Collegiate Aviation Review International

Volume 41 | Issue 2 Proceedings of the 2023 UAA Annual Conference, Article #3

12-14-2023

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

Carolina Anderson
Embry-Riddle Aeronautical University

Marisa Aguiar
Purdue University Global

The purpose of the research was to transform a non-statistical risk score model composed of 12 Safety Performance Indicators (SPIs) 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. These risk score outputs provide a keen insight into the overall level of risk within the organization.

Recommended Citation:
Anderson, C. & Aguiar, M. (2023). Development of a safety performance decision-making tool for flight training organizations. *Collegiate Aviation Review International, 41*(2), 208-216. Retrieved from http://ojs.library.okstate.edu/osu/index.php/CARI/article/view/9651/8538

Introduction

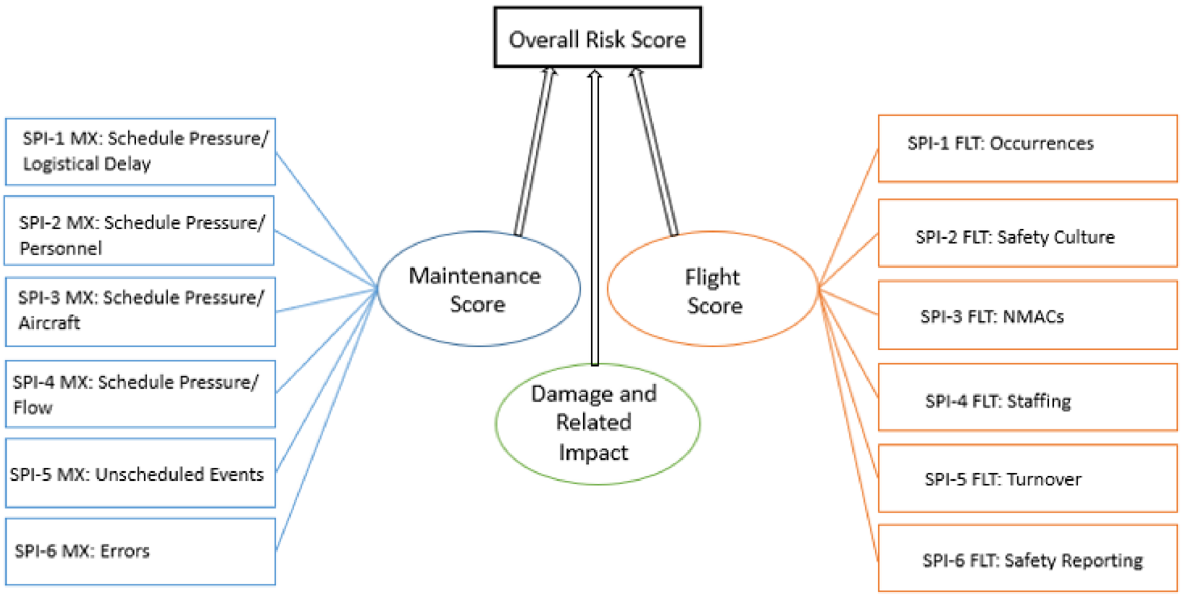
With the introduction and requirement of a Safety Management System (SMS) in aviation, the focus is shifting from traditional forms of reactive data collection and analysis toward approaches and techniques that bolster and improve the effectiveness of the organization's SMS. A vital portion of this process 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 occurs (Pierobon, 2016).

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 flight schools, 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. These risk score outputs provide a keen insight into the overall level of risk within the organization.

Theoretical Framework

The theoretical framework driving the research was founded upon a model developed by Anderson et al. (2020); a sequential, mixed-method design study was conducted, including a qualitative data collection and analysis phase, followed by a quantitative data collection and analysis phase. Subject Matter Experts (SMEs) in 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 data, see Anderson et al. (2020). 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 (see Figure 2).

Figure 2.
Diagram of the non-statistical model developed by Anderson et al. (2020) composed of SPIs and associated indicators.



Methodology

Monte Carlo simulation methodologies were used to build a safety decision-making tool based on SPIs determined by Anderson et al. (2020) to represent flight risk within 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 used the quantitative method to convert a non-statistical model into a safety decision-making tool, utilizing Monte Carlo simulation; this simulation will allow 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.

Population and Sample

The target population to which the model generalizes is large, collegiate 14 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 (FAA, 2017). The sampling data used to determine the probability distributions of the uncontrollable input variables within the model consisted of two years of operational data from both flight and maintenance operations dating from September 2017 to September 2019 for a flight training organization in the United States.

The study conducted simulation runs based on the true operational ranges specified below to simulate the range of operating conditions possible within a flight training organization with varying levels of resources with respect to personnel (Aviation Maintenance Technicians and Instructor Pilots), students, and aircraft:

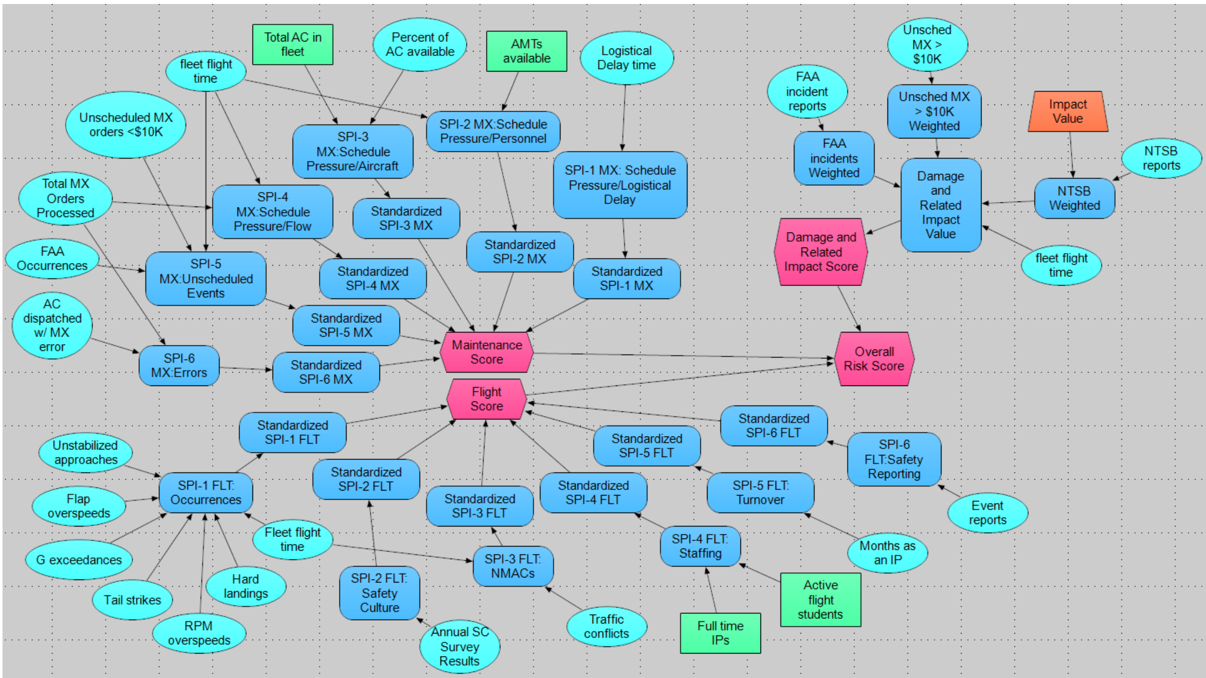
- Aviation Maintenance Technicians available: 14-35
- Aircraft available: 50-82
- Full-time Instructor Pilots: 100-200
- Active Flight Students: 335-1300

These ranges were selected because they reflect the higher and lower operational limits of the sample data drawn for the organization. The model could easily be adapted for use in any flight training organization with flight data acquisition abilities and an operational SMS.

Design of the Mathematical Model

Figure 2 depicts the structural definition of the model used for the Montecarlo simulation. 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 and depict calculation nodes producing the results of the model. The orange trapezoid represents a value that is input as a constant. The impact value was input into the model as a constant value of 1, indicating no damage or injuries incurred was selected for the purpose of this study. The pink hexagons represent the risk score output variables.

Figure 2.
Structural definition of the model.



Data Analysis Approach

Various trials of the model were completed using different random number generator seed values to confirm the output of the simulation, which 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 assess the model's reliability (Hoyt, 1941).

The study simulated 10,000 trials for a given scenario with manipulated controllable input values. 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.

Results

Validity Testing

Three verification scenarios of the model were conducted to test validity. The shape of the distributions of the uncontrollable input variables from all the verification trials is the same as the distributions drawn from the raw data sample (see Appendix A).

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 flight training organizations. 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.

Each trial was computed using the specified controllable input variables, capturing the output 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 (see Table 1).

Table 1
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.

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. 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. (See Table 2).

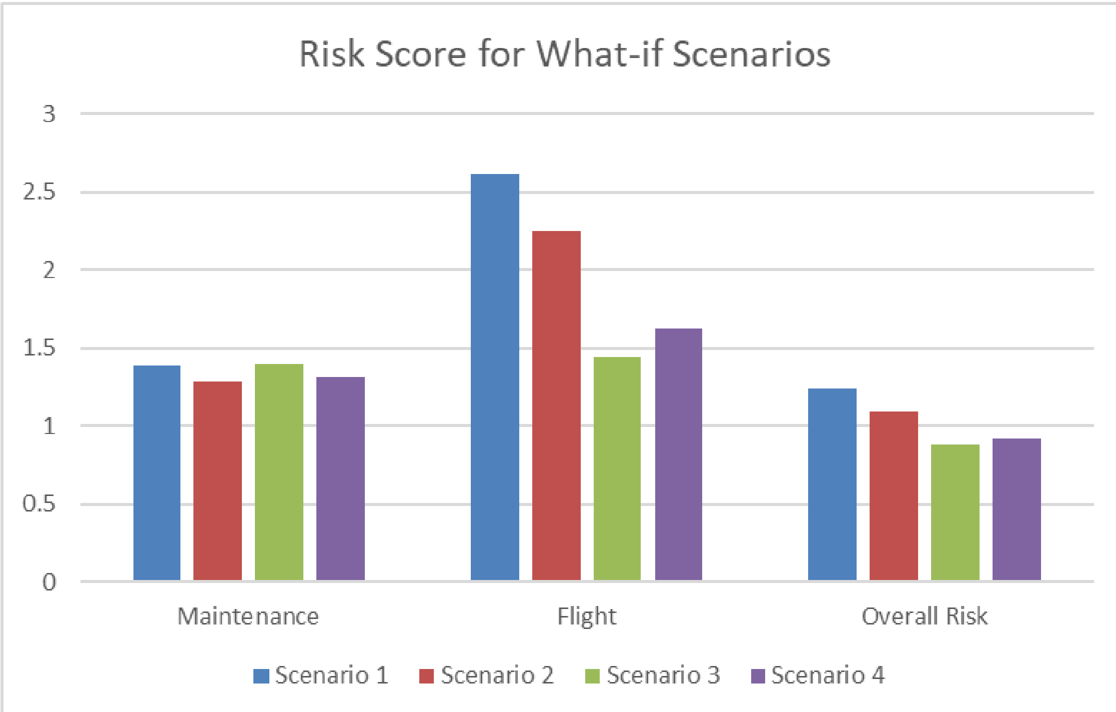
What-if Scenario 2 was conducted with the intent of simulating a scenario similar to What-if Scenario 1; however, in What-if Scenario 2, the number of personnel, including AMTs and instructor pilots, was increased from 14 AMTs to 22 and 100 instructor pilots to 138. The expenditures, consisting of aircraft and active flight students, remained high. 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, reducing 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 (See Table 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 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 (See Table 2).

Finally, What-if Scenario 4 was conducted with the intent of simulating a scenario similar to What-if Scenario 3; however, the 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. 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 (See Table 2).

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. 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 constant throughout; thus, no visual comparisons were made. 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 number of expenditures remains low. Although intuitive, this demonstrates the real-world utility of the model (see Figure 3).

Figure 3.
Maintenance, Flight, and Overall Risk Score What-if Scenario Comparison Chart



Discussion and Conclusions

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, the number of aircraft and students was high. The lowest risk score for flight and lowest overall risk occurred in What-if Scenario 3, where the level of personnel was high, and the number of aircraft and students was low.

Changes to the controllable input variables are reflected by variations to the risk score outputs, demonstrating the utility and predictive potential of 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 this flight training organization should strive to maintain an appropriate balance of high personnel to low expenditures to maintain the optimum level of operational safety.

References

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

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

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

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.

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

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