use, the controllable input values used to generate the what-if scenarios within the Monte Carlo simulation model were determined based on permutational variations drawn from the ranges of normal operating conditions specific to CFR Part 141 flight training organizations, depicted in Table 6. The selection of specific scenarios for the study focused on manipulating the controllable input variables: the number of aviation maintenance technicians available, the total number of aircraft in the fleet, the number of full-time instructor pilots, and the number of active flight students. The output of the model included probability curves depicting how changes to the controllable input variables impacted the flight score, maintenance score, damage and related impact score, and overall risk score output. The controllable input values for the four What-if Scenarios can be found in Table 17 of Chapter 5 of this dissertation.

*Step 6- Run the simulation.* Once the conceptual model was created in Analytica and the variable parameters and distributions were defined, the software ran the simulation model with 10,000 trials. The model utilizes Analytica® by Lumina Decision Systems as the software to complete the simulation. The Analytica software defines the mathematical model using a flowchart-type graphical representation and defines distributions for use as input data while providing the processing environment for repeated trails. The software also collects and organizes output data from each simulation trial to statistically analyze, examine, and compare scenario results. The simulation model predicts safety by rendering outputs based on four primary, controllable input variables: aviation maintenance technicians available, total aircraft in fleet, number of active flight students, and the number of full-time equivalent instructor pilots. These controllable input variables were manipulated to assess how changes to the controllable inputs impacted the outputs, or risk scores.

*Step 7- Analyze the results.* The results of the scenario were then analyzed to determine how the defined changes to the input variables impacted the outputs: the flight score, the maintenance score, damage and related impact score, and the overall risk score. Analysis of the results, including model validation, was conducted by performing a descriptive statistical analysis of the output tables for all trials within a particular scenario. The output tables contained the calculated probability of data from the various scenarios. The output tables were then organized in graphical format depicting the relationships between the controllable variables and the subsequent changes to the flight score, maintenance score, and overall risk score outputs. A sensitivity analysis was utilized to test edge cases, based on data from the various scenarios, to determine if the model could be modified to increase the overall sensitivity of the risk score outputs.

*Step 8- Next scenario/stop.* From this point, the criteria for the next simulated scenario could be defined and run until a sufficient number of scenarios have been completed. The next step required a decision to be made between returning to Step 4- Define the simulation scenarios, or returning to Step 1 and repeating the steps required to run another scenario. For the purpose of this dissertation, upon completion of the fourth What-if Scenario, the decision was made to stop adding scenarios.

**Apparatus and materials.** The software utilized for the Monte Carlo simulation was Analytica Educational Professional release 4.6.1.30 by Lumina Decisions Systems. This software allows researchers to model the uncertainty and variability of the input variables within the model. With Analytica, the researcher can graphically design the model simulation, as depicted by the screenshot of the simulation model from Analytica in Figure 5. Microsoft Excel 2013 was used to process the data and to analyze and illustrate characteristics of the intermediate input data, or SPIs, generated by the algorithms in the Analytica model. Microsoft Excel 2013 was also used for post-hoc testing and analysis.

**Sources of the data.** The sample of data used to determine the probability distributions for the uncontrollable input variables was drawn from a two-year sample of operational flight and maintenance data ranging from September 2017 to September 2019 from a large, collegiate CFR Part 141 flight training organization in the southeastern United States. The time period of September 2017 to September 2019 was selected to accurately capture probability distributions that are representative of the most current operating conditions, following the academic calendar, for a large, collegiate CFR Part 141 flight training organization. Utilizing probability distributions that are representative of the most current operating conditions enhances the validity of the model. The sample of data was analyzed in MiniTab Statistical Software to obtain Goodness of Fit tests to determine the probability distributions of the data sample to use within the Monte Carlo simulation. The study did not involve any human subjects or experimentation.

## Ethical Considerations

Using simulations to support executive decision-making introduces various types of ethical concerns related to the reliability and validity of the model. According to Barlow (2009), models are often deliberately built and used to form the basis for various forms of analysis using simulation techniques, the results of which are used to support organizational decision-making; the consequences associated with supporting executive decision-making via modeling techniques is the potential impact on innocent third parties. The "utility of a (simulation) study depends on the quality of the model and the skill of the modeler" (Barlow, 2009, p. 433). This testifies to the fundamental limitation of the modeling and simulation process − the development of a model and simulation provides no guarantee of a valid or successful outcome. Therefore, there is an ethical obligation to ensure the reliability and validity of both the SPIs driving the model, as well as the safety decision-making tool itself, before application and implementation within a CFR Part 141 flight training organization.

## Data Analysis Approach

**Reliability assessment method.** Various trials of the model were ran using different random number generator seed values to confirm the output of the simulation produced consistent results across trials. The distributions of the output variables were compared with descriptive statistics from simulation to simulation to demonstrate consistency. ANOVA testing was used to test for differences across sets of results (Hoyt, 1941).

To assess the model's reliability, the outputs were compared. Arbitrarily selected random number generator seed values were chosen to guarantee a different sequence of random numbers is produced for each trial. The seed value establishes the starting position in Analytica's random number generation function. This tests the model to determine if the results produced were consistent. ANOVA testing was conducted to determine if significant differences existed between the outputs of the reliability tests.

**Validity assessment method.** Typically, model validation occurs by utilizing two separate activity threads where one thread is used to ensure the mathematical calculations produced the expected results, and the other thread is used to compare the probability outputs of the model to similar models. In this case, the challenge with establishing a formal comparison of results between this model and the models developed in other studies is that no other studies directly address the same research questions. Rather, little work has been done in the realm of predictive modeling for large, collegiate CFR Part 141 flight training organizations. Researchers from the Brazilian low-cost carrier and Southwest Airlines are currently developing a similar model for assessing risk in CFR Part 121 operations, but their models are reactive in nature rather than predictive. Both the Brazilian low-cost carrier and Southwest Airlines have yet to publish their findings. Therefore, model validation occurred via the use of Subject Matter Experts using a standardized expert elicitation questionnaire distributed in a survey format (Anderson et al., 2020). Expert elicitation is the process of acquiring probabilistic belief statements from experts in a particular domain to assist in the process of quantifying uncertainty (Colson & Cooke, 2018). Inter-rater reliability was determined by calculating Kappa values. Since more than two experts were utilized, the use of Fleiss' Kappa was most appropriate (Stemler & Tsai, 2008).

The mathematical formulae used within the SPIs were derived from the formulae developed by Anderson et al. (2020). The formulae developed by Anderson et al. (2020) were established and validated via the expert elicitation process based on feedback from SMEs. To ensure no error occurred during the process of inputting the mathematical computations into the model, Verification Scenarios 1, 2, and 3 were conducted to ensure the random number generators produced a set of data values that is representative of the raw data sample. Each node within the model was assessed and manually verified to ensure the expected results.

**Data analysis process.** Both Microsoft Excel and Minitab 19 Statistical software were used for basic statistical analysis. The model produced a set of probability curves demonstrating the operating conditions within a large, collegiate CFR Part 141 flight training organization given different values of the controllable inputs. The study ran the simulation with 10,000 trials for a given scenario with manipulated controllable input values. Analytica rendered the results of each scenario in graphical and statistical formats capturing the output from each scenario in separate result matrices. The mean, standard deviation, maximum, and minimum values were used to determine the impact on either the flight or maintenance score and the overall risk score. ANOVA testing was also used to test for differences across sets of results (Hoyt, 1941). A Generalized Sensitivity Analysis (GSA) (Spear & Hornberger, 1980) was conducted to analyze the results of the What-if Scenarios. GSA is a technique that considers the sensitivity of model outputs to model inputs by separating the input parameter values into two distributions: those that created results that exceeded a specific threshold (“failed”) and those that created results that were below the threshold (“pass”). Separating the model output into two sample sets allows for the evaluation of the two sample sets as a function of any predetermined input parameter selected to represent a threshold of safe operation. GSA can also detect the presence of high output values for specific ranges of input parameters better than the other methods (Makino, McKenna, & Wakasugi, 2001). Conducting a GSA on the results of the what-if scenarios will allow for an enhanced, in-depth assessment of the resulting uncertainty within the model with respect to the effects of input parameter uncertainty.

**Summary**

The research builds a safety decision-making tool to evaluate what-if scenarios to evaluate how the changes to controllable inputs affect the SPIs determined by Anderson et al. (2020) to represent flight risk within large, collegiate CFR Part 141 flight training organizations. This research combined with former researcher efforts (Anderson et al., 2020) has provided the basis and expanded architecture used to build this model. Utilizing a quantitative methodology, the goal of the study was to expand the non-statistical model developed by Anderson et al. (2020) using Monte Carlo simulation to develop a safety decision-making tool to run what-if scenarios to assess how variations to the controllable input variables impacted the level of operational risk within an organization's flight department. Monte Carlo analysis was applied to enable the model to defensibly handle uncertainty in several key input variables while enabling the model to describe the range of possible outcomes given a set of controllable inputs to the model.

The target population to which the model generalizes is large, collegiate CFR Part 141 flight training organizations within the United States operating under the specifications defined by the FAA under Title 14 of the Code of Federal Regulations Part 141 (Federal Aviation Administration, 2017). The sampling frame consisted of the operational data from both flight and maintenance operations for a large, collegiate CFR Part 141 flight training organization in the southeastern United States. The sample data used to determine the probability distributions of the uncontrollable input variables within the model was based on a two-year sample of operational data from SPIs developed by Anderson et al. (2020) for a large, collegiate CFR Part 141 flight training organization in the southeastern United States.

The study conducted simulation runs based on specified ranges to simulate the range of operating conditions possible within large, CFR Part 141 operations with varying levels of resources concerning personnel (AMTs and IPs), students, and aircraft. These ranges were chosen because they are representative of flight training operations within large, collegiate CFR Part 141 flight training operations (Anderson et al., 2020). The selection of scenarios for the study was based on permutational variations of typical operating conditions within a large, collegiate CFR Part 141 training operation and included manipulation of the controllable input variables to simulate various operating conditions. The controllable input variables were manipulated to effectively simulate the operating conditions of a large, collegiate CFR Part 141 flight training organization. These ranges were chosen because they were found to be representative of real-world flight training operations within large, collegiate CFR Part 141 flight training organizations.

The software utilized for the Monte Carlo simulation was Analytica Educational Professional release 4.6.1.30 by Lumina Decisions Systems. Microsoft Excel 2013 was used to process the data and analyze and illustrate characteristics of the intermediate input data, or SPIs, generated by the algorithms in the Analytica model. Microsoft Excel 2013 was used for post-hoc testing and analysis.

Various trials of the model were ran using different random number generator seed values to confirm the output of the simulation produced consistent results across trials. The distributions of the output variables were compared with descriptive statistics from simulation to simulation to demonstrate consistency. ANOVA testing was used to test for differences across sets of results (Hoyt, 1941). To assess the model's reliability, the outputs were compared. Arbitrarily selected random number generator seed values were chosen to guarantee a different sequence of random numbers is produced for each trial.

Concerning model validation, the challenge with establishing a formal comparison of results between this model and the models developed in other studies is that no other studies directly address the same research questions. Rather, little work has been done in the realm of predictive modeling for large, collegiate CFR Part 141 flight training organizations. This research utilizes the validated equations drawn from the non-statistical model developed by Anderson et al. (2020) for the mathematical inputs driving the computational nodes, including the SPIs, the Flight Score, Maintenance Score, Damage and Related Impact Score, and the Overall Risk Score, as the foundation to develop the safety performance decision-making tool.

The model produced a set of probability curves demonstrating the operating conditions within a large, collegiate CFR Part 141 flight training organization given different values of the controllable inputs. The study ran the simulation with 10,000 trials for a given scenario with manipulated controllable input values to identify the sensitivity of the results to specific probabilistic inputs within the model. Analytica rendered the results of each scenario in multiple graphical forms capturing the output from each scenario in separate result matrices. The mean, standard deviation, maximum, and minimum values were used to determine the impact on either the flight or maintenance score and the overall risk score. ANOVA testing was also used to test for differences across sets of results (Hoyt, 1941). A Generalized Sensitivity Analysis (GSA) was conducted to evaluate the results of the what-if scenarios and determine if the sensitivity of the model could be improved (Spear & Hornberger, 1980).

# CHAPTER IV

# RESULTS

Chapter 3 described the steps necessary to transform a nonstatistical model composed of domain-specific SPIs into a safety performance decision-making tool, using Monte Carlo simulation, to run what-if scenarios to assess how variations to the controllable input variables impact the level of operational risk within a large, collegiate CFR Part 141 flight training organization. Determining the probability distributions of the uncontrollable input variables from the sample data allowed for the nonstatistical model to be transformed into a predictive, safety performance decision-making tool.

This chapter describes the results in four general sections. The first three sections answer the first research question – how can the SPI model developed by Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020) be transformed into a predictive, safety performance decision-making tool with the ability to run what-if scenarios? Section one details the output of the verification testing process. Section two includes data describing the reliability test results of the model. Section three depicts the validity test results of the model. The fourth section demonstrates the utility of the model both statistically and graphically in response to research question two – how do changes to the controllable input variables impact the Overall Risk Score?

## Demographic Information

The sample data used to determine the probability distributions of the uncontrollable input variables for the Monte Carlo simulation was comprised of two years of operational flight and maintenance data from September 2017 to September 2019 from a large, collegiate 14 CFR Part 141 flight training organization in the southeastern United States. The sample is 14 CFR Part 141 flight training organizations. Descriptive statistics of the raw data sample can be found in Table 7. The demographic distribution of the sample 14 CFR Part 141 flight training organization can be found in Figure 6. Operating at a capacity of approximately 7,000 flight hours per month, the demographic results of the sample fall within the normal range of operating conditions determined to be representative of 14 CFR Part 141 flight training organizations by SMEs in the area of CFR Part 141 flight training organizations.

Table 7

*Descriptive Statistics of the Raw Data Sample*

| SPI | Variable | Lower Limit | Higher Limit | *Mean* | *SD* |
|---|---|---|---|---|---|
| 1-MX | Logistical delay time | 100 | 310 | 203.8579 | 46.7893 |
| 2-MX | *AMTs Available** | 14 | 35 | 21 | 3.5033 |
| | Fleet flight time | 4000 | 13500 | 7365.717 | 1674.774 |
| 3-MX | Percent of AC available | 70 | 100 | 83.8003 | 4.6361 |
| | *Total AC available** | 50 | 82 | 62.236 | 6.2056 |
| 4-MX | Fleet flight time | 4000 | 13500 | 7365.717 | 1674.774 |
| | Total MX orders processed | 100 | 1200 | 514.9677 | 118.706 |
| 5-MX | Unscheduled MX orders <$10K | 300 | 1000 | 468.1397 | 132.7093 |
| | FAA occurrences | 0 | 40 | 6.32 | 4.7847 |
| | Fleet flight time | 4000 | 13500 | 7365.717 | 1674.774 |
| 6-MX | Total MX orders processed | 100 | 1200 | 514.9677 | 118.706 |
| | AC dispatched w/ MX error | 0 | 2 | 0.12 | 0.3317 |
| 1-FLT | Unstable approaches | 0 | 946 | 78.0129 | 229.9836 |
| | Flap overspeeds | 0 | 3 | 0.56 | 0.7118 |
| | G exceedances | 0 | 3 | 0.44 | 1.0033 |
| | Tail strikes | 0 | 10 | 1.64 | 1.9339 |
| | RPM overspeeds | 0 | 3 | 0 | 0 |
| | Hard landings | 0 | 7 | 1.2 | 1.6583 |
| | Fleet flight time | 4000 | 13500 | 7365.717 | 1674.774 |

| SPI | Variable | Lower Limit | Higher Limit | *Mean* | *SD* |
|---|---|---|---|---|---|
| 2-FLT | Annual SC survey results | 1 | 5.76 | 4.6 | 0.0181 |
| 3-FLT | Traffic conflicts | 0 | 18 | 8.04 | 3.0752 |
| | Fleet flight time | 4000 | 13500 | 7365.717 | 1674.774 |
| 4-FLT | *Full-time Ips** | *100* | *200* | 138 | 8.8600 |
| | *Active flight students** | *335* | *1300* | 656 | 179.8793 |
| 5-FLT | Months as an IP | 0 | 12 | 10 | 0 |
| 6-FLT | Event reports | 25 | 150 | 67.3372 | 20.5756 |
| Damage & Related Impact | FAA incident reports | 0 | 3 | 0.2 | 0.4082 |
| | Unsched MX > $10K | 0 | 3 | 0.96 | 1.5133 |
| | NTSB reports | 0 | 3 | 0.16 | 0.3742 |
| | Fleet flight time | 4000 | 13500 | 7365.717 | 1674.774 |

**Controllable input variable*

**Model Verification Testing**

The simulation used Analytica 64-bit Educational Professional software Release 4.6.1.30 by Lumina Decision Systems. To ensure the model’s algorithms were accurately entered in the simulation software, the content of each node of the model depicted in Figure 5 was verified for consistency with the model equations depicted in Table 5.

Input nodes, comprised of probability distribution data, were statistically and graphically examined to substantiate the output conformed to each input’s specific distribution profile, as determined by a two-year sample of raw operational flight and maintenance data ranging from September 2017 to September 2019 from a large, collegiate CFR Part 141 flight training organization in the southeastern United States. Computational nodes, depicted by light blue rounded rectangles in Figure 5, were verified by comparing the node’s simulated output to the results of manual calculations drawn from the sample data. There are 22 uncontrollable inputs that were supplied as random numbers within the bounds of their specified probability distributions. These inputs can be found in Table 3. For model verification purposes, the output of each of these distributions is examined below from a simulation run with 10,000 trials.

Three Verification Scenarios were conducted. Within Verification Scenario 1, the values selected to serve as controllable input variable values in Table 8 were determined by calculating the mean value for each variable of the sample data. The purpose of using the mean value of each variable from the sample data was to ensure the output of the model was representative of the CFR Part 141 flight training organization’s true operating conditions determined by the raw data sample.

The values for the controllable input variables in Verification Scenario 2 were drawn from the low values of the operational ranges for CFR Part 141 flight training organizations depicted in Table 6, whereas the controllable input variables for Verification Scenario 3 were drawn from high operational range values. High and low range values were selected to represent the varying operational capacities of the target population. By conducting simulation trials that model a range of available resources concerning personnel, students, and aircraft, decision-makers could then determine the optimal level of resources necessary to meet operational demands while staying above a predetermined level of acceptable risk, thereby maintaining safety.

Table 8

*Verification Scenario 1 Controllable Input Values*

| Controllable Input | Value |
|---|---|
| AMTs available | 22 |
| Aircraft available | 56 |
| Full-time instructor pilots (Ips) | 138 |
| Active flight students | 681 |

*Note*. Source: Raw data means. Sample: 10,000; Random seed: 99

Table 9 depicts the output values and the shape of the distribution for each uncontrollable input variable in Verification Scenario 1 extracted from the outputs of the model. The shape of the distributions of the uncontrollable input variables from Verification Scenario 1 is the same as the distributions drawn from the raw data sample. The higher and lower limits of the raw data sample were included for comparison purposes (Anderson et al., 2020).

For each SPI, the higher limit was calculated by analyzing the two-year sample of data for a specific SPI, finding the operational month with the highest data point value and dividing the highest value by the operational month with the lowest data point value. A lower limit was determined by reversing the equation, and dividing the lowest value over a two-year span of sample data by the highest value. As determined by the model output for Verification Scenario 1, the mean values for all 22 uncontrollable inputs fell within the boundaries of the lower and higher limits of the raw data.

Table 9

*Verification Scenario 1 Comparison Input*

| | | Input Variable Distributions | | | | Raw Data Sample | |
|---|---|---|---|---|---|---|---|
| SPI | Variable | Min Value | Max Value | Mean | Distribution Shape | Lower Limit | Higher Limit |
| 1-MX | Logistical delay time | 102 | 297 | 212 | Weibull | 100 | 310 |
| 2-MX | *AMTs Available** | 22 | 22 | | | 14 | 35 |
| | Fleet flight time | 4006 | 13300 | 7602 | Normal | 4000 | 13500 |
| 3-MX | Percent of AC available | 70 | 100 | 85 | Logistic | 70 | 100 |
| | *Total AC available** | 56 | 56 | | | 50 | 82 |
| 4-MX | Fleet flight time | 4006 | 13300 | 7602 | Normal | 4000 | 13500 |
| | Total MX orders processed | 100 | 800 | 532 | Logistic | 100 | 1200 |
| 5-MX | Unscheduled MX orders <$10K | 415 | 500 | 577 | Binomial | 300 | 1000 |
| | FAA occurrences | 1 | 49 | 6 | Geometric | 0 | 40 |
| | Fleet flight time | 4006 | 13300 | 7602 | Normal | 4000 | 13500 |
| 6-MX | Total MX orders processed | 107 | 1036 | 535 | Logistic | 100 | 1200 |
| | AC dispatched w/ MX error | 0 | 1 | 0.05 | Bernoulli | 0 | 2 |
| 1-FLT | Unstable approaches | 6 | 767 | 156 | Lognormal | 0 | 946 |
| | Flap overspeeds | 0 | 3 | 0.55 | Poisson | 0 | 3 |
| | G exceedances | 0 | 3 | 0.42 | Poisson | 0 | 3 |
| | Tail strikes | 0 | 7 | 2.72 | Poisson | 0 | 10 |
| | RPM overspeeds | 0 | 1 | 0.5 | Poisson | 0 | 3 |
| | Hard landings | 0 | 5 | 1.7 | Poisson | 0 | 7 |
| | Fleet flight time | 4006 | 13300 | 7602 | Normal | 4000 | 13500 |
| 2-FLT | Annual SC survey results | 4.6 | 4.6 | 4.6 | Certain | 1 | 5.76 |
| 3-FLT | Traffic conflicts | 1 | 25 | 10 | Binomial | 0 | 18 |
| | Fleet flight time | 4006 | 13300 | 7602 | Normal | 4000 | 13500 |
| 4-FLT | *Full-time Ips** | *138* | *138* | | | *100* | *200* |
| | *Active flight students** | *681* | *681* | | | *335* | *1300* |
| 5-FLT | Months as an IP | 10 | 10 | 10 | Certain | 0 | 12 |
| 6-FLT | Event reports | 39 | 108 | 67 | Negative Binomial | 25 | 150 |

| | | Input Variable Distributions | | | | Raw Data Sample | |
|---|---|---|---|---|---|---|---|
| SPI | Variable | Min Value | Max Value | SPI | Variable | Min Value | Max Value |
| Damage & Related Impact | FAA incident reports | 0 | 1 | 0.2 | Binomial | 0 | 3 |
| | Unsched MX > $10K | 0 | 3 | 0.91 | Poisson | 0 | 3 |
| | NTSB reports | 0 | 1 | 0.34 | Binomial | 0 | 3 |
| | Fleet flight time | 4006 | 13300 | 7602 | Normal | 4000 | 13500 |

**Controllable input variable*

Once the uncontrollable input variables were verified to be representative of the raw data based on the shape of the probability distribution outputs, the minimum, maximum, and mean values for each calculation node, or SPI, were compared with the lower and higher limits of the raw data, shown in Table 10. Close inspection indicated the model’s output, including the maximum and minimum values, were generally lower than the lower and higher limits of the raw data; however, the mean values of the SPIs all fall within the bounds of the raw data.

Table 10

*Verification Scenario 1: SPI Comparison Outputs*

| | SPI Distributions | | | Raw Data Sample | |
|---|---|---|---|---|---|
| SPI | Min | Max | Mean | Lower Limit | Higher Limit |
| SPI-1 MX | 1.712 | 4.874 | 3.532 | 1.6667 | 5.1667 |
| SPI-2 MX | 0.0016 | 0.0055 | 0.003 | 0 | 0.00875 |
| SPI-3 MX | 1.25 | 1.786 | 1.516 | 0.8537 | 2 |
| SPI-4 MX | 0.0121 | 0.1888 | 0.0731 | 0.0074 | 0.3 |
| SPI-5 MX | 0.0369 | 0.1393 | 0.0698 | 0 | 0.26 |
| SPI-6 MX | 0 | 0.005 | 0.0001 | 0 | 0.02 |
| | | | | | |
| SPI-1 FLT | 0.001 | 0.1799 | 0.0224 | 0 | 0.0302 |
| SPI-2 FLT | 4.6 | 4.6 | 4.6 | 1 | 5.76 |
| SPI-3 FLT | 0.0002 | 0.0034 | 0.0013 | 0 | 0.0045 |
| SPI-4 FLT | 4.9348 | 4.9348 | 4.9348 | 2 | 8 |
| SPI-5 FLT | 10 | 10 | 10 | 0 | 36 |
| SPI-6 FLT | 35 | 103 | 67 | 0 | 200 |

The next step in the verification process included examining each standardized SPI as well as the Damage and Related Impact variable, depicted in Figure 5 as a blue rounded rectangular computational node. To accurately feed into a standardized risk score output ranging from 0-5, the model fed each SPI computational node into an individual standardized SPI node. The output for each standardized SPI computational node, as well as the standardized Damage and Related Impact variable, were compared with the lower and higher limits of the raw data. This output can be found in Table 11. Results indicated that the mean values of each standardized SPI and the standardized Damage and Related Impact variable fell between the lower and higher limits of the raw data, further verifying the accuracy of the model.

Table 11

*Verification Scenario 1: Standardized SPI Comparison Outputs*

| | SPI Distributions | | | Raw Data Sample | | |
|---|---|---|---|---|---|---|
| SPI | Min | Max | Mean | Min | Max | Mean |
| SPI-1 MX | 0 | 5 | 2.878 | 0.8883 | 4.8501 | 2.5948 |
| SPI-2 MX | 2.112 | 4.609 | 3.717 | 2.1639 | 4.2139 | 3.285 |
| SPI-3 MX | 1.73 | 4.064 | 2.882 | 1.1229 | 3.4422 | 2.8129 |
| SPI-4 MX | 0 | 5 | 1.724 | 0.2209 | 1.3921 | 1.103 |
| SPI-5 MX | 0 | 5 | 1.601 | 0.3281 | 1.7218 | 1.2976 |
| SPI-6 MX | 0 | 5 | 0.0974 | 0 | 0.4562 | 0.0512 |
| SPI-1 FLT | 0 | 5 | 0.5917 | 0.0189 | 2.0034 | 0.5187 |
| SPI-2 FLT | 1.218 | 1.218 | 1.218 | 1.2185 | 1.2185 | 1.2185 |
| SPI-3 FLT | 0 | 5 | 1.702 | 0.3660 | 2.0590 | 1.1916 |
| SPI-4 FLT | 2.446 | 2.446 | 2.446 | 0.4610 | 3.9552 | 2.4291 |
| SPI-5 FLT | 3.611 | 3.611 | 3.611 | 3.6111 | 3.6111 | 3.6111 |
| SPI-6 FLT | 0 | 5 | 2.382 | 0.8000 | 4.0250 | 3.108 |
| Damage & Related Impact | 0 | 0.4197 | 0.084 | 0 | 0.6233 | 0.1220 |

Table 12

*Verification Scenario 1: Risk Score Output Comparisons*

| | Output Variable Distributions | | | Manual Calculation | |
|---|---|---|---|---|---|
| Risk Score Output | Min | Max | Mean | Min | Max |
| Maintenance Score | 1.007 | 2.805 | 1.49 | 0.9272 | 1.7378 |
| Flight Score | 1.121 | 3.466 | 1.781 | 1.3347 | 2.0705 |
| Damage & Related Impact Score | 0 | 0.4197 | 0.084 | 0 | 0.3349 |
| Overall Risk Score | 0.7336 | 1.609 | 1.015 | 0.7854 | 1.1698 |

*Note.* The mean model output values fall within the minimum and maximum ranges manually calculated based on the raw data, verifying the model's calculations produced viable output values.

Finally, Table 12 depicted the mean, minimum, and maximum values for the Risk Score probability density outputs for maintenance, flight, the Damage and Related Impact, and the Overall Risk Score for the operation as a whole. The risk score outputs were manually calculated using the raw data sample values for two years of operational flight and maintenance data from a CFR Part 141 flight training organization in the southeastern United States. For verification purposes, the maximum and minimum values were used. The mean output for all four of the controllable risk score outputs fell between the maximum and minimum values of the raw data sample. Thus, the output values calculated by Analytica, specifically the mean values, fell within the bounds of the manual calculations of the outputs, given the input values used for verification testing. The resulting outputs produced the following distribution of values shown below in Figures 6, 7, 8, and 9.

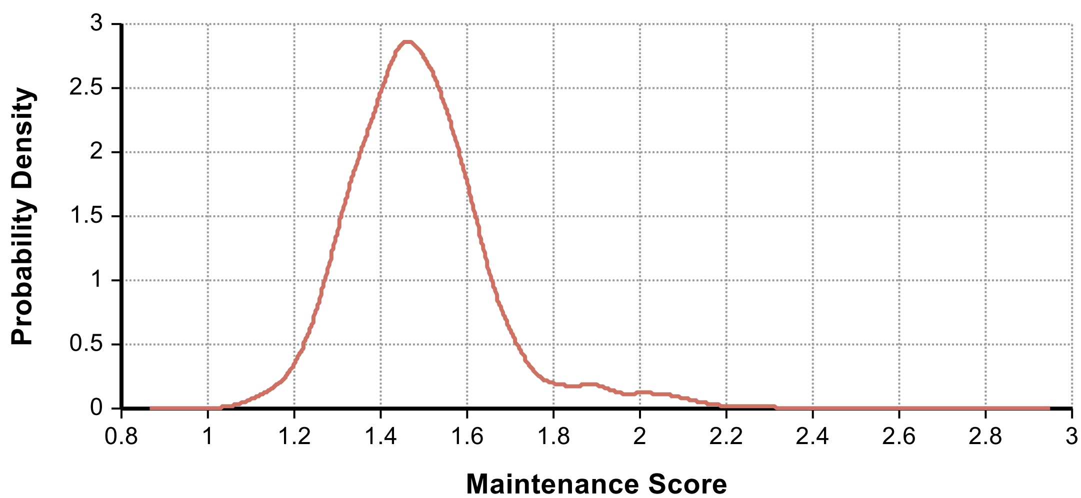



*Figure 6.* Probability density distribution of the Maintenance Score in Verification Scenario 1.

Figure 6 demonstrates the resulting probability density distribution output of the Maintenance Score in Verification Scenario 1. Results portrayed a mean risk score output of 1.49, indicating a safe level of maintenance operation under the specifications for the controllable input variables. The shape of the distribution visually indicates the vast range of potential output scores resulting from running the simulation through 10,000 trials.

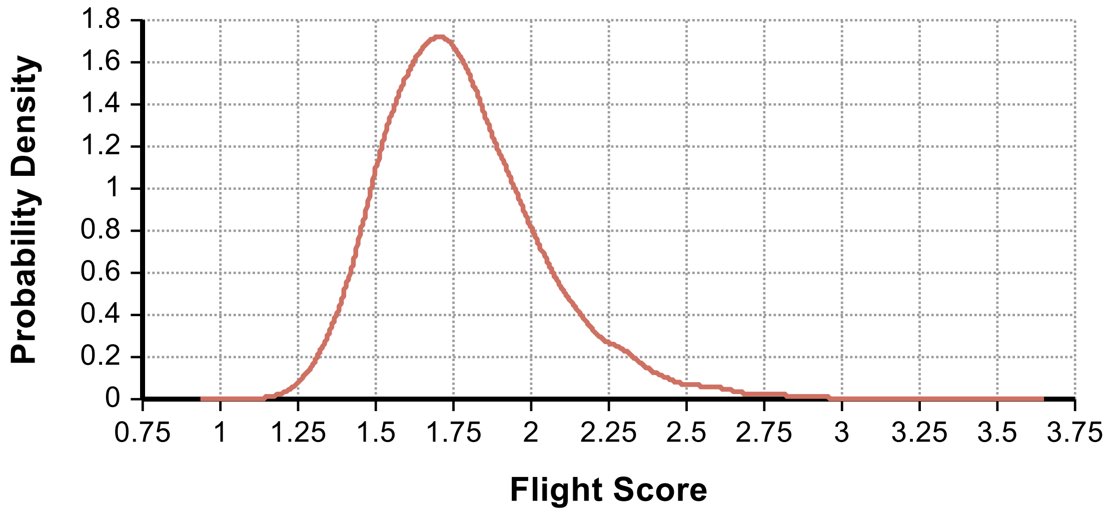



*Figure 7.* Probability density distribution of the Flight Score in Verification Scenario 1.

Figure 7 demonstrates the probability density distribution output of the Flight Score in Verification Scenario 1. Results portrayed a mean risk score output of 1.781 indicating a safe level of flight operation under the specifications for the controllable input variables with the outputs centered close to the mean; however, when compared to the mean risk score output of the Maintenance Score, the Flight Score output is slightly riskier.

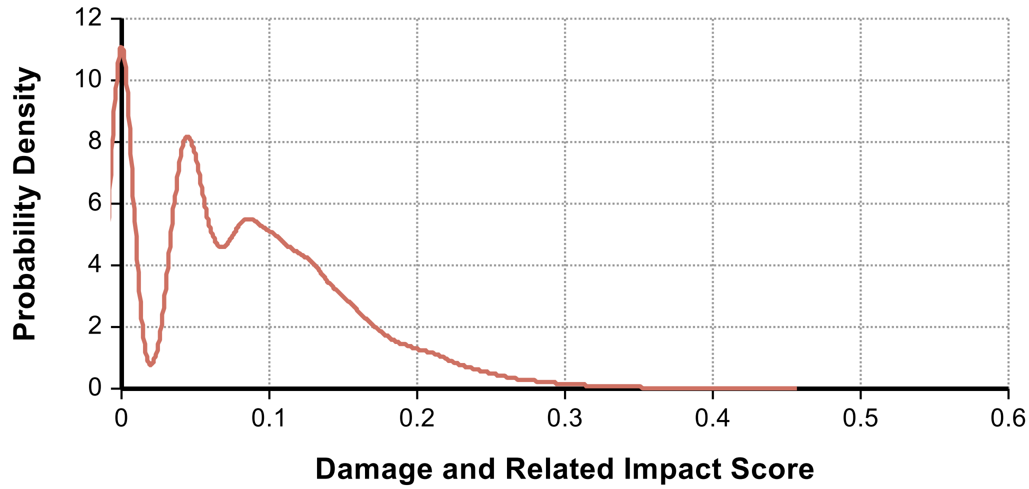



*Figure 8.* Probability density distribution of the Damage & Related Impact Score in Verification Scenario 1.

Figure 8 reveals the probability density distribution output of the Damage and Related Impact Score in Verification Scenario 1. The mean risk score output for the Damage & Related Impact Score was 0.084 indicating a safe operation. The erratic shape of the distribution is due to the infrequency of NTSB reports, FAA incident reports, unscheduled maintenance events greater than $10,000, and a static Impact Value of 1 indicating no accidents or incidents. However, these values were assigned high weights due to their importance within the system. The combination of infrequent occurrence and high weighted values produced the erratic distribution of the Damage & Related Impact Score.

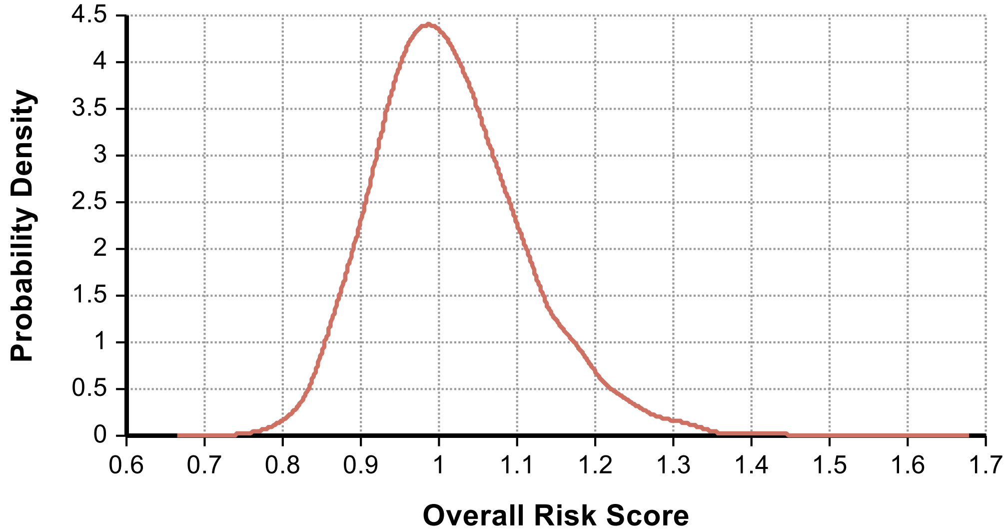



*Figure 9.* Probability density distribution of the Overall Risk Score in Verification Scenario 1.

Figure 9 shows the probability density distribution output of the Overall Risk Score in Verification Scenario 1. Results portrayed a mean overall risk score output of 1.015, indicating a safe level of the overall operation under the specifications for the controllable input variables.

To ensure no programming error occurred, two additional Verification Scenarios were performed using different controllable input variables. Verification Scenario 2 was conducted using the range lows as values for the controllable input variables. Verification Scenario 3 utilized the range highs as values for the controllable input variables. The controllable input variables used in Verification Scenarios 2 and 3 can be found in Table 13.

Table 13

*Verification Scenarios 2 and 3 Controllable Input Values*

| Controllable Input | Verification Scenario 2 Value | Verification Scenario 3 Value |
|---|---|---|
| AMTs available | 14 | 35 |
| Aircraft available | 50 | 82 |
| Active flight students | 335 | 1300 |
| Full-time instructor pilots (Ips) | 100 | 200 |

*Note*. Source: Operational range highs and lows; Sample: 10,000; Random seed: 99

The results, depicted as risk score outputs for the Maintenance Score, Flight Score, Damage and Related Impact Score, and the Overall Risk Score, is depicted in Tables 14 and 15. Overall, the model produced the results expected based on the controllable input variable specifications. The individual outputs for each uncontrollable input variable, SPI comparison outputs, and standardized SPI outputs from Verification Scenarios 2 and 3 can be found in Appendix B (Tables B1-6) and C (Figures C1-8).

Table 14

*Verification Scenario 2: Risk Score Output Comparisons*

| | Model | | | Manual Calculation | |
|---|---|---|---|---|---|
| Risk Score Output | Min | Max | Mean | Min | Max |
| Maintenance Score | 1.168 | 3.006 | 1.667 | 0.9272 | 1.7378 |
| Flight Score | 0.9561 | 3.301 | 1.616 | 1.3347 | 2.0705 |
| Damage & Related Impact Score | 0 | 0.4197 | 0.084 | 0 | 0.3349 |
| Overall Risk Score | 0.7321 | 1.628 | 1.021 | 0.7854 | 1.1698 |

*Note.* The mean model output values fall within the minimum and maximum ranges manually calculated based on the raw data, verifying the model's calculations produced viable output values.

Table 15

*Verification Scenario 3: Risk Score Output Comparisons*

| | Model | | | Manual Calculation | |
|---|---|---|---|---|---|
| Risk Score Output | Min | Max | Mean | Min | Max |
| Maintenance Score | 0.66 | 2.387 | 1.106 | 0.9272 | 1.7378 |
| Flight Score | 1.284 | 3.629 | 1.944 | 1.3347 | 2.0705 |
| Damage & Related Impact Score | 0 | 0.4197 | 0.084 | 0 | 0.3349 |
| Overall Risk Score | 0.6828 | 1.517 | 0.9486 | 0.7854 | 1.1698 |

*Note.* The mean model output values fall within the minimum and maximum ranges manually calculated based on the raw data, verifying the model's calculations produced viable output values.

**Reliability Testing**

Monte Carlo simulation modeling uses randomly selected numbers from predetermined probability distributions to produce data outputs in the form of probability distributions to account for the uncertainty inherent to the 22 uncontrollable input variables. Testing was conducted across multiple trials using various random number generator seed values to ensure the results remained consistent across trials. The mean probability output represents the forecasted level of operational risk on a standardized 0-5 risk scale for the Flight Score, Maintenance Score, Damage and Related Impact, and an Overall Risk Score representative of the operation as a whole over 10,000 trials of the simulation model.

The model was tested using various numbers of trial iterations ranging from 10 trials up to 30,000 trials. Although the results varied, the results were nearly identical after 10,000 trials for a given test. Ultimately, this study used 10,000 trials. To evaluate the reliability of the model, the study compared the results of three different iterative runs of the model—each using a unique seed value to ensure a different sample of random numbers for the uncontrollable input variables. The controllable input values for the three different runs of the model are the same as those used in Verification Scenario 1 (see Table 8). Analyzing the output with different seed values allows for the model to be verified for consistency in its results.

In each scenario, 10,000 trials were executed, and three arbitrarily selected random number generator seed values were selected to ensure the model produced a different set of random numbers across trials. The seed value determines the starting position in the random number generation; thus, different seed values cause the software to produce different samples of random numbers within the simulation. Using different samples of random numbers tests the model to see if it produces consistent results regardless of the starting point, or seed value.

To reflect the operating conditions of a large, collegiate CFR Part 141 flight training organization, the values chosen for the four controllable input variables were based on the mean values drawn from two years of operational flight and maintenance data. Again, the values for the four controllable input variables are as follows:

- Aviation maintenance technicians available: 22
- Total aircraft in fleet: 56
- Full-time instructor pilots: 138
- Active flight students: 681

Table 16 depicts the results of the reliability testing using different seed values. For each group of results, three different seed values generated three different samples of random numbers. Thus, the model ran 10,000 trials, producing 10,000 results for each of the three different samples of random numbers. Table 16 also shows the mean and standard deviation of the outputs for each of these runs. No significant differences appeared among the different sets of results indicating the results are statistically reliable. This study used ANOVA to test for differences across the three groups (Hoyt, 1941). The ANOVA F-statistic and P-value for each set of results can be found in Table 16.

Table 16

*Comparison of Results with Different Random Number Seed Values*

| Output | Seed Value | Mean | Standard Deviation | ANOVA *F* | ANOVA *P*-value |
|---|---|---|---|---|---|
| Maintenance Score | 99 | 1.49 | 0.1686 | 3.6446 | 0.3071 |
| | 50 | 1.491 | 0.1606 | | |
| | 10 | 1.492 | 0.1638 | | |
| Flight Score | 99 | 1.781 | 0.2627 | 81 | 0.0704 |
| | 50 | 1.784 | 0.2628 | | |
| | 10 | 1.792 | 0.2692 | | |
| Damage & Related Impact Score | 99 | 0.0835 | 0.0687 | 0.25 | 0.7048 |
| | 50 | 0.0829 | 0.0692 | | |
| | 10 | 0.0833 | 0.0680 | | |
| Overall Risk Score | 99 | 1.015 | 0.0978 | 36 | 0.1051 |
| | 50 | 1.016 | 0.0958 | | |
| | 10 | 1.018 | 0.0986 | | |

*Note*. No significant differences appear among the different sets of results; thus, the results are considered statistically reliable.

Assumptions for ANOVA were also tested. The large sample size of the simulated data meets the normality assumption. Levene's testing verified the satisfaction of the homogeneity assumption. A non-significant Levene's statistic test ($p > 0.05$) indicates the homogeneity of variance among the test groups. As shown in Table 16, the *p*-values for all cases are greater than 0.05, indicating there are no significant differences among the three samples; therefore, the results produced by the model are statistically reliable.

**Validity Testing**

The challenge with establishing a formal comparison of results between this model and the models developed in other studies is that no other studies directly address the same research questions. Additionally, little work has been done in the realm of predictive modeling specific to large, collegiate CFR Part 141 flight training organizations. This research utilizes the validated equations drawn from the non-statistical model developed by Anderson et al. (2020) for the mathematical inputs driving the computational nodes, including the SPIs, the Flight Score, Maintenance Score, Damage and Related Impact Score, and the Overall Risk Score, as the foundation to develop the safety performance decision-making tool.

The peer-reviewed research conducted by Anderson et al. (2020) validated the non-statistical model and associated equations via the use of Subject Matter Experts using a standardized expert elicitation survey questionnaire. Expert elicitation was used to establish inter-rater reliability for the assessment of SME evaluations. The Fleiss' kappa value was 0.0360, indicating a fair level of agreement among raters. Qualitative feedback was solicited and SMEs were asked to provide any comments or feedback on the model and equations driving the model to justify their rating scores. SMEs were in a high level of agreement relative to the overall utility of the model in providing a quantitative indicator of flight risk for large, collegiate CFR Part 141 flight training organizations. Thus, the equations driving the predictive, safety performance decision-making tool developed in this dissertation have been previously validated through the peer-reviewed research conducted by Anderson et al. (2020).

Additionally, three Verification Scenarios of the model were conducted using the validated equations determined by Anderson et al. (2020). Within Verification Scenario 1, the values selected to serve as controllable input variable values were determined by calculating the mean value for each variable of the sample data. The purpose of using the mean value of each variable from the sample data was to ensure the output of the model was representative of the CFR Part 141 flight training organization's true operating conditions determined by the raw data sample. Whereas the values for the controllable input variables in Verification Scenario 2 were drawn from the low values of the operational ranges for CFR Part 141 flight training organizations, the controllable input variables for Verification Scenario 3 were drawn from high operational range values. High and low range values were selected to represent the varying operational capacities of the target population. Demonstrating the capability of the model using a wide range of available resources further enhances the validity of the findings.

**Monte Carlo Simulation Results**

To demonstrate the utility of the safety performance decision-making tool for real-world use, the controllable input values used to generate the what-if scenarios within the Monte Carlo simulation model were determined based on permutational variations of ranges of normal operating conditions specific to CFR Part 141 flight training organizations. These ranges can be found in Table 6. These permutations were conducted by varying the level of personnel, including available aviation maintenance technicians and instructor pilots, as low, moderate, or high. Similarly, permutations of resource expenditures, including aircraft available and active flight students, were also varied by degree of low, moderate, or high. Low values consisted of the lowest possible range values, moderate values consisted of the median value, and high range values consisted of the highest value of the predetermined, true, operational ranges for a large, collegiate CFR Part 141 flight training organization.

The Analytica software tool computed each trial using the specified controllable input variables listed in Table 17, capturing the output from each trial in a separate results matrix for each trial. This allowed the model to compute the risk score outputs, depicted as probability results, for the controllable input values given for each simulation trial.

Table 17

*Controllable Inputs for What-if Scenarios 1, 2, 3, and 4*

| What-if Scenario | Controllable Input | Value | Description |
|---|---|---|---|
| Scenario 1 | AMTs | 14 | Low personnel, high expenditures |
| | Aircraft | 82 | |
| | IPs | 100 | |
| | Students | 1300 | |
| Scenario 2 | AMTs | 22 | Moderate personnel, high expenditures |
| | Aircraft | 82 | |
| | IPs | 138 | |
| | Students | 1300 | |
| Scenario 3 | AMTs | 35 | High personnel, low expenditures |
| | Aircraft | 50 | |
| | IPs | 200 | |
| | Students | 335 | |
| Scenario 4 | AMTs | 35 | High personnel, moderate expenditures |
| | Aircraft | 56 | |
| | IPs | 200 | |
| | Students | 681 | |

*Note*. AMTs = Aviation maintenance technicians; Aircraft = Aircraft available; IPs= Full-time instructor pilots; Students = Active flight students.

Table 18

*Results of What-if Scenarios 1, 2, 3 and 4*

| What-if Scenario | Output | Mean (*M*) | Standard Deviation (*SD*) |
|---|---|---|---|
| Scenario 1 | Maintenance Score | 1.39 | 0.1683 |
| | Flight Score | 2.621 | 0.2566 |
| | Damage & Related Impact Score | 0.0835 | 0.0687 |
| | Overall Risk Score | 1.237 | 0.0967 |
| Scenario 2 | Maintenance Score | 1.283 | 0.1578 |
| | Flight Score | 2.248 | 0.2566 |
| | Damage & Related Impact Score | 0.0835 | 0.0687 |
| | Overall Risk Score | 1.092 | 0.0951 |
| Scenario 3 | Maintenance Score | 1.396 | 0.1601 |
| | Flight Score | 1.441 | 0.2566 |
| | Damage & Related Impact Score | 0.0835 | 0.0687 |
| | Overall Risk Score | 0.8845 | 0.0955 |
| Scenario 4 | Maintenance Score | 1.317 | 0.1563 |
| | Flight Score | 1.621 | 0.2566 |
| | Damage & Related Impact Score | 0.0835 | 0.0687 |
| | Overall Risk Score | 0.9149 | 0.0949 |

What-if Scenario 1 was conducted with the intent of simulating a scenario where personnel, including AMTs and instructor pilots, are low, but the necessary expenditures, including aircraft and active flight students, are high. The probability density distribution output for What-if Scenario 1 can be found in Figures 10, 11, 12, and 13. Based on the specific controllable input variables used, results indicated What-if Scenario 1 had the highest mean value for the Overall Risk Score and the Flight Score, indicating a higher level of operational risk associated with conditions where a flight instructor capacity of

100 full-time instructors is not adequate to meet the demands of 1300 flight students, increasing the level of operational risk, specifically in the flight department. Although this is intuitive, it demonstrates the utility of the model for real-world use.

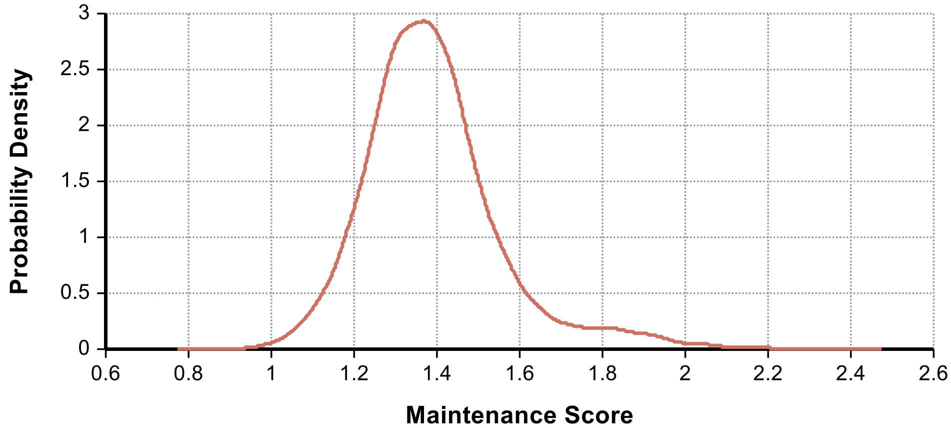



*Figure 10.* Probability density distribution of the Maintenance Score in What-if Scenario 1.

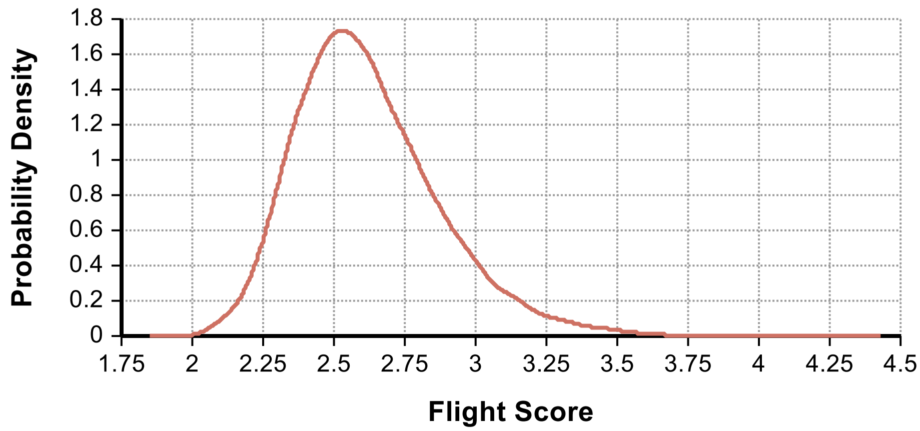



*Figure 11.* Probability density distribution of the Flight Score in What-if Scenario 1.

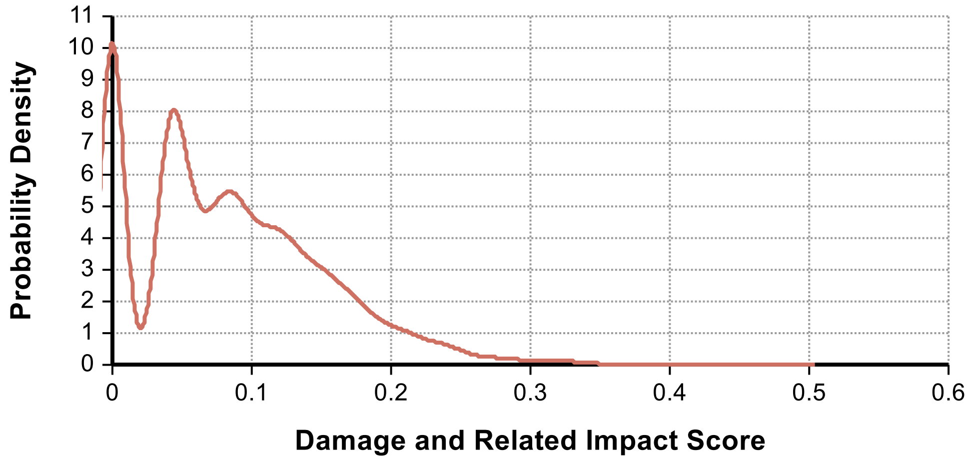



*Figure 12.* Probability density distribution of the Damage and Related Impact Score in What-if Scenario 1. Output scores between -1 and 0 are representative of occurrences in which there were no incidents to report.

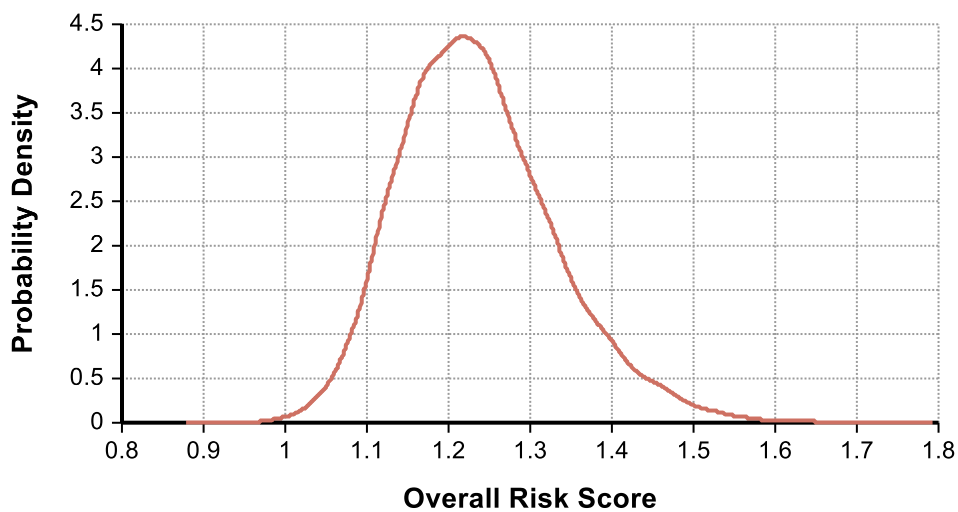



*Figure 13.* Probability density distribution of the Overall Risk Score in What-if Scenario 1.

What-if Scenario 2 was conducted with the intent of simulating a scenario similar to What-if Scenario 1; however, in What-if Scenario 2, the number of personnel, including AMTs and instructor pilots, was increased from 14 AMTs to 22 and 100 instructor pilots to 138. The expenditures, consisting of aircraft and active flight students, remained high. The probability density distribution output for What-if Scenario 2 can be found in Figures 14, 15, 16, and 17. Intuitively, both the Flight and Maintenance Scores improved from What-if Scenarios 1 to 2 indicating a reduction in the level of operational risk by closing the gap between the number of instructor pilots and active flight students, thus lowing the Overall Risk Score. The lowest Maintenance Score occurred in What-if Scenario 2 indicating the ratio of 22 technicians to 82 aircraft is optimal.

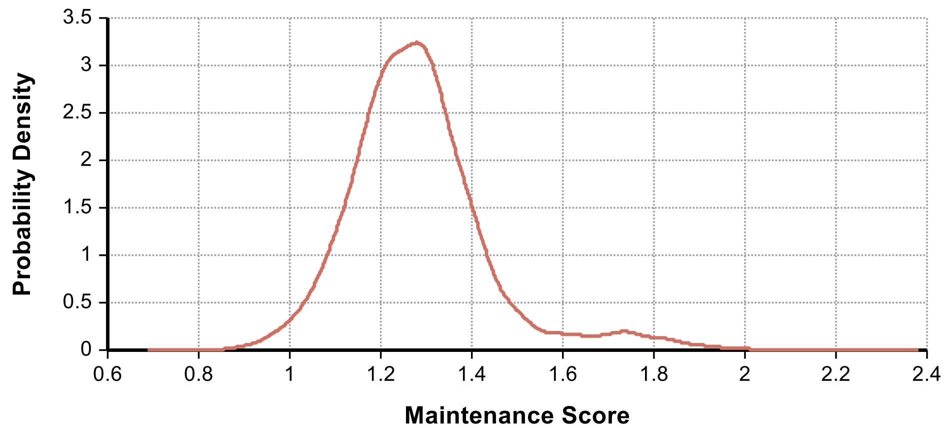



*Figure 14.* Probability density distribution of the Maintenance Score in What-if Scenario 2.

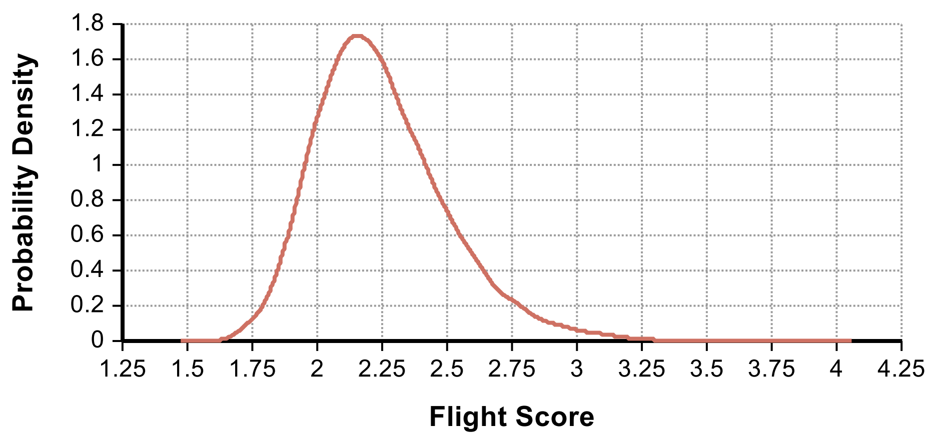


*Figure 15.* Probability density distribution of the Flight Score in What-if Scenario 2.

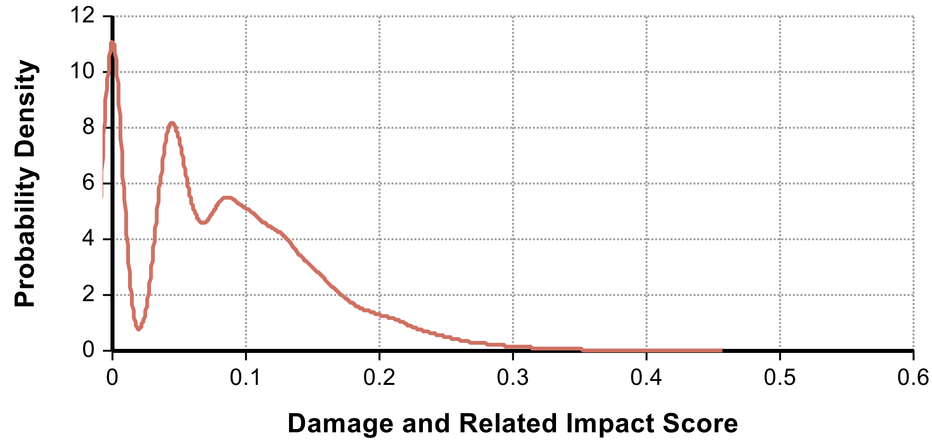



*Figure 16.* Probability density distribution of the Damage and Related Impact Score in What-if Scenario 2.

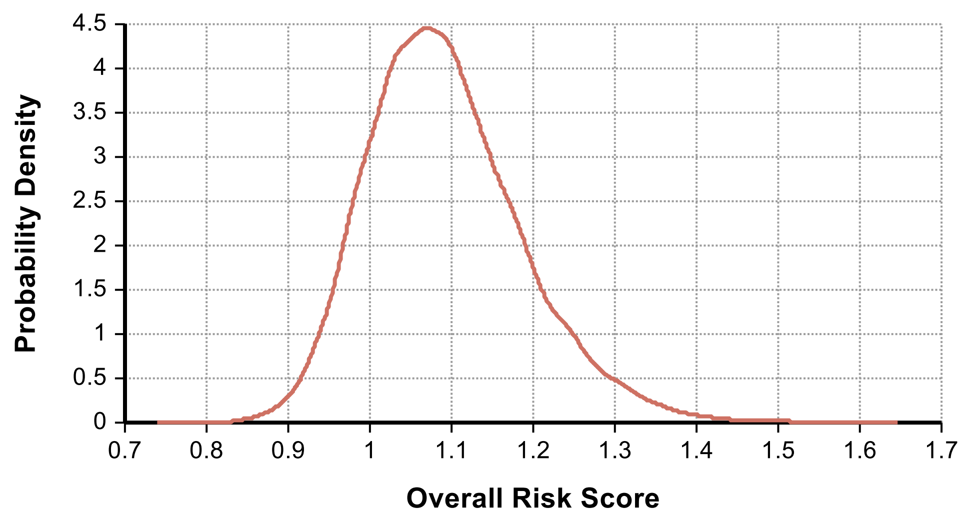



*Figure 17.* Probability density distribution of the Overall Risk Score in What-if Scenario 2.

What-if Scenario 3 was conducted with the intent of simulating a scenario opposite of What-if Scenarios 1 and 2 where there is an excess of personnel and a low level of expenditures, including a low number of flight students and few aircraft available. The probability density distribution output for What-if Scenario 3 can be found in Figures 18, 19, 20, and 21. The excess of personnel drove the Maintenance Score up from the previous trials indicating an excess of available maintenance technicians increased the level of risk within the maintenance department, negatively impacting safety. The Flight Score was the lowest in What-if Scenario 3 indicating a 1:1 ratio of instructor pilots to flight students is optimal. Of all four What-if Scenarios, What-if Scenario 3 had the lowest Overall Risk Score (*M* = 0.8845, *SD* = 0.0955) indicating the safest level of operating conditions compared to the other three trials.

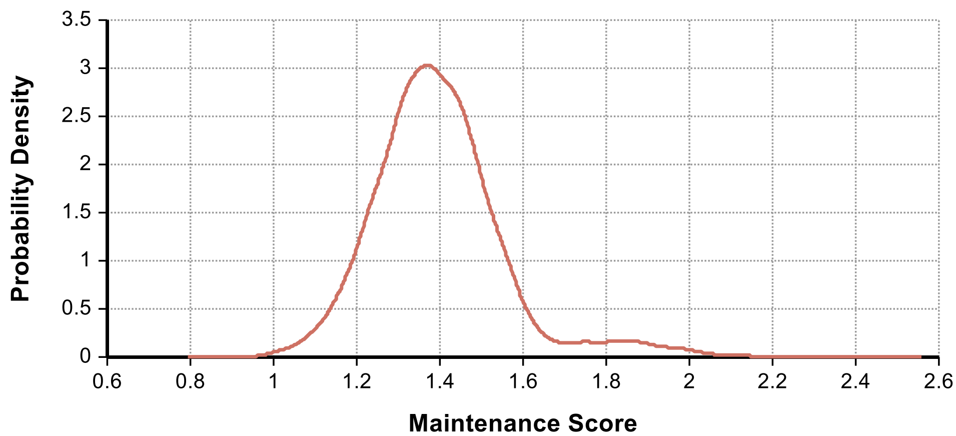



*Figure 18.* Probability density distribution of the Maintenance Score in What-if Scenario 3.

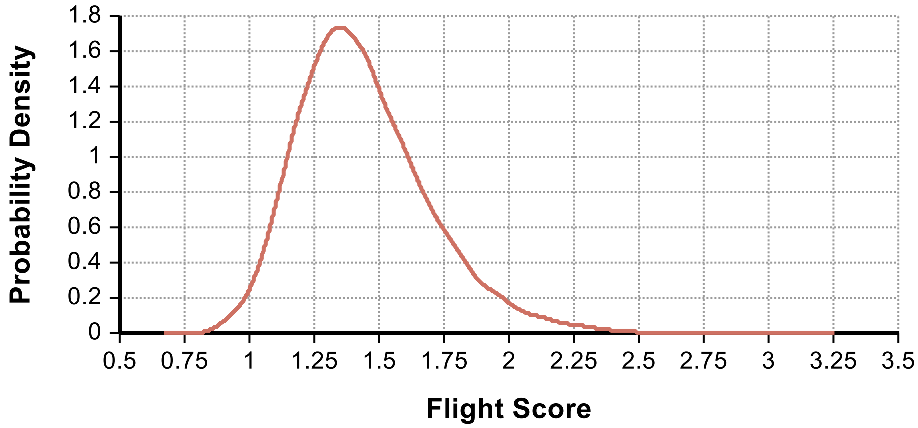



*Figure 19.* Probability density distribution of the Flight Score in What-if Scenario 3.

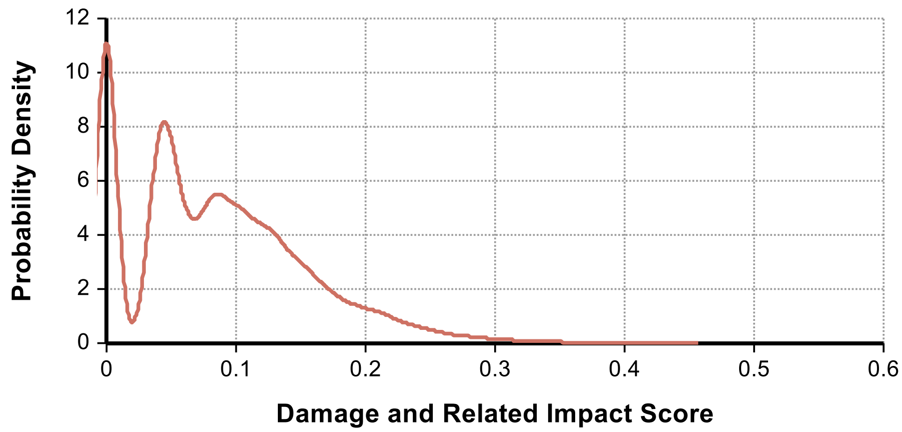



*Figure 20.* Probability density distribution of the Damage and Related Impact Score in What-if Scenario 3.

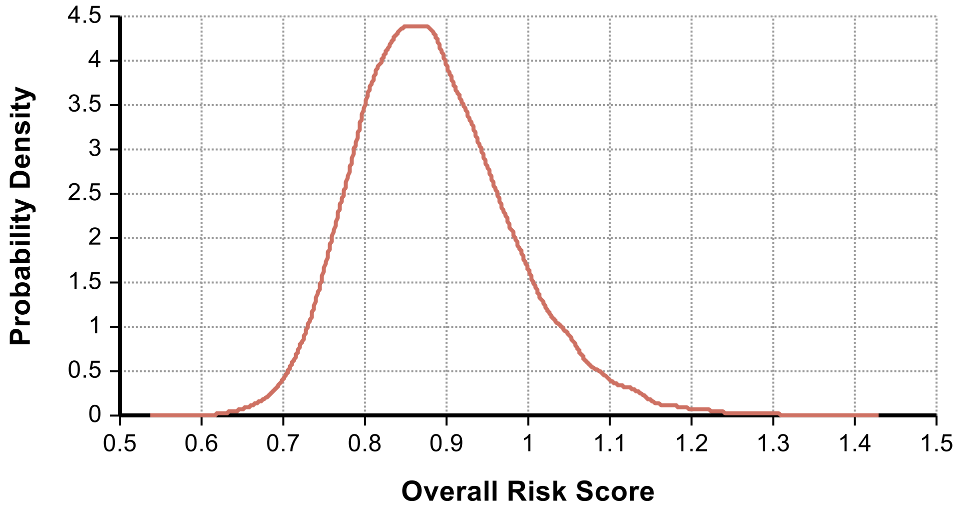



*Figure 21.* Probability density distribution of the Overall Risk Score in What-if Scenario 3.

Finally, What-if Scenario 4 was conducted with the intent of simulating a scenario similar to What-if Scenario 3; however, in regard to the expenditures, aircraft was increased from 50 to 56, and the number of flight students was increased from 335 to 681. The amount of available personnel remained high. The probability density distribution output for What-if Scenario 4 can be found in Figures 22, 23, 24, and 25. Within What-if Scenario 4, the Flight Score increases from 1.441 to 1.621 indicating the level of risk increases as the gap between the number of personnel and expenditures closes.


*Figure 22.* Probability density distribution of the Maintenance Score in What-if Scenario 4.

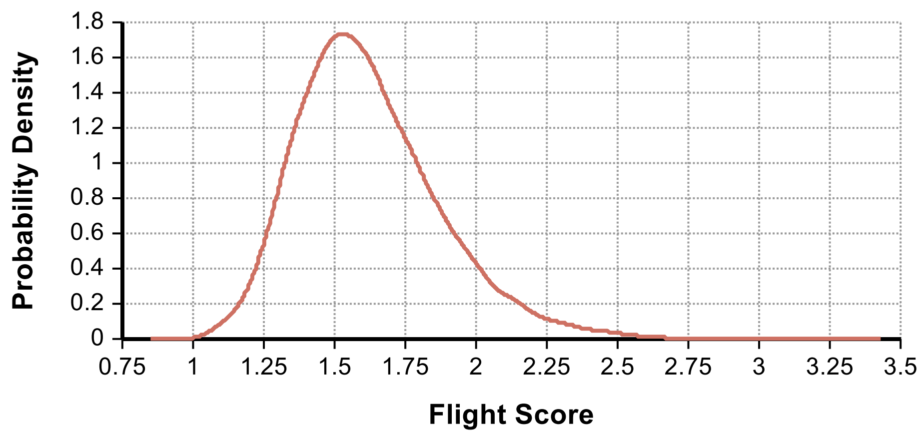



*Figure 23.* Probability density distribution of the Flight Score in What-if Scenario 4.

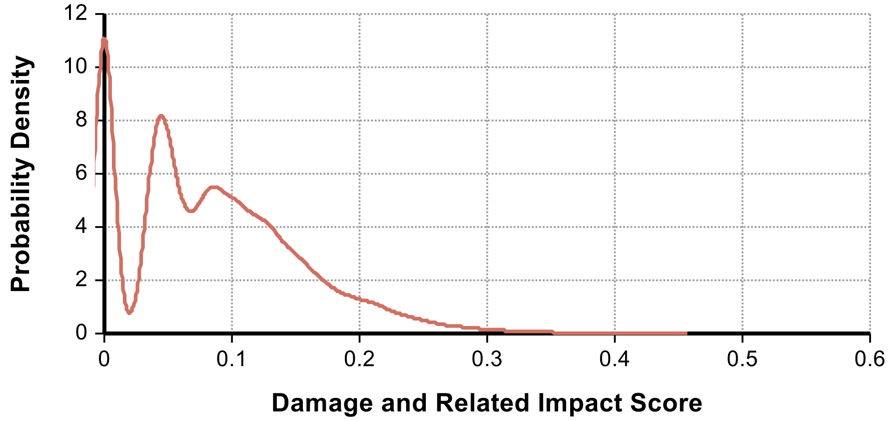



*Figure 24.* Probability density distribution of the Damage and Related Impact Score in What-if Scenario 4.

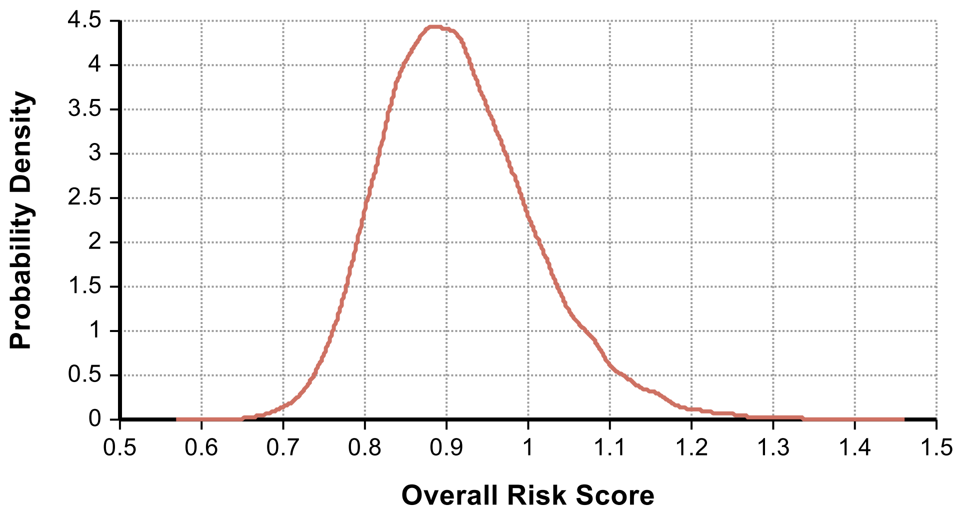



*Figure 25.* Probability density distribution of the Overall Risk Score in What-if Scenario 4.

Table 19 depicts a comparison of the mean scores and standard deviations for What-if Scenarios 1, 2, 3, and 4 to compare the mean risk score outputs demonstrating how changes to the inputs lead to differences in the risk score outputs. Figures 26, 27, and 28 depict visual comparisons of the risk score outputs for What-if Scenarios 1, 2, 3, and 4 categorized by maintenance, flight, and overall risk score outputs. The x-axis displays the risk score outputs for the model across scenarios, and the y-axis represents the probability of occurrence in percentages.

Table 19

*What-if Scenario Comparisons*

| Output Score | What-if Scenario 1 *M (SD)* | What-if Scenario 2 *M (SD)* | What-if Scenario 3 *M (SD)* | What-if Scenario 4 *M (SD)* |
|---|---|---|---|---|
| Maintenance | 1.39 (*0.17*) | 1.283(*0.16*) | 1.396(*0.16*) | 1.317 (*0.16*) |
| Flight | 2.621 (*0.26*) | 2.248 (*0.26*) | 1.441 (0.26) | 1.621 (*0.26*) |
| Damage & Related Impact | 0.084 (*0.07*) | 0.084 (*0.07*) | 0.084 (*0.07*) | 0.084 (*0.07*) |
| Overall Risk | 1.237 (*0.10*) | 1.092 (*0.10*) | 0.8845 (*0.10*) | 0.9149 (*0.09*) |

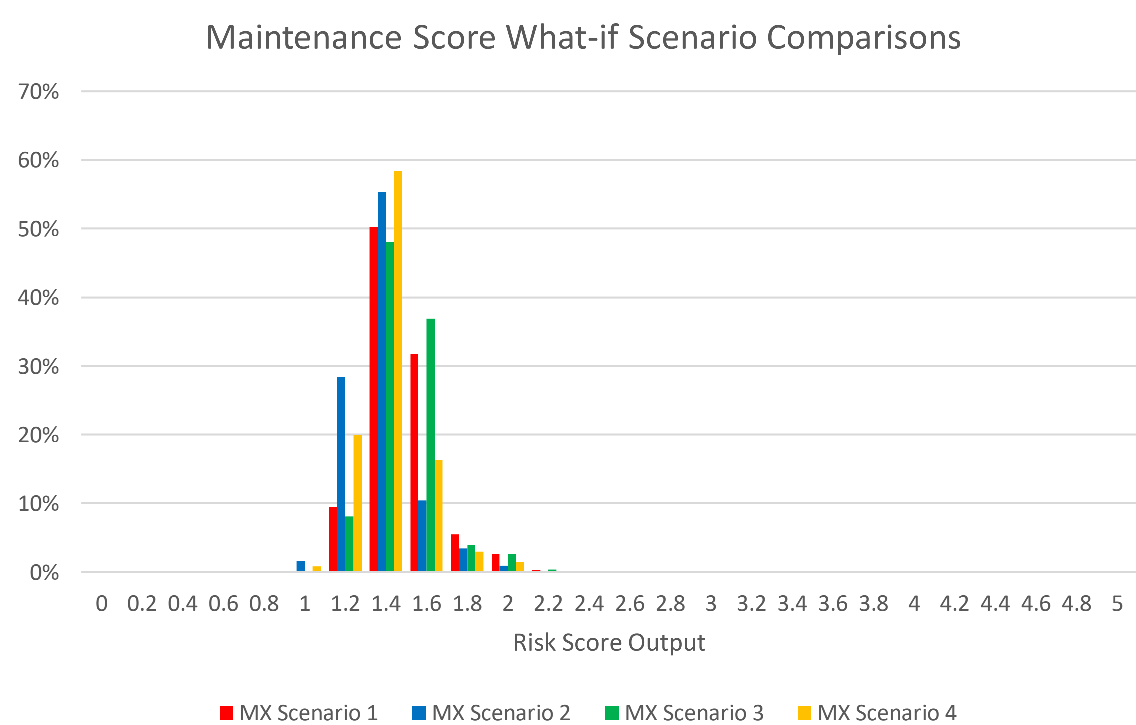



Figure 26. Maintenance Score What-if Scenario Comparison Chart

Results indicate the lowest risk score for maintenance occurred in What-if Scenario 2, where the level of personnel was moderate, yet expenditures, including aircraft and students, were high.

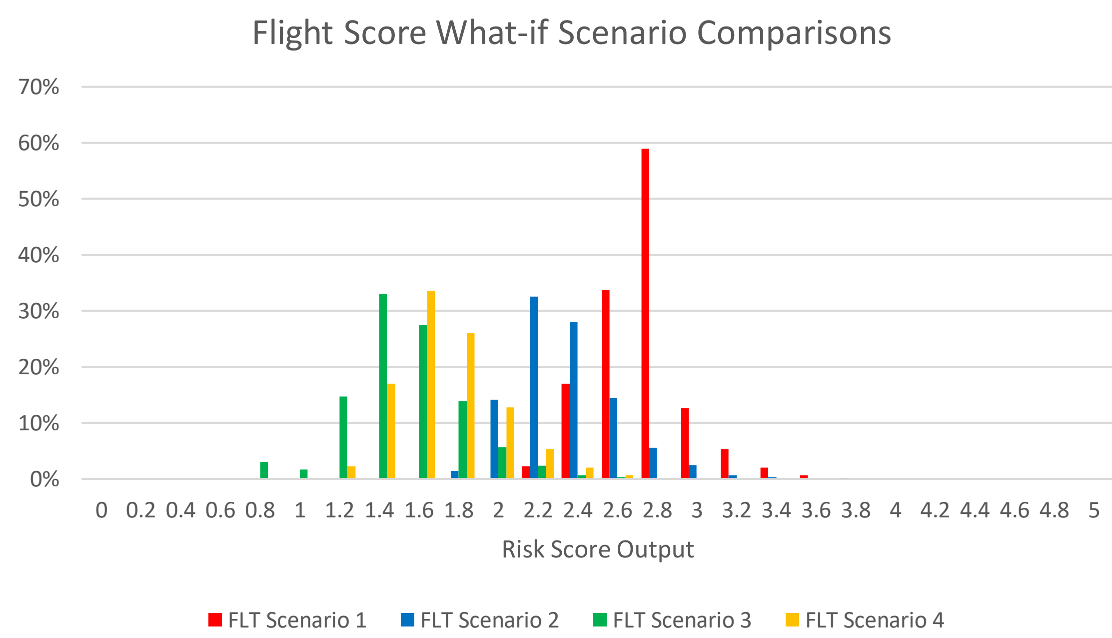



Figure 27. Flight Score What-if Scenario Comparison Chart

The lowest risk score for flight occurred in What-if Scenario 3, where the level of personnel was high, and expenditures were low. The Damage and Related Impact Score remained consistent throughout; thus, no visual comparisons were made.

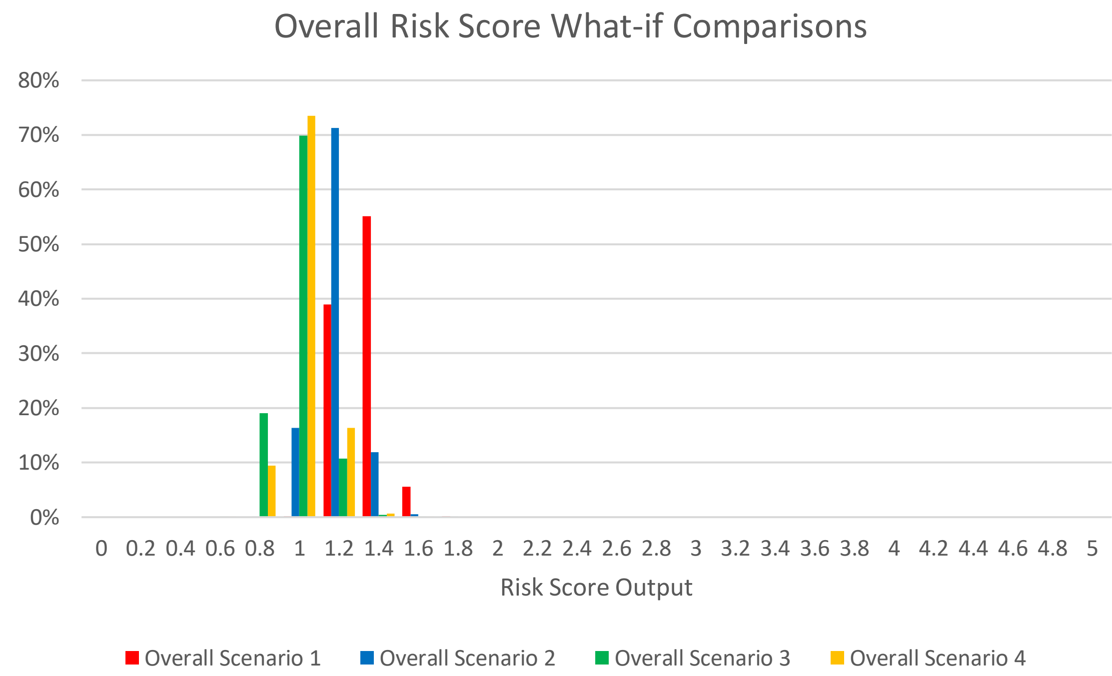



Figure 28. Overall Risk Score What-if Scenario Comparison Chart

What-if Scenario 3 also had the lowest Flight Score and Overall Risk Score, indicating operations are at the lowest level of risk when the level of personnel is high, yet the amount of expenditures remains low. Although intuitive, this demonstrates the real-world utility of the model.

A Generalized Sensitivity Analysis (GSA) was conducted using a one-factor-at-a-time approach. The purpose of conducting a GSA was to locate sensitive parameters, or those that have the greatest effect on the model, and non-sensitive parameters, or those input variables causing stagnation of the model. Findings of the GSA indicated a lack of sensitivity within the Damage and Related Impact Score. This may be partially due to the Impact Factor feeding into the model as a constant variable with a definition of 1, chosen to represent a scenario where no damage and no injuries have occurred. Due to the obscurity of accidents in CFR Part 141 flight training operations; injuries, fatalities, and the extent of damage are situationally specific and thus challenging variables to predict. Due to this, a constant of 1 was used, indicating no injuries and no damage to people or property occurred to demonstrate model utility. However, the sensitivity of Overall Risk Score output did not change by removing the Impact Value constant variable.

To improve the overall sensitivity of the model, What-if Scenarios 1, 2, 3, and 4 were rerun as Sensitivity Trial Scenarios 1, 2, 3, and 4 with the Damage and Related Impact Score pathway removed. However, removing the Damage and Related Impact Score and associated input variables also required an adaption to the Overall Risk Score equation and model weights. The adapted equation for the Overall Risk Score is portrayed below. The model weights for Maintenance and Flight were changed from 0.3 to 0.5 to accommodate the removal of the Damage and Related Impact score, which had a weight of 0.4.

$$(Maintenance\ Score * 0.5) + (Flight\ Score * 0.5)$$

Figure 29 depicts the conceptual layout of the model in Analytica with the Damage and Related Impact Score pathway and associated input variables removed to determine if the sensitivity of the model improves. Results of Sensitivity Trial Scenarios 1, 2, 3, and 4 can be found in Table 20.

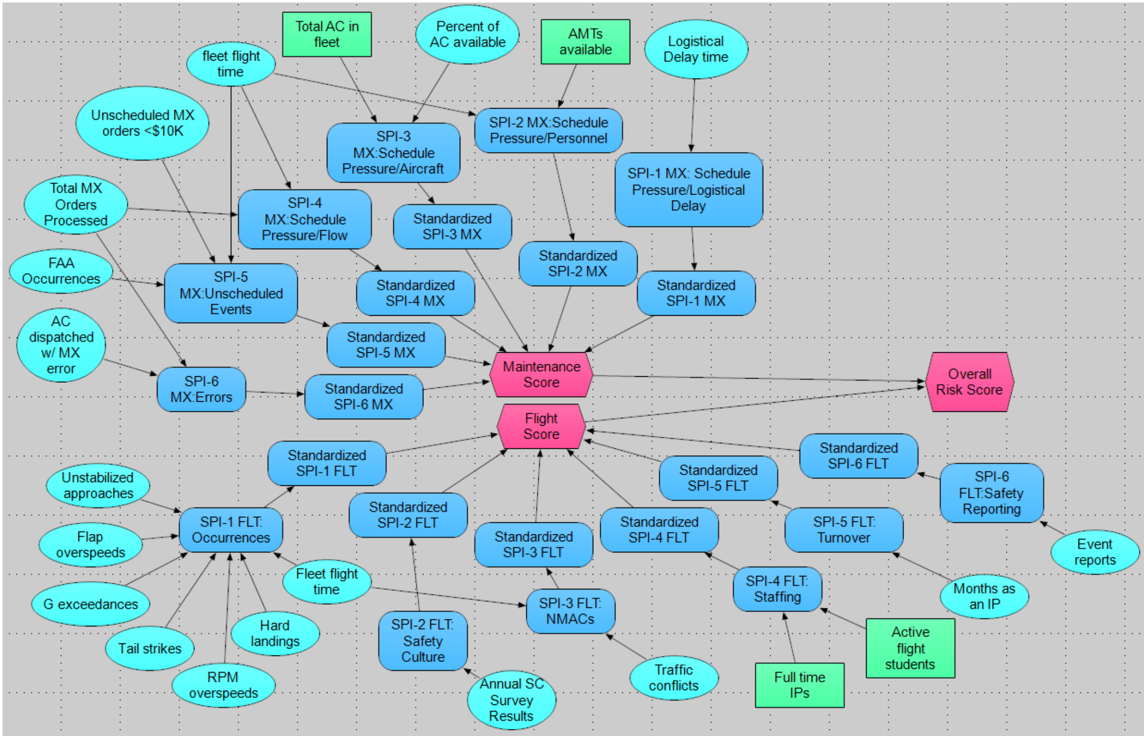


*Figure 29.* Conceptual layout of the model in Analytica for sensitivity analysis.

Table 20

*Results of Sensitivity Trial Scenarios 1, 2, 3, and 4*

| Scenario | Maintenance Score *M (SD)* | Flight Score *M (SD)* | Overall Risk Score *M (SD)* |
|---|---|---|---|
| What-if Scenario 1 | 1.39 (*0.17*) | 2.621 (*0.26*) | 1.237 (*0.10*) |
| Sensitivity Trial Sceanrio 1 | 1.39 (*0.17*) | 2.621 (*0.26*) | 2.005 (*0.15*) |
| What-if Scenario 2 | 1.283(*0.16*) | 2.248 (*0.26*) | 1.092 (*0.10*) |
| Sensitivity Trial Scenario 2 | 1.283(*0.16*) | 2.248 (*0.26*) | 1.765 (*0.15*) |
| What-if Scenario 3 | 1.396(*0.16*) | 1.441 (0.26) | 0.8845 (*0.10*) |
| Sensitivity Trial Scenario 3 | 1.396(*0.16*) | 1.441 (0.26) | 1.419 (*0.15*) |
| What-if Scenario 4 | 1.317 (*0.16*) | 1.621 (*0.26*) | 0.9149 (*0.09*) |
| Sensitivity Trial Scenario 4 | 1.317 (*0.16*) | 1.621 (*0.26*) | 1.469 (*0.15*) |

Overall, results of the Sensitivity Trial Scenarios indicated that removing the Damage and Related Impact Score pathway did little to improve the sensitivity of the model. As demonstrated in Table 20, removing the Damage and Related Impact Score pathway only impacts the Overall Risk Score. However, with this pathway removed, the Overall Risk Score outputs are slightly higher, capturing an increased level of risk than they had been within the What-if Scenarios.

Removing the Damage and Related Impact Score pathway restricts the utility of the model by failing to account for the key variables included by Anderson et al. (2020) in the non-statistical model due to its value in depicting the overall level risk associated with the operation at a particular given time. The Damage and Related-Impact variable, although reactive in nature and challenging to accurately forecast, provides an external perception of the risk associated with the whole operation and should remain a valuable portion of the safety decision-making tool.

**Summary**

Using Monte Carlo simulation, a safety decision-making tool was developed to assess how changes to the controllable input variables impact the level of operational risk within a large, collegiate CFR Part 141 flight training organization. Before model execution, input nodes supplying distribution data were examined to ensure the output produced by the model aligns with the predetermined probability distributions of the uncontrolled input variables, as determined by a two-year sample of raw operational flight and maintenance data ranging from September 2017 to September 2019 from a large, collegiate CFR Part 141 flight training organization in the southeastern United States. The output of each computational node of the model was verified by comparing the node’s output to the results of manual calculations drawn from the two-year sample of operational data. There were 22 uncontrollable inputs to the model.

For model verification purposes, the output of each of these distributions was examined from a simulation run with 10,000 trials. Three Verification Scenarios were conducted. The values for the controllable input variables in Verification Scenario 2 were drawn from the low values of the operational ranges for CFR Part 141 flight training organizations, and the controllable input variables for Verification Scenario 3 were drawn from high operational range values. High and low range values were selected to represent the varying operational capacities of the target population. To ensure no programming error occurred during the construction of the model, two additional Verification Scenarios were performed using different controllable input variables.

Reliability Testing was performed using different random number generator seed values to verify the model produced consistent results. The study compared the output from three different runs of the model—each using a unique seed value to ensure a different sample of random numbers for the uncontrollable input variables, which remained the same across trials. Based on the results of ANOVA output, no significant differences appeared among the different sets of results, indicating the results are statistically reliable.

This research utilizes the validated equations drawn from the non-statistical model developed by Anderson et al. (2020) for the mathematical inputs driving the computational nodes, including the SPIs, the Flight Score, Maintenance Score, Damage and Related Impact Score, and the Overall Risk Score, as the foundation to develop the safety performance decision-making tool. The peer-reviewed research conducted by

Anderson et al. (2020) validated the non-statistical model and associated equations via the use of Subject Matter Experts using a standardized expert elicitation survey questionnaire. Thus, the equations driving the predictive, safety performance decision-making tool developed in this dissertation have been previously validated through the peer-reviewed research conducted by Anderson et al. (2020).

What-if Scenarios 1, 2, 3, and 4 were conducted to demonstrate the utility of the safety performance decision-making tool for real-world use. The controllable input values used to generate the what-if scenarios within the Monte Carlo simulation model were determined based on permutational variations of the ranges of normal operating conditions for the target population − large, collegiate CFR Part 141 flight training organizations. Each what-if scenario ran the model through 10,000 trials to generate the output datasets. Comparison of the four trials effectively demonstrated the utility and potential for the safety performance decision-making tool.

Results of the GSA indicated that removing the Damage and Related Impact Score pathway does improve the sensitivity of the model; however, the improvement is very minor. Removing the Damage and Related Impact Score pathway restricts the utility of the model by failing to account for the key variables included by Anderson et al. (2020) in the non-statistical model due to its value in depicting the overall level of risk associated with the operation at a particular given time.

.

# CHAPTER V

## DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

This chapter discusses the results described in Chapter IV and addresses the research questions from Chapter I. This chapter examines the data produced by the simulation model developed for this study, discusses the analysis of the data, and identifies the study's conclusions. Finally, this chapter discusses the limitations of the study and provides recommendations for future research.

The purpose of the research was to create and validate a safety performance decision-making tool to transform a non-statistical model composed of 12 SPIs determined by Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020) to be most indicative of flight risk specific to 14 CFR Part 141 flight training organizations into a predictive, safety performance decision-making tool. The model uses what-if scenarios to evaluate how changing controllable input variables affect the level of operational risk within the system, portrayed within the model as the risk score outputs.

The study derived the outputs, or risk scores, from a Monte Carlo simulation model. A Monte Carlo simulation accounts for the uncertainties present within the real-world operating conditions of a complex, collegiate CFR Part 141 flight training organization. The model created for this study produced probability distribution output data to provide critical, safety decision-making information on the level of operational risk associated with manipulating the following controllable input variables: number of aviation maintenance technicians available, number of aircraft available, number of full-time instructor pilots, and the number of active flight students. The data driving the distributions for the uncontrollable input variables found in Table 3 were drawn from a two-year sample of operational flight and maintenance data from a large, collegiate CFR Part 141 flight training organization in the southeastern United States.

## Discussion

To effectively create and validate a safety decision-making tool, it was first necessary to define both the scenarios and the input values to be used by the model. Since this study focused on creating a model specific to large, collegiate Title 14 CFR Part 141 flight training organizations, the selected permutational scenarios intended to represent the vast range of operating conditions for collegiate, 14 CFR Part 141 flight training organizations. The probability distributions used for the uncontrollable input variables were also drawn from the same two-year sample of operational flight and maintenance data from a large, collegiate flight training organization in the southeastern United States. Thus, the risk score outputs of the model are specific to the operating conditions of the particular CFR Part 141 flight training organization used within the sample. However, the equations driving the predictive model have been validated in the peer-reviewed literature by Anderson et al. (2020) indicating the model could easily be adapted for immediate use by any collegiate, Title 14 CFR Part 141 flight training organization with data collection capabilities and an active SMS by determining the appropriate uncontrollable input distributions specific to that organization's operating conditions.

The following section will address the research questions driving the study, explain how findings are supported, and describe how the findings fit into the existing body of knowledge surrounding predictive modeling for large, collegiate CFR Part 141 flight training organizations.

> Research Question 1: How can the SPI model developed by Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020) be transformed into a predictive, safety performance decision-making tool with the ability to run what-if scenarios?

To address Research Question 1, this dissertation outlines the process of transforming a non-statistical risk assessment model developed by Anderson et al. (2020) composed of 12 domain-specific SPIs and associated equations into a predictive, safety performance decision-making tool using a two-year sample of operational flight and maintenance data from a large, collegiate CFR Part 141 flight training organization in the southeastern United States to determine uncontrollable input probability distributions and demonstrate the utility of the model for real-world use. The safety performance decision-making tool created for this dissertation utilizes what-if scenarios to simulate how changes to the four controllable input variables influence the risk scores, or outputs.

Documented within the first three sections in Chapter IV, verification, reliability, and validity testing was either discussed or conducted on the safety performance decision-making tool to ensure findings were supported. Using Analytica 64-bit Educational Professional software Release 4.6.1.30 by Lumina Decision Systems, three Verification Scenarios were run on the predictive model. Within Verification Scenario 1, the values selected to serve as controllable input values were determined by calculating the mean value for each variable based on the two-year sample of raw data. The purpose of using mean values for comparison purposes was to ensure the output of the model was representative of the raw sample data from the CFR Part 141 flight training organization. The values of the controllable input variables in Verification Scenario 2 were drawn from the low values of the operational ranges. Finally, values for the controllable input variables in Verification Scenario 3 were drawn from high operational range values for CFR Part 141 flight training organizations. High and low range values were selected to represent the varying operational capacities of the target population. Results indicated that the simulation model's mean output value fell between the higher and lower limits of the raw data sample. Overall, the model produced the results expected based on the controllable input variable specifications, effectively verifying the efficacy of the transition from a non-statistical risk assessment model in a predictive, safety performance decision-making tool.

To further support the findings, reliability testing was conducted on the output of the simulation model. The outputs from three different runs of the model were compared—each using 10,000 trials and a unique seed value to ensure a different sample of random numbers for the uncontrollable input variables. Analyzing the output with different seed values allows for the model to be verified for consistency in its results despite the changes produced by the random number generator. Mean probability was the key output for this model. The mean probability output represents the forecasted level of operational risk on a standardized 0-5 risk scale for the Flight Score, Maintenance Score, Damage and Related Impact Score, and Overall Risk Score. The results of the reliability trials were analyzed using ANOVA to test for differences across the three groups (Hoyt, 1941). No significant differences appeared among the different sets of results indicating the results are statistically reliable.

The challenge with results comparison between this model and the models developed in other studies is that no other studies directly address the same research questions. Additionally, little work has been done in the realm of predictive modeling specific to large, collegiate CFR Part 141 flight training organizations, leaving a deficit of validated models for comparison. However, this research utilized the peer-reviewed and validated equations drawn from the non-statistical model developed by Anderson et al. (2020) for the mathematical inputs driving the computational nodes, including the SPIs, the Flight Score, Maintenance Score, Damage and Related Impact Score, and the Overall Risk Score, as the foundation to develop the safety performance decision-making tool.

Research Question 2: How do changes to the controllable input variables impact the overall risk score?

To address Research Question 2 and demonstrate the utility of the model for real-world use, distributions and ranges of values were utilized to simulate the many thousands of potential outcomes within the what-if scenarios allowing for an assessment of how the changes to the controllable input variables impact the risk scores. The controllable input values used to generate the what-if scenarios within the Monte Carlo simulation model were determined based on permutational variations of the range of normal operating conditions specific to CFR Part 141 flight training organizations. These permutations were conducted by varying the level of personnel, concerning available aviation maintenance technicians and instructor pilots, as low, moderate, or high. Similarly, permutations of resource expenditures, including aircraft available and active flight students, were also varied by degree of low, moderate, or high. Low values consisted of the lowest range values, moderate values consisted of the median value, and high range values consisted of the highest potential value of the predetermined, true, operational ranges for a large, collegiate CFR Part 141 flight training organization.

To support the findings and demonstrate the utility of the safety performance decision-making tool, four What-if Scenarios were conducted by manipulating the controllable input variables, or resources including personnel, students, and aircraft. What-if Scenario 1 was run with the intent of simulating a scenario where personnel, with regard to AMTs and instructor pilots, are low, but the necessary expenditures, consisting of aircraft and active flight students, was high. Within What-if Scenario 2, the number of personnel, including AMTs and instructor pilots, was increased from 14 AMTs to 22 and 100 instructor pilots to 138. The expenditures, consisting of aircraft and active flight students, remained high. What-if Scenario 3 was conducted with the intent of simulating a scenario opposite of What-if Scenarios 1 and 2 where there is an excess of personnel and a low level of expenditures, including a low number of flight students and available aircraft. Within What-if Scenario 4, the number of expenditures in terms of aircraft was increased from 50 to 56, and the number of flight students was increased from 335 to 681. The amount of available personnel remained high.

Table 21

*What-if Scenario Comparisons*

| | What-if Scenario 1 | What-if Scenario 2 | What-if Scenario 3 | What-if Scenario 4 |
|---|---|---|---|---|
| Output Score | *M (SD)* | *M (SD)* | *M (SD)* | *M (SD)* |
| Maintenance | 1.39 (*0.17*) | 1.283(*0.16*) | 1.396(*0.16*) | 1.317 (*0.16*) |
| Flight | 2.621 (*0.26*) | 2.248 (*0.26*) | 1.441 (0.26) | 1.621 (*0.26)* |
| Damage & Related Impact | 0.084 (*0.07*) | 0.084 (*0.07*) | 0.084 (*0.07*) | 0.084 (*0.07*) |
| Overall Risk | 1.237 (*0.10*) | 1.092 (*0.10*) | 0.8845 (*0.10*) | 0.9149 (*0.09*) |

Table 21 depicts a comparison of What-if Scenarios 1, 2, 3, and 4, demonstrating how changes to the inputs lead to differences in the risk score outputs. Results of the four What-if Scenarios indicate the lowest risk score for maintenance occurred in What-if Scenario 2, where the level of personnel was moderate, yet expenditures, concerning aircraft and students, were high. The lowest risk score for flight occurred in What-if Scenario 3, where the level of personnel was high, and expenditures were low. The Damage and Related Impact Score remained consistent throughout. What-if Scenario 3 also had the lowest Flight Score and Overall Risk Score, indicating operations are at the lowest level of risk and optimum level of safety among trials under the following specifications:

- Aviation Maintenance Technicians available: 35
- Aircraft available: 50
- Instructor Pilots: 200
- Active Flight Students: 335

As demonstrated by the mean probability output data produced by the simulation model, changes to the controllable input variables are reflected by variations to the risk score outputs demonstrating the utility and potential for the safety performance decision-making tool. The risk score outputs produced from the what-if scenarios could then be utilized by safety personnel and administration to make more informed safety-related decisions, based on the mean level of operational risk predicted without expending unnecessary resources. The lowest Overall Risk Score occurs in What-if Scenario 3, indicating CFR Part 141 flight training organizations should strive to maintain an appropriate balance of high personnel to low expenditures to maintain the optimum level of operational safety.

This research fits into the existing body of knowledge surrounding the area of predictive aviation safety assessment techniques by providing detailed insight into the process of transitioning from traditionally reactive SPIs into a safety performance decision-making tool with forecasting abilities specific to large, collegiate CFR Part 141 flight training organizations. The extant literature indicated a deficit of these predictive, domain-specific safety performance decision-making tools. Thus, this reusable model pioneers the way for the inclusion of validated safety performance decision-making tools into the risk management component of flight training organizations' SMS.

**Conclusions**

This dissertation demonstrated the process of transitioning from a non-statistical model composed of domain-specific, yet reactive, SPIs into a safety performance decision-making tool with forecasting abilities for safety decision-making purposes, specific to CFR Part 141 flight training organizations improving the risk management component of CFR Part 141 flight training organizations' Safety Management System (SMS). Figure 5 illustrates the conceptual layout and structural definitions of the model in Analytica from Lumina Decision Systems. ANOVA testing found no significant differences between sets of results, indicating the model is statistically reliable. As the mathematical inputs driving the computational nodes, or SPIs, are drawn from peer-reviewed and previously validated research conducted by Anderson et al. (2020), the model is considered valid. Finally, What-if Scenarios 1, 2, 3, and 4 were conducted to effectively demonstrate the utility of the safety performance decision-making tool in influencing the risk score outputs.

**Theoretical contributions.** This dissertation describes the process of transforming a nonstatistical model composed of domain-specific SPIs into a safety performance decision-making tool. It also extends the previously validated non-statistical model composed of SPIs determined by Anderson et al. (2020) to be most indicative of flight risk specific to large, collegiate CFR Part 141 flight training organizations to create a new, predictive, safety performance decision-making tool with the ability to run what-if scenarios. Determining the probability distributions of the uncontrollable input variables from the sample data allowed for the nonstatistical model to be transformed into a predictive, safety performance decision-making tool.

The study demonstrates the utility of Monte Carlo simulation as a viable approach for handling input parameters with varying levels of uncertainty to assist in administrative, safety decision-making. Describing the potential outcomes as a range of outcomes provides insight into how potential changes to controllable inputs affect the level of risk within the system while acknowledging the results of actually making real-world changes to the system may vary due to the uncertainties involved.

The model will also provide a mechanism for expanding the breadth of knowledge related to optimizing resources from both flight and maintenance operations to enhance operational safety for CFR Part 141 flight training organizations. Further, a thorough review of the extant literature indicated a gap in the process of transitioning from traditionally reactive SPIs into safety performance decision-making tools with forecasting abilities for safety decision-making purposes specific to large, collegiate CFR

Part 141 flight training operations. With the literature indicating a deficit of a validated safety decision-making tool specific to CFR Part 141 flight training operations, this research has filled this gap by providing a validated, safety decision-making tool, specific to CFR Part 141 operations, to advance the applications of proactive, rather than reactive, aviation safety assessment techniques by modeling the potential of the system without compromising resources.

**Practical contributions.** From a practical standpoint, this research will aid in shaping the current understanding of the factors most substantially contributing to flight risk within large, collegiate CFR Part 141 flight training organizations, thereby improving overall flight safety. As a safety decision-making tool, the model could also be used by the administration within a large, collegiate CFR Part 141 flight training organization to rationalize hiring, technology acquisition, and other safety-related initiatives by modeling the potential of modifying resources, or controllable inputs, without the risk associated with actually expending the organization's resources.

With a consistent stream of data updated on a monthly basis, CFR Part 141 flight training organizations could utilize this safety decision-making tool to understand the impact altering the ratios of resources-to-expenditures has on the level of operational risk present within the flight department, maintenance department, and the operation overall. Results of the What-if Scenarios and Sensitivity Trial Scenarios indicated the trial with the lowest risk scores was What-if Scenario 3 and Sensitivity Trial 3. In both trials, the controllable input values were 35 AMTs to 50 aircraft and 200 full-time instructor pilots to 335 flight students. Demonstrably, CFR Part 141 flight training organizations could lower their levels of risk, thereby improving their overall safety, by maintaining conditions where there is enough personnel staffed to accommodate the level of expenditures, including aircraft and active flight students.

For this dissertation, the data supplying the probability distributions for the uncontrollable input variables are drawn from a two-year sample of operational data ranging from 2017-2019 for a large, collegiate 14 CFR Part 141 flight training organization. However, the model could be adapted for use in any CFR Part 141 flight training organization with data acquisition capacities and an operational SMS simply by modifying the input value probability distributions to reflect the operating conditions of the selected 14 CFR Part 141 flight training organization. Providing collegiate CFR Part 141 flight training organizations with a safety decision-making tool will enhance the risk management component of the operation's SMS by taking an increasingly proactive approach to safety by providing insight into the impact changes to operating conditions may have on the safety of the overall operation determined by evaluating the quantitative risk score outputs. The ability to forecast operating conditions using Monte Carlo simulation will allow CFR Part 141 flight training organizations to make better informed safety-related decisions while optimizing efficiency without compromising safety.

**Limitations of the Findings**

The research was limited to the creation and validation of a safety performance decision-making tool utilizing Monte Carlo simulation to transform a non-statistical model composed of the ten SPIs determined by Anderson et al. (2020) into a predictive, safety decision-making tool capable of running what-if scenarios to determine how changes to input variables affect the levels of operational risk within the organization. The variables used in this model are limited to those found to be most relevant to measuring flight risk in a large, collegiate CFR Part 141 flight training organization by SMEs in the areas of both flight and maintenance (Anderson et al., 2020). The four controllable input variables selected for use in the simulation are just four pieces of a large and complex system. As demonstrated within What-if Scenarios 1-4, manipulating these controllable inputs does not drastically impact the risk score outputs, as the ranges of normal operating conditions used to determine the values for the controllable input variables may have not been broad enough to capture more dynamic variations to the risk score outputs.

## Recommendations

The results of this study demonstrated the creation and validation of a safety performance decision-making tool. The safety performance decision-making tool should be utilized by safety personnel and administrators to make more informed safety-related decisions, based on the level of risk predicted by the manipulation of controllable input variables within the what-if scenarios, without expending unnecessary organizational resources.

**Recommendations for large, collegiate 14 CFR Part 141 flight training organizations**. Large, collegiate 14 CFR Part 141 flight training organizations should improve and streamline their operational data collection capabilities and storage to ensure the model is provided with accurate data to determine the uncontrollable input probability distributions. Additionally, 14 CFR Part 141 flight training organizations should utilize a larger sample of raw operational flight and maintenance data to ensure the accuracy of the probability distributions for the uncontrollable inputs and the predictive utility of the model. Finally, large, collegiate 14 CFR Part 141 flight training organizations should explore the potential of utilizing different controllable input variables for use within the model.

CFR Part 141 flight training organizations could utilize this safety decision-making tool to run what-if scenarios to understand the impact of altering the quantity of resources and expenditures, with regard to the number of AMTs available, the number of aircraft available, the number of full-time instructor pilots, and the number of active flight students and the influence these changes make on the level of operational risk for the flight department, maintenance department, and the operation overall. CFR Part 141 flight training organizations could also use the model to determine an acceptable level of risk particular to their operation based on the manipulation of resources. Results of the model, based on the probability distributions drawn from a two-year sample of operational data, indicated the trial with the lowest risk scores occurred when there is enough personnel staffed to accommodate a low level of expenditures. In both trials, the controllable input values were 35 AMTs to 50 aircraft and 200 instructor pilots to 335 flight students simulating a scenario where the level of personnel is high but expenditures are low. To reduce the level of overall risk within the organization, CFR Part 141 flight training organizations should evaluate their current ratios of AMTs to aircraft and instructor pilots to flight students to maintain an optimized level of balance and direct financial resources to accommodate an operation where the level of personnel is high yet expenditures are low.

**Recommendations for future research.** Future research should focus on opportunities to further explore both the capabilities of the model and options for improving the accuracy of the model’s predictions. The ranges of normal operating conditions used to determine the values for the controllable input variables may have not been broad enough to capture the potential for more dynamic variations to the risk score outputs. Future research could focus on expanding the range of operational values when determining controllable input variables to assess how changes to the risk score outputs are impacted with a more expansive range of operating conditions.

In an attempt to increase the predictive potential of the model, future research should reevaluate the Damage and Related Impact variable, as it is composed of variables that are reactive in nature, making this SPI challenging to predict. Increasing the predictive accuracy of the Damage and Related Impact variable may increase the sensitivity of the Overall Risk Score output. Future research should also explore the potential of including additional controllable input variables, thereby leaving less up to chance. Future research should aim to improve the overall utility of the model for 14 CFR Part 141 flight training organizations by incorporating clear, measurable human performance variables into the model, assuming the data is available. To enhance the robustness of the model, future research should consider incorporating the three indicators (NAC, Weather, and ATC Delay), included in the original model by Anderson et al. (2020), due to their potential correlations with the SPIs and their unpredictable influence on day-to-day flight operations.

Additionally, future research should explore the potential of incorporating machine learning techniques to allow for the data supplying the probability distributions for the uncontrolled input variables to be updated on a regular basis eliminating the need to manually update the distributions. This will improve the accuracy and predictive capabilities of the model. As monte carlo simulation can be used to quantify risk, future research should also consider the alternative approach of utilizing optimization techniques to further minimize risk.

**Summary**

The purpose of the dissertation was to create and validate a safety performance decision-making tool to transform a reactive safety model into a predictive, safety performance decision-making tool, specific to large, collegiate Title 14 CFR Part 141 flight training organizations, to increase safety and aid in operational decision-making. The validated safety decision-making tool uses what-if scenarios to assess how changes to the controllable input variables impact the overall level of operational risk within an organization’s flight department.

SPIs from the non-statistical SPI model developed by Anderson et al. (2020) were used to create the safety performance decision-making tool, as these SPIs are most indicative of operational flight risk for a 14 CFR Part 141 flight training organization. However, a 14 CFR flight training organization could utilize its own unique SPIs by determining the probability distributions of the uncontrollable input variables, further enhancing the generalizability of the safety performance decision-making tool. Anderson et al. (2020) created and validated, via expert elicitation a non-statistical model composed of SPIs from both flight and maintenance operations and their relevant formulae based on two years of operational flight and maintenance data. The SPIs from the non-statistical model developed by Anderson et al. (2020) were used as the foundation to develop a safety performance decision-making tool based on the input variables for the chosen SPIs. Monte Carlo simulation was conducted and run to enable the SPI model to handle uncertainty in some of the key, influential variables.

As a safety decision-making tool, the model could also be used by the administration within a large, collegiate CFR Part 141 flight training organization to rationalize hiring, technology acquisition, and other safety-related initiatives by modeling the potential of modifying resources, or controllable inputs, without the risk associated with actually expending the organization’s resources. The model could also be adapted for use in any flight training organization with data acquisition capabilities and an active SMS.

The research methodology has been designed to address the following research questions:

1. How can the SPI model developed by Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020) be transformed into a predictive, safety performance decision-making tool with the ability to run what-if scenarios?
2. How do changes to the controllable input variables impact the overall risk performance score?

To address Research Question 1, Monte Carlo simulation was utilized to transform the non-statistical risk assessment model composed of SPIs developed by Anderson et al. (2020) into a predictive, safety performance decision-making tool. In response to Research Question 2, distributions and ranges of values were utilized to simulate the many thousands of potential outcomes within the what-if scenarios allowing for an assessment of how the changes to the controllable input variables impact the overall level of operational risk. After manipulating the controllable input variables, or resources with regard to personnel, students, and aircraft, the probability distribution output from the what-if scenarios then allows safety personnel and administration to make more informed safety-related decisions, based on the level of risk predicted by the what-if scenarios, without expending unnecessary resources.

The target population to which the model generalizes is large, collegiate CFR Part 141 flight training organizations within the United States operating under the specifications defined by the FAA within Title 14 of the Code of Federal Regulations Part 141 (Federal Aviation Administration, 2017). The sampling frame consisted of two-years of operational data from both flight and maintenance operations dating from September 2017 to September 2019 for a large, collegiate CFR Part 141 flight training organization in the southeastern United States. The sample data used to determine the probability distributions of the uncontrollable input variables within the model was comprised of two years of operational flight and maintenance data from a large, collegiate 14 CFR Part 141 flight training organization in the southeastern United States.

To ensure simulation scenarios are representative of the target population, true operational ranges representative of large, collegiate 14 CFR Part 141 flight training organizations in the United States were used to enhance the generalizability of the model. The study conducted simulation runs based on the true operational ranges to simulate the range of operating conditions possible within large, collegiate CFR Part 141 flight training organizations with varying levels of resources regarding personnel (Aviation Maintenance Technicians and Instructor Pilots), students, and aircraft.

The software utilized for the Monte Carlo simulation was Analytica Educational Professional release 4.6.1.30 by Lumina Decisions Systems. Microsoft Excel 2013 was used to process the data and to analyze and illustrate characteristics of the intermediate input data, or SPIs, generated by the algorithms in the Analytica model. Microsoft Excel 2013 was also used for post-hoc testing and analysis.

There are 22 uncontrolled inputs to the model specified as probability distributions. Three Verification Scenarios were conducted. Reliability Testing was performed with various numbers of trial runs and random number generator seed values to ensure consistent results despite the changing random number generator. To test the model reliability, the study compared the output from three different runs of the model—each using a unique seed value to ensure a different sample of random numbers for the uncontrollable input variables, which remained the same across trials. Based on the results of ANOVA output, no significant differences appeared among the different sets of results, indicating the results are statistically reliable.

The peer-reviewed research conducted by Anderson et al. (2020) validated the non-statistical model and associated equations via the use of Subject Matter Experts using a standardized expert elicitation survey questionnaire. Thus, the equations driving the predictive, safety performance decision-making tool developed in this dissertation have been previously validated through the peer-reviewed research conducted by Anderson et al. (2020).

What-if Scenarios 1, 2, 3, and 4 were conducted to demonstrate the utility of the safety performance decision-making tool for real-world use; the controllable input values used to generate the what-if scenarios within the Monte Carlo simulation model were determined based on permutational variations of the ranges of normal operating conditions for the target population − large, collegiate CFR Part 141 flight training organizations. Comparison of the four trials effectively demonstrated the utility and potential for the safety performance decision-making tool.

CFR Part 141 flight training organizations could utilize this safety decision-making tool to run what-if scenarios to understand the impact of altering the quantity of resources and expenditures, in terms of the number of AMTs available, the number of aircraft available, the number of full-time instructor pilots, and the number of active flight students and the influence these changes make on the level of operational risk for the flight department, maintenance department, and the operation overall. As the focus of this dissertation was on the process of transforming a reactive model into a safety performance decision-making tool, a 14 CFR flight training organization could utilize its own unique SPIs by determining the probability distributions of the uncontrollable input variables, further enhancing the generalizability of the safety performance decision-making tool.

CFR Part 141 flight training organizations could also use the model to determine an acceptable level of risk particular to their operation based on the manipulation of resources. Results of the model, based on the probability distributions drawn from a two-year sample of operational data, indicated the trial with the lowest risk scores occurred when there is enough personnel staffed to accommodate a moderate amount of expenditures. In both trials, the controllable input values were 35 AMTs to 50 aircraft and 200 instructor pilots to 335 flight students simulating a scenario where the level of personnel is high but expenditures are low. To reduce the level of overall risk within the organization, CFR Part 141 flight training organizations should evaluate their current ratios of AMTs to aircraft and instructor pilots to flight students to maintain an optimized level of balance and direct financial resources to accommodate an operation where the level of personnel is high compared to expenditures.

# REFERENCES

Adjekum, D. K. (2014). Safety culture perceptions in a collegiate aviation program: A systematic assessment. *Journal of Aviation Technology and Engineering, 3*(2), 44-56.

Ale, B. J. M., Bellamy, L. J., Cooke, R. M., Goossens, L. H. J., Hale, A. R., Roelen, A. L. C., & Smith, E. (2006). Towards a causal model for air transport safety – an ongoing research project. *Safety Science, 44*, 657-673.

Anderson, C. L., Aguiar, M. D., Truong, D., Friend, M. A., Williams, J., & Dickson, M. T. (2020). Development of a risk indicator score card for a large flight training department. *Safety Science*, *131*, 1-11.

Ayres, D., Schmutte, J. S., & Stanfield, J. (2017). Expect the unexpected: Risk assessment using Monte Carlo simulations. *Journal of Accountancy.* Retrieved from https://www.journalofaccountancy.com/issues/2017/nov/risk-assessment-using-monte-carlo-simulations.html

Barlow, J. (2009). Simplification: Ethical implications for simulation. Paper presented at the *18th World IMACS/MODSIM Congress.* Cairns, Australia. 13-17 July 2009.

Bastos, L. C. M. (2005). *Risk management model for on-demand Part 135 (air taxi) operators* (Unpublished master's thesis). University of Central Missouri.

Belkhouche, F. (2013). Modeling and calculating the collision risk for air vehicles. *IEEE Transactions on Vehicular Technology, 62*(5), 2031-2041.

Blair, E. (2017). Strategic safety measures: Seven key benefits. *Professional Safety, 62*(2), 32-39.

Careddu, L., Costantino, F., & Di Gravio, G. (2008). Simulation model of airport runway incursions. *Proceedings of the 20th European Modeling & Simulation Symposium,* 243-248.

Cassens, R. E. (2010). *Elements related to teaching pilots aeronautical decision making* (Unpublished master's thesis). Purdue University.

Cassens, R. E. (2015). *The effectiveness of behavior-based safety in the flight training environment* (Unpublished doctoral dissertation). Purdue University.

Chen, C. F., & Chen, S. C. (2014). Measuring the effects of safety management system practices, morality leadership and self-efficacy on pilots' safety behaviors: Safety motivation as a mediator. *Safety Science, 62,* 376–385.

Chen, W., & Jing, L. (2016). Safety performance monitoring and measurement of civil aviation unit. *Journal of Air Transport Management, 57,* 228-233.

Cox, S. J., & Cox, T. R. (1991). The structure of employee attitudes to safety: an European example. *Work & Stress, 5,* 93-106.

Dakwat, A. L., & Villani, E. (2018). System safety assessment based on STPA and model checking. *Safety Science, 109,* 130-143.

Dekker, S. (2011). *Drift into Failure: From hunting broken components to understanding complex systems.* CRC Press.

Di Gravio, G., Mancini, M., Patriarca, R., & Costantino, F. (2015). Overall safety performance of air traffic management system: Forecasting and monitoring. *Safety Science*, 72, 351-362. doi:10.1016/j.ssci.2014.10.003

Dunn, W., & Schultis, J. K. (2011). *Exploring Monte Carlo Methods* (1st ed.). Elsevier Publishing.

Dyhrberg, M. B., & Jensen, P. L. (2004). Organizations in context: Proposal for a new theoretical approach in prescriptive accident research. *Safety Science, 42*(10), 961-977.

EUROCONTROL. (2009). Reporting and assessment of safety occurrences. *ATM-ESARR2*. Retrieved from https://www.skybrary.aero/bookshelf/books/268.pdf

European Union Aviation Safety Agency. (2016). Flight data monitoring on ATR aircraft. *ATR Training Center.* Retrieved from https://www.easa.europa.eu/sites/default/files/dfu/16T0153_ATR_FDM_2016.pdf

Faghih-Roohi, S., Xie, M., & Ng, K. M. (2014). Accident risk assessment in marine transportation via Markov modelling and Markov Chain Monte Carlo simulation. *Ocean Engineering, 91*, 363-370.

Farrance, I., & Frenkel, R. (2014). Uncertainty in measurement: A review of Monte Carlo simulation using Microsoft Excel for the calculation of uncertainties through functional relationships, including uncertainties in empirically derived constants. *Clinical Biochemical Review, 35*(1), 37-61.

Federal Aviation Administration. (2016). Order 8000.369B Safety Management System. *U.S. Department of Transportation.* Retrieved from https://www.faa.gov/documentLibrary/media/Order/FAA_Order_8000.369B.pdf

Federal Aviation Administration. (2017). *Advisory Circular 141-1B.* Retrieved from https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_141-1B.pdf

Frijters, A., & Swuste, P. (2011). *How to measure safety in construction industry*. Paper presented at the National Occupational Injury Research Symposium, Washington, DC.

Gunduz, M., Birgonul, M. T., & Ozdemir, M. (2017). Fuzzy structural equation model to assess construction site safety performance. *Journal of Construction Engineering and Management*, *143*(4), 1-16.

Hacura, A., Jadamus-Hacura, M., & Kocot, A. (2001). Risk Analysis in investment appraisal based on the Monte Carlo simulation technique. *The European Physical Journal B- Condensed Matter and Complex Systems*, *20*(4), 551-553.

Hadjimichael, M. (2009). A fuzzy expert system for aviation risk assessment. *Expert Systems with Applications*, *36*, 6512-6519.

Hoyt, C. (1941). Test reliability estimated by analysis of variance. *Psychometrika, 6*(3), 153-160.

Insua, D. R., Alfaro, C., Gomez, J., Hernandez-Coronado, P., & Bernal, F. (2019). Forecasting and assessing consequences of aviation safety occurrences. *Safety Science, 111*, 243-252.

International Civil Aviation Organization (ICAO). (2013). *ICAO safety management manual* (4th ed.). ICAO.

International Civil Aviation Organization (ICAO). (2013b). *Annex 19 to the Convention on International Civil Aviation, Safety Management* (1st ed.). ICAO.

Jackman, F. (2018). Safety performance indicators. *Flight Safety Foundation.* Retrieved from https://flightsafety.org/asw-article/safety-performance-indicators/

Jacquemart, D., & Morio, J. (2013). Conflict probability estimation between aircraft with dynamic importance splitting. *Safety Science*, *51*(1), 94-100. doi:10.1016/j.ssci. 2012.05.010

Janicak, C. (2015). *Safety metrics: Tools and techniques for measuring safety performance* (3rd ed.). Bernan Press.

Jepperson, R. L. (1991). Institutions, institutional effects, and institutionalism. *The New Institutionalism in Organizational Analysis, 6,* 143-163.

Keller, J. (2015). *Unexpected transition from VFR to IMC: An examination of training protocols to mitigate pilot gaps in knowledge and performance* (Unpublished doctoral dissertation). Purdue University.

Leung, F. H., & Savithiri, R. (2009). Spotlight on focus groups. *Canadian Family Physician, 55*(2), 218-219.

Lu, C. T., Wetmore, M., & Przetak, R. (2006). Another approach to enhance airline safety: Using management safety tools. *Journal of Air Transportation, 11*(2), 113–139.

Lumina Decision Systems, Inc. (2015). *Analytica user guide – Analytica 4.6.* Retrieved from http://downloads.analytica.com/ana/UserGuide4_6_1.pdf

Makino, H., McKenna, S. A., & Wakasugi, K. (2001). Sensitivity analysis of Monte Carlo simulation results using Kolomogrov-Smirnov *d* statistic. *Proceedings of the International Association of Mathematical Geology 2001 Annual Meeting,* 1-16.

Mendonca, F. A. C., & Carney, T. Q. (2017). A safety management model for FAR 141 approved flight schools. *Journal of Aviation Technology and Engineering, 6*(2), 33-49.

Netjasov, F., Crnogorac, D., & Pavlović, G. (2019). Potential safety occurrences as indicators of air traffic management safety performance: A network based simulation model. *Transportation Research Part C, 102,* 490-508. doi:10.1016/j.trc.2019.03.026

Oswald, D., Zhang, R. P., Lingard, H., Payam, P., & Tiendung, L. (2018). The use and abuse of safety indicators in construction. *Engineering, Construction, and Architectural Management, 25*(9), 1188-1209.

Panagopoulos, I., Atkin, C., & Sikora, I. (2017). Development of performance indicators lean-sigma framework for measuring aviation system's safety performance. *Transportation Research Procedia, 22,* 35-44.

Papadopoulos, C. E., & Yeung, H. (2001). Uncertainty estimation and Monte Carlo simulation method. *Flow Measurement and Instrumentation, 12,* 291-298.

Patriarca, R., Di Gravio, G., Cioponea, R., & Licu, A. (2019). Safety intelligence: Incremental proactive risk management for holistic aviation safety performance. *Safety Science, 118*, 551-567. doi:10.1016/j.ssci.2019.05.040

Pierobon, M. (2016). Unleashing SPIs. *Flight Safety Foundation.* Retrieved from https://flightsafety.org/asw-article/unleashing-spis/

Reason, J. T. (1997). *Managing the risks of organizational accidents* (1st ed.). Ashgate. doi:10.4324/9781315543543

Reiman, T., & Pietikäinen, E. (2010). Indicators of safety culture: Selection and utilization of leading safety performance indicators. *Strålsäkerhetsmyndigheten*, 1-72.

Safety and Reliability Society. (n. d.). *Chapter 4 Monte Carlo simulation. Applied R&M manual for defence systems part D- supporting theory.* Retrieved from https://sars.org.uk/BOK/Applied%20R&M%20Manual%20for%20Defence%20Systems%20(GR-77)/p4c04.pdf

Safety Management International Collaboration Group. (2013). Measuring safety performance guidelines for service providers. *Skybrary.* Retrieved from https://www.skybrary.aero/bookshelf/books/2395.pdf

Saaty, T. L. (1980). *The Analytic Hierarchy Process.* McGraw-Hill Education.

Spall, J. C. (2003). *Introduction to stochastic search and optimization.* Wiley.

Spear, R.C., & Hornberger, G. M. (1980). Identification of critical uncertainties via generalized sensitivity analysis. *Water Research, 14*(1), 43-49.

Stemler, S., & Tsai, J. (2008). Best practices in interrater reliability three common approaches. In Osborne, J. *Best practices in quantitative methods* (pp. 29-49). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412995627

Stolzer, A. J. & Goglia, J. J. (2015). *Safety management systems in aviation* (2st ed.). Ashgate. doi:10.4324/9781351149648

Stonesifer, C. S., Calkin, D. E., Thompson, M. P., & Kaiden, J. D. (2014). Developing an aviation exposure index to inform risk-based fire management decisions. *Journal of Forestry, 112*(6), 581-590.

Stringfellow, M. V., Leveson, N. G., & Owens, B. G. (2010). Safety-driven design for software-intensive aerospace and automotive systems. *Proceedings of the IEEE International Conference on Safety, Security, and Rescue Robotics, 98*(4), 515-525.

Stroeve, S. H., Blom, H. A. P., & 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,* 133-149. doi:10.1016/j.ress.2012.07.002

Verweijen, B., & Lauche, K. (2019). How many blowouts does it take to learn the lesson? An institutional perspective on disaster development. *Safety Science, 111,* 111-118.

.

# APPENDIX A

## Permission to Conduct Research

EMBRY-RIDDLE
Aeronautical University

**Dissertation Proposal Approval Form**

Student Name: Marisa D. Aguiar ID# 2400603

Department of: School of Graduate Studies, College of Aviation

Proposed Title: Development of a Safety Performance Decision-Ma

**Committee Approval ( Please print names. **State University affiliation if other than ERAU.)**

Chairperson: Carolina L. Anderson **University:

Dept: Aeronautical Science Signature:

---

Member: Dothang Truong **University

Dept: School of Graduate Studies : Signature: Dothang Truong Digitally signed by Dothang Truong Date: 2020.06.02 14:40:32 -04'00'

---

Member: Kenneth P. Byrnes **University:

Dept: Flight Training Signature: Kenneth Byrnes Digitally signed by Kenneth Byrnes Date: 2020.06.03 08:05:53 -04'00'

---

Member: **University:

Dept: Signature:

---

External Member: Gregory S. Woo **Affiliation:

Dept: Signature: Digitally signed by Gregory S. Woo, Ph.D.

---

**Administrative Approval:**

Dothang Truong Digitally signed by Dothang Truong Date: 2020.06.02 14:40:44 -04'00'

Program Coordinator

Steven Hampton Digitally signed by Steven Hampton DN: cn=Steven Hampton, o=Embry-Riddle Aeronautical University, ou=Associate Dean for Research and Graduate Studies, email=hamptons@erau.edu, c=US Date: 2020.06.11 11:08:02 -04'00'

Department Chair

Alan Stolzer Digitally signed by Alan Stolzer Date: 2020.06.11 15:10:40 -04'00'

Dean

EMBRY-RIDDLE
Aeronautical University

Your research may require regulatory oversight. Approval from a regulatory oversight committee may be necessary before any research is conducted.

Does the proposal involve research with any human/animal subjects? YES ☐ NO ☑

If "Yes" Indicate date approved and IRB NUMBER: ____________________

Project Number ____________________ Date ____________________

If the IRB determines that your project DOES NOT NEED approval, please provide appropriate documentation from that office.

I, Marisa D. Aguiar, affirm that the research for my doctorate degree will be conducted in agreement with ethical standards at Embry-Riddle Aeronautical University and that my dissertation will be original. I will provide unambiguous attribution for the thought and the words of other scholars eventually appearing in the work. I understand that failure to provide clear credit in this way can result in severe penalties, including separation from the University and revocation of a degree.

I also understand that regulatory oversight for my research may be required and that I should contact the Institutional Review Board for assistance.

Marisa Aguiar Digitally signed by Marisa Aguiar
Date: 2020.06.02 09:41:47 -04'00'

Student Signature

# APPENDIX B

## Tables

Table B1

*Verification Scenario 2 Comparison Output*

| | | Model Output | | | Raw Data | |
|---|---|---|---|---|---|---|
| SPI | Variable | Min Value | Max Value | Mean | Lower Limit | Higher Limit |
| 1-MX | Logistical delay time | 100 | 290 | 212 | 100 | 310 |
| 2-MX | *AMTs Available** | *14* | *14* | | 14 | 35 |
| | Fleet flight time | 4006 | 13300 | 7602 | 4000 | 13500 |
| 3-MX | Percent of AC available | 70 | 100 | 85 | 70 | 100 |
| | *Total AC in fleet** | *50* | *50* | | 50 | 82 |
| 4-MX | Fleet flight time | 4006 | 13300 | 7602 | 4000 | 13500 |
| | Total MX orders processed | 100 | 800 | 532 | 100 | 1200 |
| 5-MX | Unscheduled MX orders <$10K | 425 | 582 | 500 | 300 | 1000 |
| | FAA occurrences | 1 | 49 | 6 | 0 | 40 |
| | Fleet flight time | 4006 | 13300 | 7602 | 4000 | 13500 |
| 6-MX | Total MX orders processed | 100 | 799 | 532 | 100 | 1200 |
| | AC dispatched w/ MX error | 0 | 1 | 0.05 | 0 | 2 |
| 1-FLT | Unstable approaches | 6 | 767 | 156 | 0 | 946 |
| | Flap overspeeds | 0 | 3 | 0.54 | 0 | 3 |
| | G exceedances | 0 | 3 | 0.42 | 0 | 3 |
| | Tail strikes | 0 | 7 | 3 | 0 | 10 |
| | RPM overspeeds | 0 | 1 | 0.5 | 0 | 3 |
| | Hard landings | 1 | 5 | 1.7 | 0 | 7 |
| | Fleet flight time | 4006 | 13300 | 7602 | 4000 | 13500 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2-FLT | Annual SC survey results | 4.6 | 4.6 | 4.6 | 1 | 5.76 |
| 3-FLT | Traffic conflicts | 2 | 14 | 9 | 0 | 18 |
| | Fleet flight time | 4006 | 13300 | 7602 | 4000 | 13500 |
| 4-FLT | *Full time Ips** | *100* | *100* | | *100* | *200* |
| | *Active flight students** | *335* | *335* | | *335* | *1300* |
| 5-FLT | Months as an IP | 10 | 10 | 10 | 0 | 12 |
| 6-FLT | Event reports | 39 | 108 | 67 | 25 | 150 |
| Damage & Related Impact | FAA incident reports | 0 | 1 | 0.12 | 0 | 3 |
| | Unsched MX > $10K | 0 | 3 | 0.9 | 0 | 3 |
| | NTSB reports | 0 | 1 | 0.335 | 0 | 3 |
| | Fleet flight time | 4006 | 13300 | 7602 | 4000 | 13500 |

Table B2

*Verification Scenario 2: SPI Comparison Output*

| | Model Output | | | Raw Data | |
|---|---|---|---|---|---|
| SPI | Min | Max | Mean | Lower Limit | Higher Limit |
| SPI-1 MX | 1.7125 | 4.828 | 3.5321 | 1.6667 | 5.1667 |
| SPI-2 MX | 0.0010 | 0.0035 | 0.0019 | 0 | 0.00875 |
| SPI-3 MX | 1.4 | 2 | 1.698 | 0.8537 | 2 |
| SPI-4 MX | 0.0121 | 0.1888 | 0.073 | 0.0074 | 0.3 |
| SPI-5 MX | 0.0369 | 0.1393 | 0.0697 | 0 | 0.26 |
| SPI-6 MX | 0 | 0.005 | 0.0001 | 0 | 0.02 |
| | | | | | |
| SPI-1 FLT | 0.0011 | 0.1737 | 0.0224 | 0 | 0.0302 |
| SPI-2 FLT | 4.6 | 4.6 | 4.6 | 1 | 5.76 |
| SPI-3 FLT | 0.0002 | 0.0039 | 0.0013 | 0 | 0.0045 |
| SPI-4 FLT | 4.9348 | 4.9348 | 4.9348 | 2 | 8 |
| SPI-5 FLT | 10 | 10 | 10 | 0 | 36 |
| SPI-6 FLT | 35 | 103 | 67 | 0 | 200 |

Table B3

*Verification Scenario 2: Standardized SPI Comparison Output*

| | Model Output | | | Raw Data | |
|---|---|---|---|---|---|
| SPI | Min | Max | Mean | Lower Limit | Higher Limit |
| SPI-1 MX | 0 | 5 | 2.878 | 0.8883 | 4.8501 |
| SPI-2 MX | 3.411 | 5 | 4.433 | 2.1639 | 4.2139 |
| SPI-3 MX | 2.384 | 4.999 | 3.674 | 1.1229 | 3.4422 |
| SPI-4 MX | 0 | 5 | 1.724 | 0.2209 | 1.3921 |
| SPI-5 MX | 0 | 5 | 1.601 | 0.3281 | 1.7218 |
| SPI-6 MX | 0 | 5 | 0.0974 | 0 | 0.4562 |
| SPI-1 FLT | 0 | 5 | 0.5917 | 0.0189 | 2.0034 |
| SPI-2 FLT | 1.218 | 1.218 | 1.218 | 1.2185 | 1.2185 |
| SPI-3 FLT | 0 | 5 | 1.702 | 0.3660 | 2.0590 |
| SPI-4 FLT | 1.125 | 1.125 | 1.125 | 0.4610 | 3.9552 |
| SPI-5 FLT | 3.611 | 3.611 | 3.611 | 3.6111 | 3.6111 |
| SPI-6 FLT | 0 | 5 | 2.382 | 0.8000 | 4.0250 |
| Damage & Related Impact | 0 | 0.4197 | 0.084 | 0 | 0.0002 |

Table B4

*Verification Scenario 3 Comparison Output*

| | | Model Output | | | Raw Data | |
|---|---|---|---|---|---|---|
| SPI | Variable | Min Value | Max Value | Mean | Lower Limit | Higher Limit |
| 1-MX | Logistical delay time | 100 | 290 | 212 | 100 | 310 |
| 2-MX | *AMTs Available** | *35* | *35* | | 14 | 35 |
| | Fleet flight time | 4006 | 13300 | 7606 | 4000 | 13500 |
| 3-MX | Percent of AC available | 70 | 100 | 85 | 70 | 100 |
| | *Total AC in fleet** | *82* | *82* | | 50 | 82 |
| 4-MX | Fleet flight time | 4006 | 13300 | 7602 | 4000 | 13500 |
| | Total MX orders processed | 100 | 800 | 532 | 100 | 1200 |
| 5-MX | Unscheduled MX orders <$10K | 425 | 582 | 500 | 300 | 1000 |
| | FAA occurrences | 1 | 49 | 6 | 0 | 40 |
| | Fleet flight time | 4006 | 13300 | 7606 | 4000 | 13500 |
| 6-MX | Total MX orders processed | 100 | 799 | 532 | 100 | 1200 |
| | AC dispatched w/ MX error | 0 | 1 | 0.05 | 0 | 2 |
| 1-FLT | Unstable approaches | 6 | 767 | 156 | 0 | 946 |
| | Flap overspeeds | 0 | 3 | 0.54 | 0 | 3 |
| | G exceedances | 0 | 3 | 0.42 | 0 | 3 |
| | Tail strikes | 0 | 7 | 3 | 0 | 10 |
| | RPM overspeeds | 0 | 1 | 0.5 | 0 | 3 |
| | Hard landings | 1 | 5 | 1.7 | 0 | 7 |
| | Fleet flight time | 4006 | 13300 | 7606 | 4000 | 13500 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 2-FLT | Annual SC survey results | 4.6 | 4.6 | 4.6 | 1 | 5.76 |
| 3-FLT | Traffic conflicts | 2 | 14 | 9 | 0 | 18 |
| | Fleet flight time | 4006 | 13300 | 7606 | 4000 | 13500 |
| 4-FLT | *Full time Ips** | *200* | *200* | | *100* | *200* |
| | *Active flight students** | *1300* | *1300* | | *335* | *1300* |
| 5-FLT | Months as an IP | 10 | 10 | 10 | 0 | 12 |
| 6-FLT | Event reports | 39 | 108 | 67 | 25 | 150 |
| Damage & Related Impact | FAA incident reports | 0 | 1 | 0.12 | 0 | 3 |
| | Unsched MX > $10K | 0 | 3 | 0.9 | 0 | 3 |
| | NTSB reports | 0 | 1 | 0.335 | 0 | 3 |
| | Fleet flight time | 4006 | 13300 | 7606 | 4000 | 13500 |

Table B5

*Verification Scenario 3: SPI Comparison Output*

| | Model Output | | | Raw Data | |
|---|---|---|---|---|---|
| SPI | Min | Max | Mean | Lower Limit | Higher Limit |
| SPI-1 MX | 1.7125 | 4.8741 | 3.526 | 1.6667 | 5.1667 |
| SPI-2 MX | 0.0036 | 0.0087 | 0.0048 | 0 | 0.00875 |
| SPI-3 MX | 0.8537 | 1.219 | 1.035 | 0.8537 | 2 |
| SPI-4 MX | 0.0121 | 0.1888 | 0.0734 | 0.0074 | 0.3 |
| SPI-5 MX | 0.0369 | 0.1393 | 0.0697 | 0 | 0.26 |
| SPI-6 MX | 0 | 0.0053 | 0.0001 | 0 | 0.02 |
| | | | | | |
| SPI-1 FLT | 0.0001 | 0.1799 | 0.6167 | 0 | 0.0302 |
| SPI-2 FLT | 4.6 | 4.6 | 4.6 | 1 | 5.76 |
| SPI-3 FLT | 0.0002 | 0.0035 | 0.0013 | 0 | 0.0045 |
| SPI-4 FLT | 6.5 | 6.5 | 6.5 | 2 | 8 |
| SPI-5 FLT | 10 | 10 | 10 | 0 | 36 |
| SPI-6 FLT | 35 | 103 | 67 | 0 | 200 |

Table B6

*Verification Scenario 3: Standardized SPI Comparison Output*

| | Model Output | | | Raw Data | |
|---|---|---|---|---|---|
| SPI | Min | Max | Mean | Lower Limit | Higher Limit |
| SPI-1 MX | 0 | 5 | 2.878 | 0.8883 | 4.8501 |
| SPI-2 MX | 0 | 3.973 | 2.554 | 2.1639 | 4.2139 |
| SPI-3 MX | 0.0007 | 1.595 | 0.7873 | 1.1229 | 3.4422 |
| SPI-4 MX | 0 | 5 | 1.724 | 0.2209 | 1.3921 |
| SPI-5 MX | 0 | 5 | 1.601 | 0.3281 | 1.7218 |
| SPI-6 MX | 0 | 5 | 0.0974 | 0 | 0.4562 |
| SPI-1 FLT | 0 | 5 | 0.6167 | 0.0189 | 2.0034 |
| SPI-2 FLT | 1.218 | 1.218 | 1.218 | 1.2185 | 1.2185 |
| SPI-3 FLT | 0.0990 | 4.807 | 1.676 | 0.3660 | 2.0590 |
| SPI-4 FLT | 3.75 | 3.75 | 3.75 | 0.4610 | 3.9552 |
| SPI-5 FLT | 3.611 | 3.611 | 3.611 | 3.6111 | 3.6111 |
| SPI-6 FLT | 0 | 5 | 2.054 | 0.8000 | 4.0250 |
| Damage & Related Impact | 0 | 0.2947 | 0.0502 | 0 | 0.0002 |

# APPENDIX C

## Figures


*Figure C1*. Probability density distribution of the Maintenance Score in Verification Scenario 2.

.

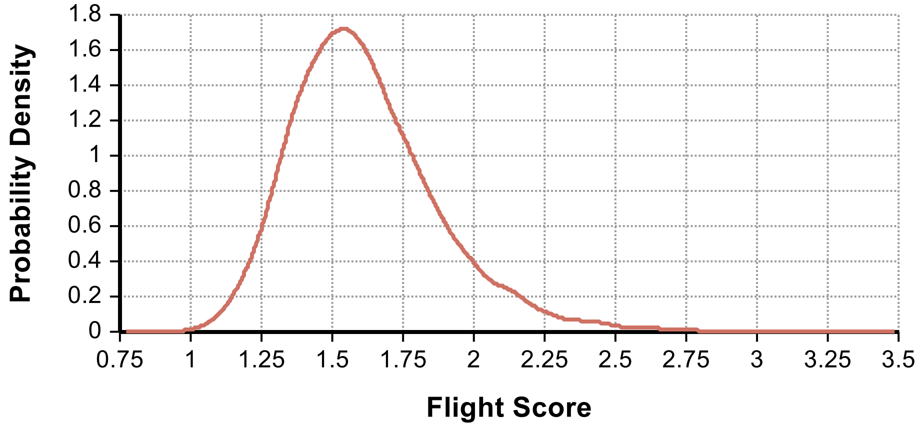



*Figure C2.* Probability density distribution of the Flight Score in Verification Scenario 2.

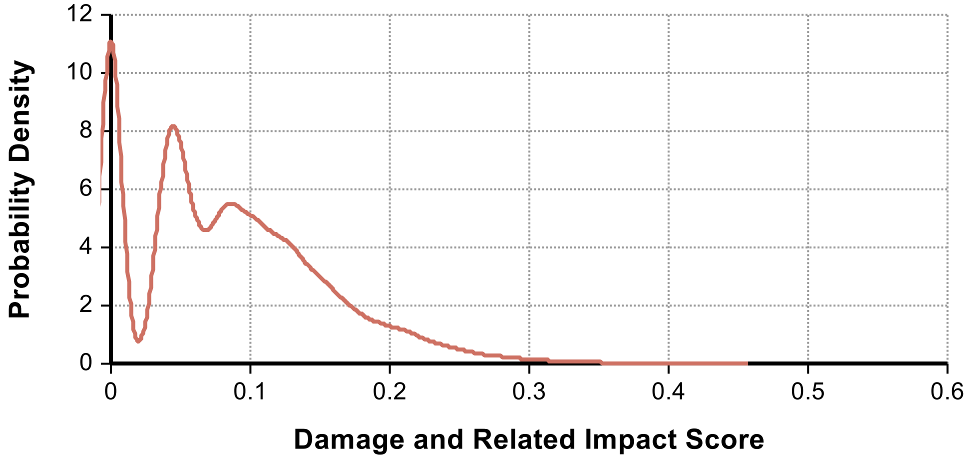



*Figure C3*. Probability density distribution of the Damage & Related Impact Score in Verification Scenario 2.

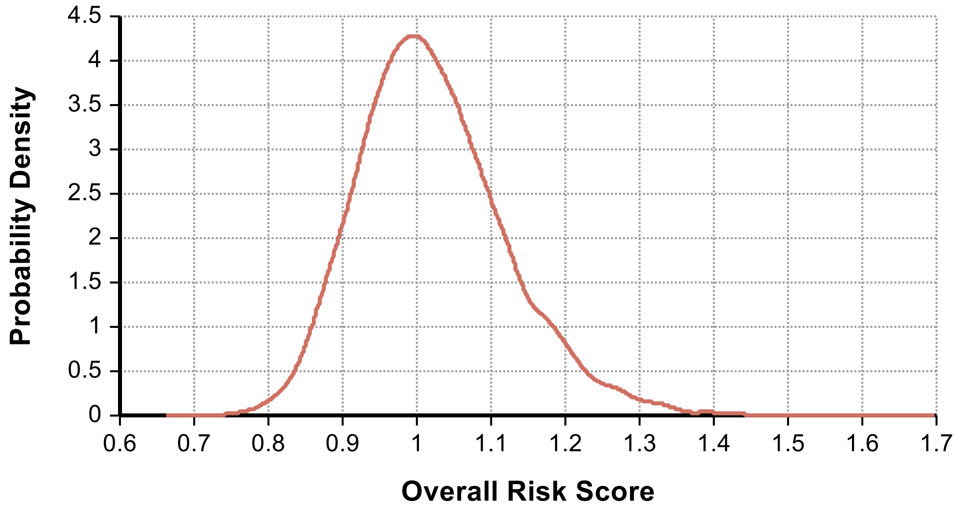



*Figure C4.* Probability density distribution of the Overall Risk Score in Verification Scenario 2.

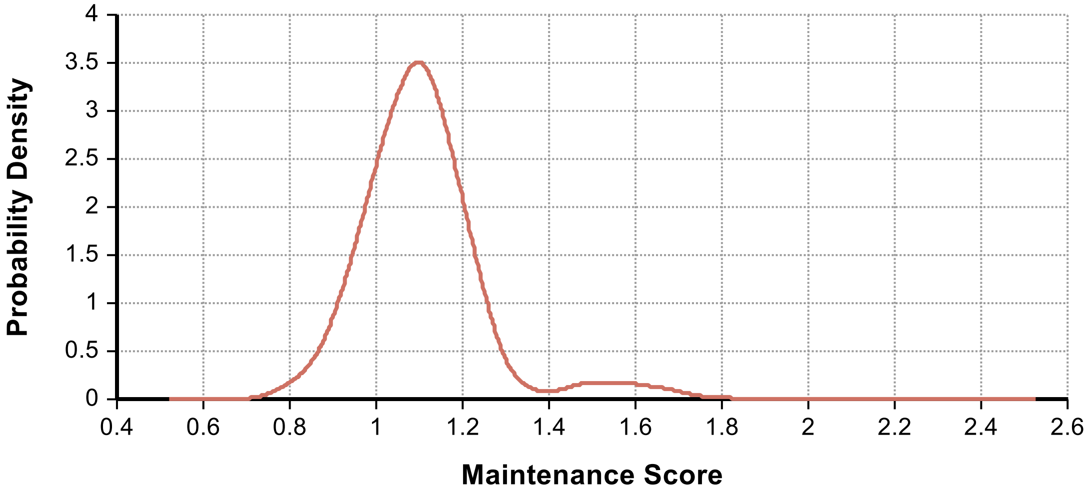



*Figure C5.* Probability density distribution of the Maintenance Score in Verification Scenario 3.

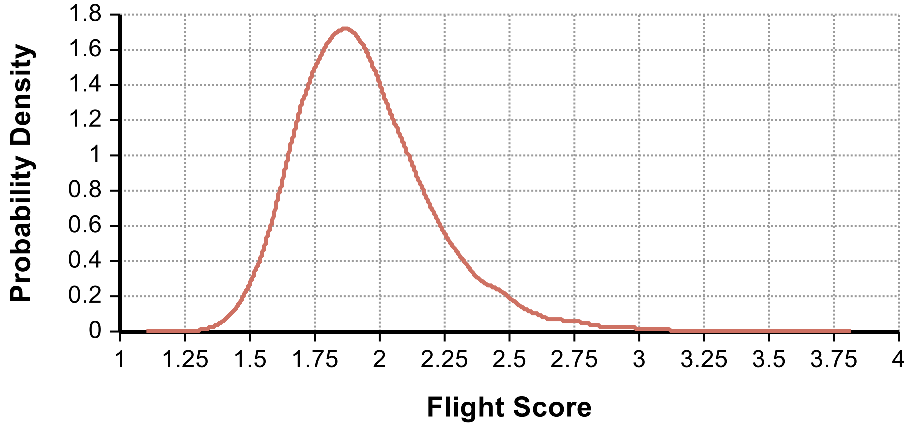



*Figure C6.* Probability density distribution of the Flight Score in Verification Scenario 3.

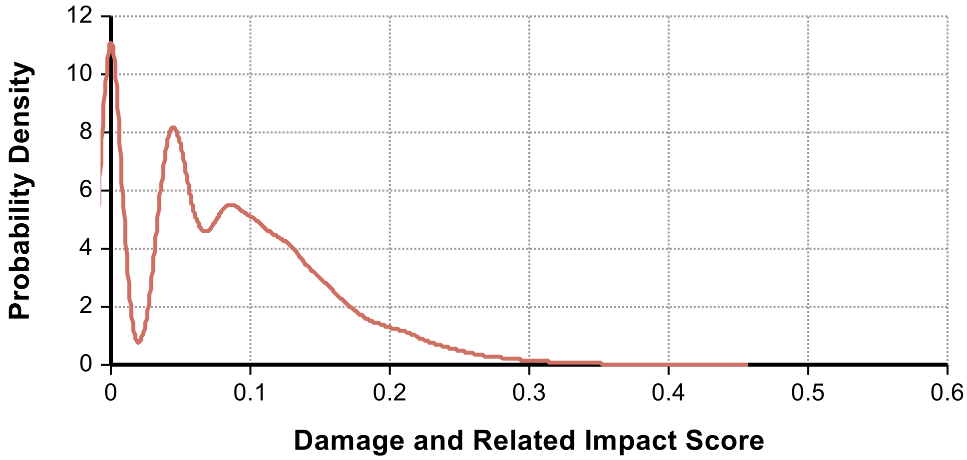



*Figure C7.* Probability density distribution of the Damage & Related Impact Score in Verification Scenario 3.

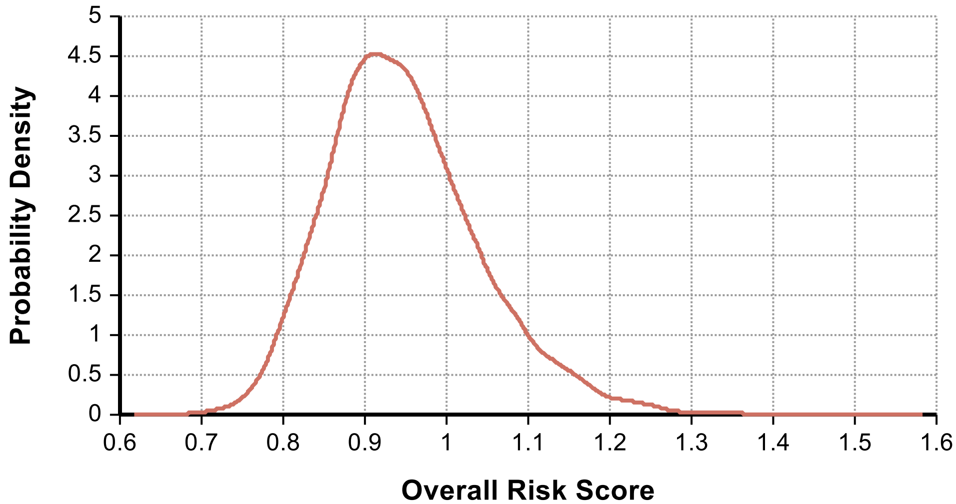



*Figure C8.* Probability density distribution of the Overall Risk Score in Verification Scenario 3.