National Training Aircraft Symposium (NTAS) 2022 - Bridging the Gap

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

Marisa D. Aguiar PhD
Embry-Riddle Aeronautical University, aguiarm@my.erau.edu

Carolina L. Anderson PhD
Embry-Riddle Aeronautical University, lenzc@erau.edu

Aguiar, Marisa D. PhD and Anderson, Carolina L. PhD, "Development of a Safety Performance Decision-Making Tool for Flight Training Organizations" (2023). *National Training Aircraft Symposium (NTAS)*. 20.
https://commons.erau.edu/ntas/2022/presentation/20

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

MARISA D. AGUIAR, PH.D.

CAROLINA L. ANDERSON, PH.D.

Project Phases

Phase 1: SPI Selection - Anderson, Aguiar, Truong, Friend, Williams, and Dickson (2020)

- 12 SPIs selected (6 for flight, 6 for Maintenance)
- Collected 2 years' worth of data
- External and internal SMEs provided feedback via expert elicitation

Phase 2: Distributions and Forecasting

- Determine data distributions
- Forecast using Monte Carlo simulation
- Run what-if scenarios

Phase 3: Implementation-Future Research

- Collect data in a semi-autonomous way via reports
- Find the best way to present the results
- Make necessary changes based on feedback

Phase 1

Objective

- Create and validate a single quantitative indicator of flight safety performance for a Flight Department, to be calculated on a periodic basis.

Purpose Statement

- Increase accuracy of the Risk Management and Safety Assurance components of a Flight Department's SMS by applying scientific principles from data analytics and safety theory
- to help justify funding of new staff positions, technology, or other safety-related initiatives
- to have the ability to run what-if scenarios to assess how changes to input variables may affect overall safety

Significance

- Deficit of validated models of flight safety performance for large flight training operations
- To understand variables contributing to flight safety for large flight training operations

Phase 2

Research Problem

- Safety monitoring based on relevant, domain-specific SPIs is still a reactive approach to safety monitoring
  - Provides a one-size-fits-all approach to safety monitoring

Purpose Statement

- Create and validate a safety performance decision-making tool transforming 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

Significance

- Improves accuracy and robustness of the SMS
- Administrative utility
- Adaptable for any CFR Part 141 with data acquisition capabilities and an active SMS
- Enhances the understanding of the relationship between resource optimization and operational safety
- Fills a gap by providing a validated safety decision-making tool specific to CFR Part 141 operations

Delimitations and Assumptions

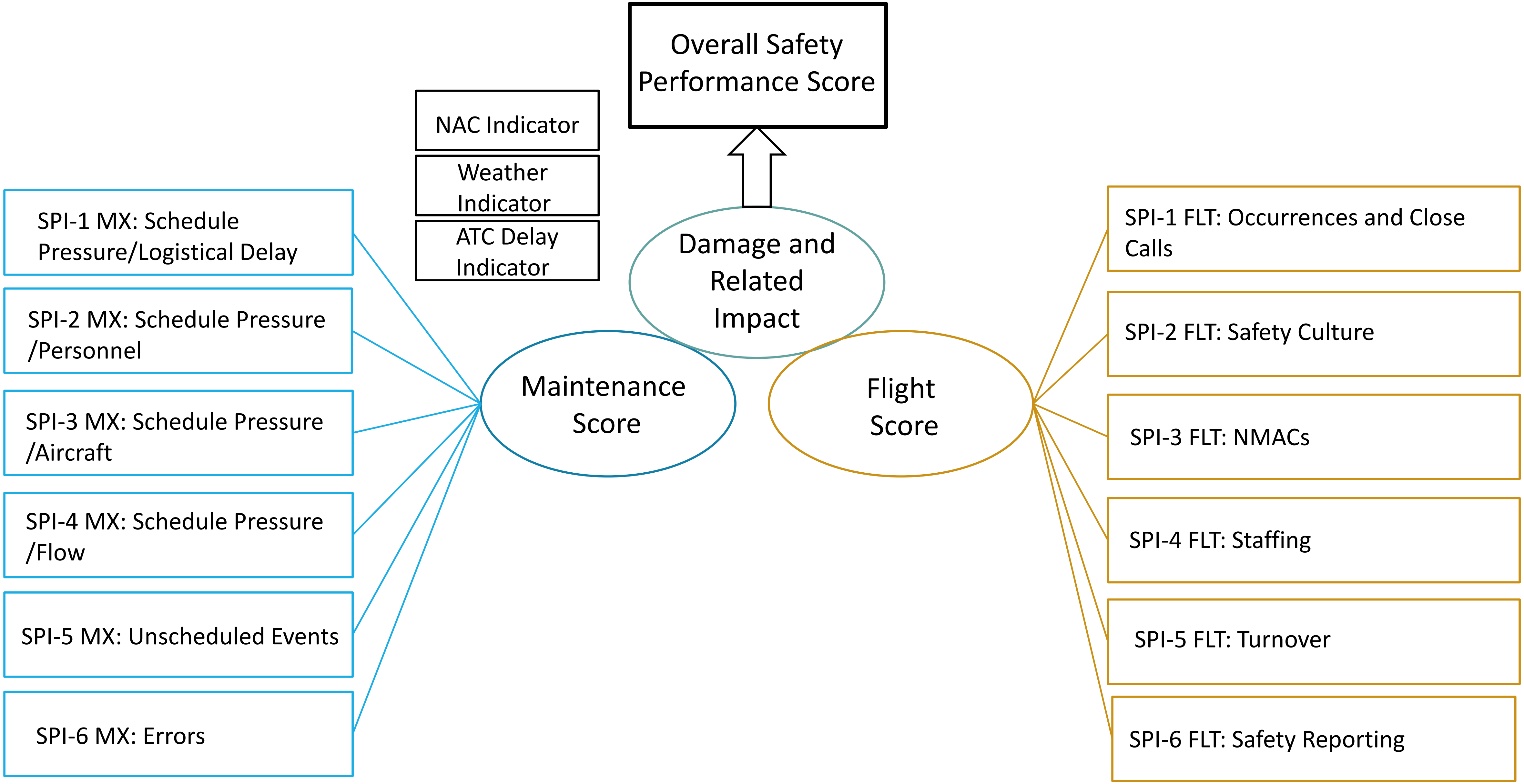
Both models are designed to measure the potential for increased or decreased flight risk for large, collegiate flight training programs within the United States using readily available flight department data

The model does **NOT** measure:

- Occupational risks
- Cases of gross negligence
- Security threats
- Human performance state measurements

A *large* CFR Part 141 assumes the following operational criteria:

- ○ At least 500 flight training students
- ○ 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 (SMS)



Overall Safety Performance Score

NAC Indicator

Weather Indicator

ATC Delay Indicator

Damage and Related Impact

Maintenance Score

Flight Score

SPI-1 MX: Schedule Pressure/Logistical Delay

SPI-2 MX: Schedule Pressure /Personnel

SPI-3 MX: Schedule Pressure /Aircraft

SPI-4 MX: Schedule Pressure /Flow

SPI-5 MX: Unscheduled Events

SPI-6 MX: Errors

SPI-1 FLT: Occurrences and Close Calls

SPI-2 FLT: Safety Culture

SPI-3 FLT: NMACs

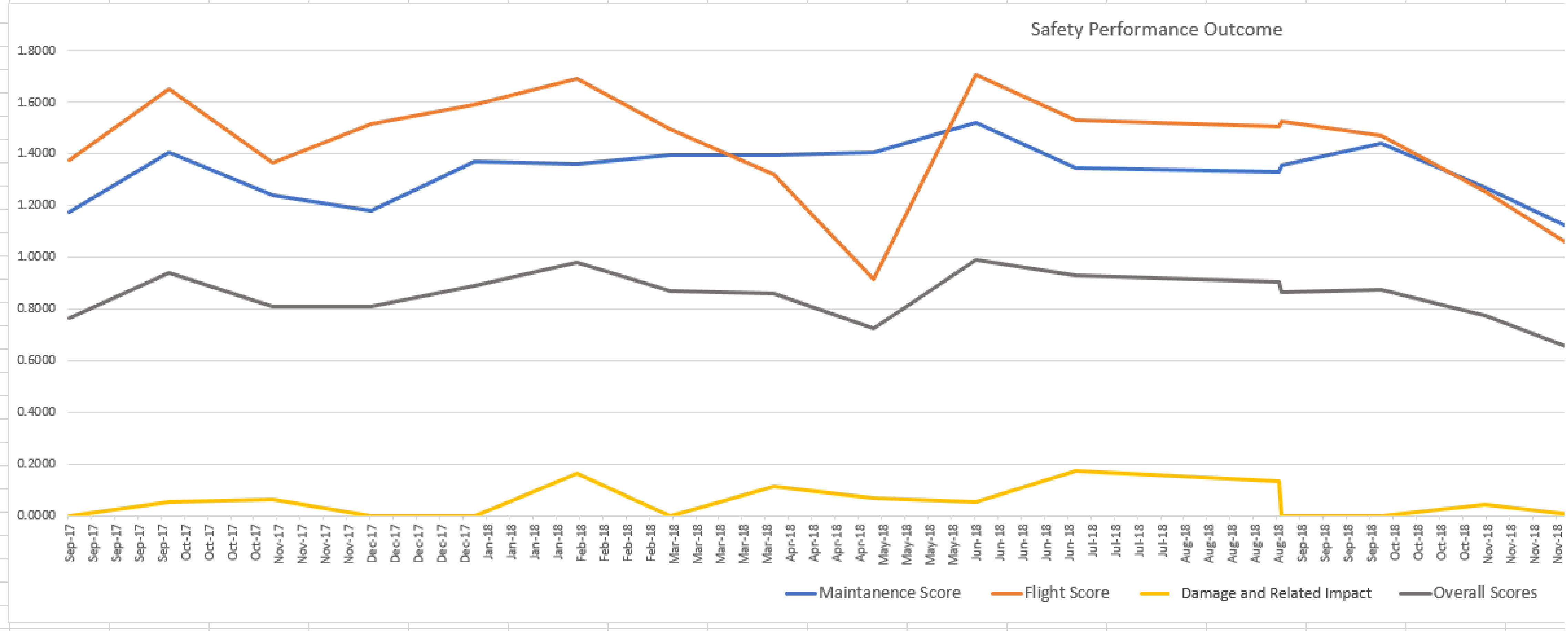
SPI-4 FLT: Staffing

SPI-5 FLT: Turnover

SPI-6 FLT: Safety Reporting

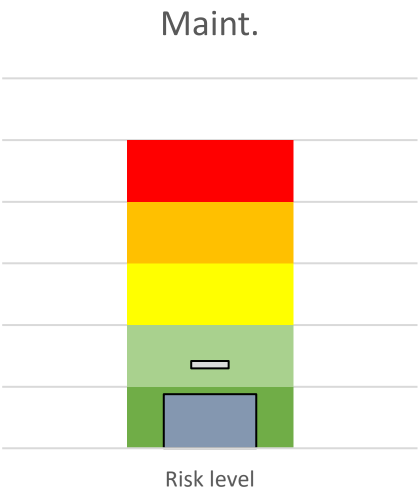
Safety Performance Outcome

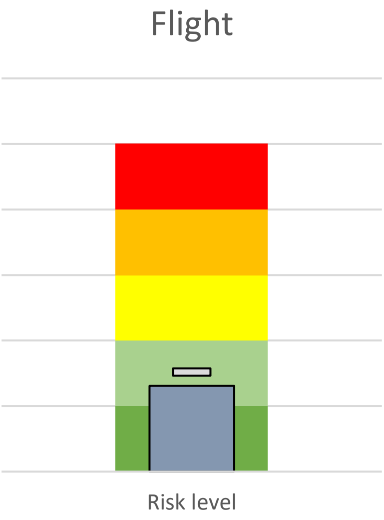
All data shown are notional. No actual data is provided and the information here does not represent actual performance of any organization or department.



Implementation

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





Research Questions

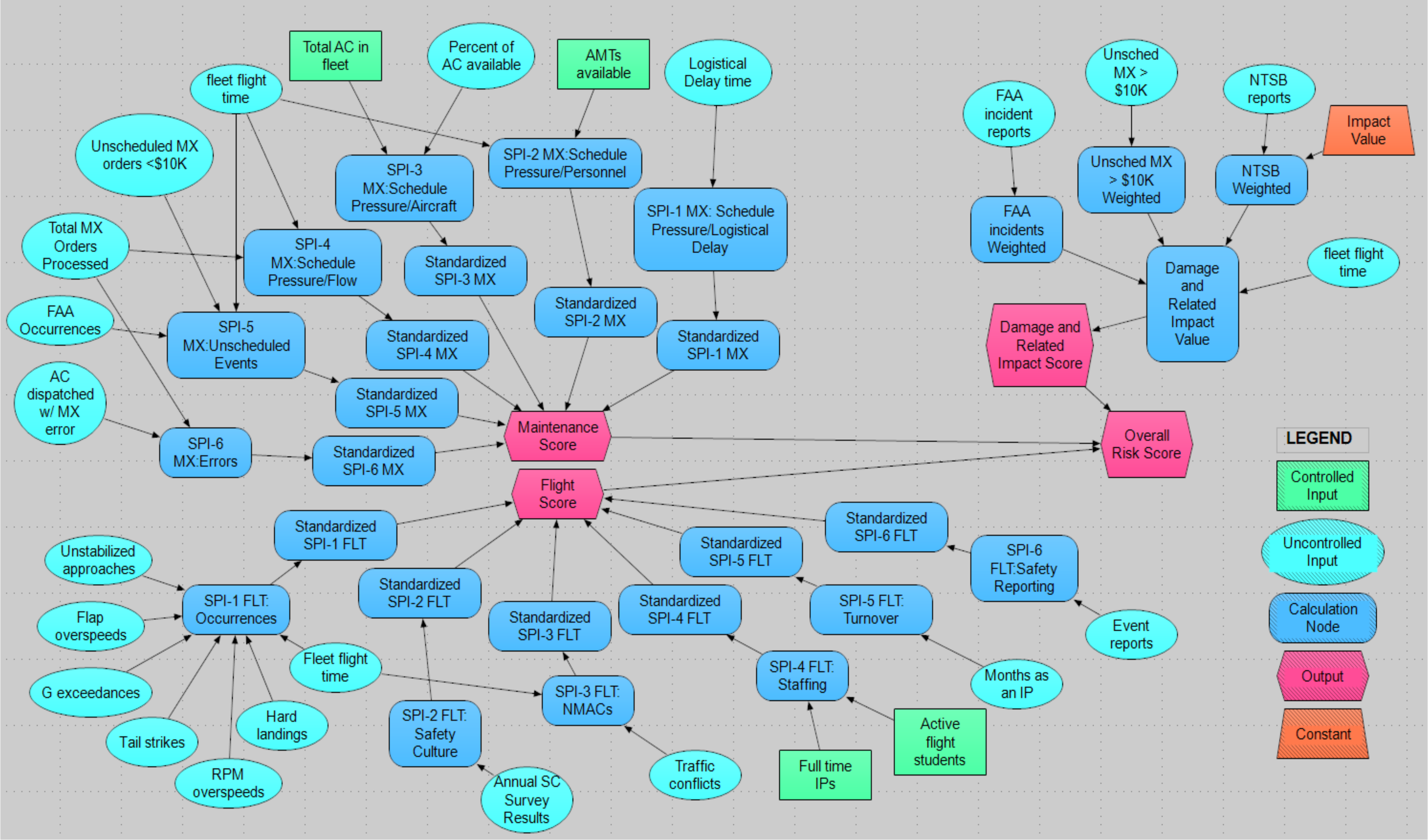
①

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 score?

Research Methodology



Quantitative Research Design

Population: large, collegiate CFR Part 141 flight training organizations within the U.S.

- Sample: Sept. 2017-2019 operational flight data from a flight training organization in the southeastern U.S.

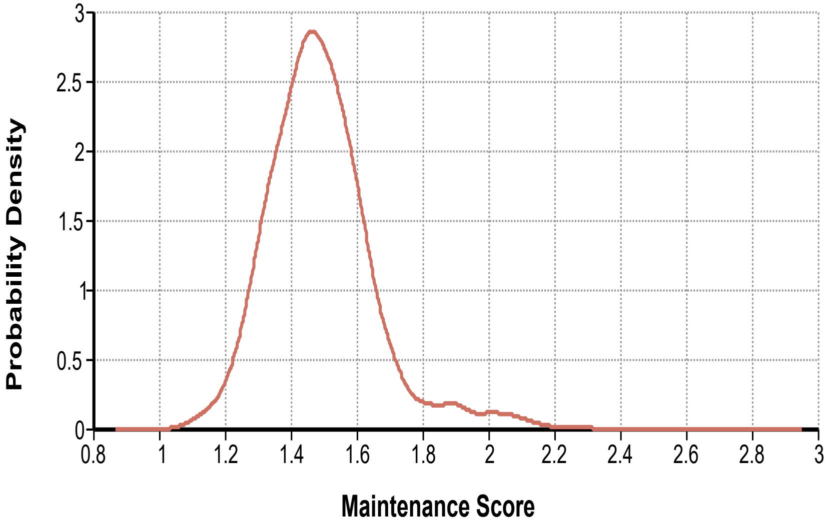
Simulation Scenarios

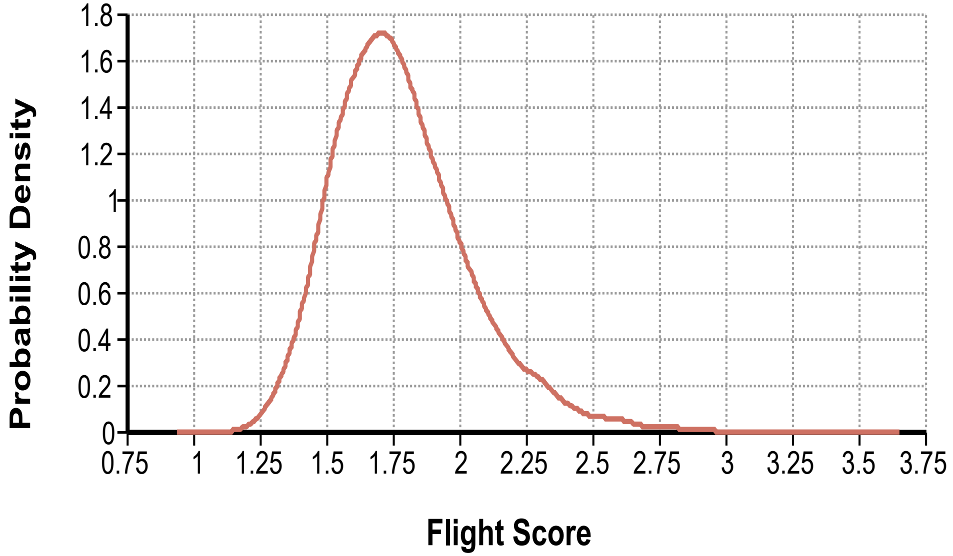
Controllable Input	Range
AMTs available	14-35
Aircraft available	50-82
Full-time instructor pilots (Ips)	100-200
Active flight students	335-1300

Data Analysis Approach
Validity Assessment
Reliability Assessment: ANOVA
Data Analysis

Results: Verification Scenarios

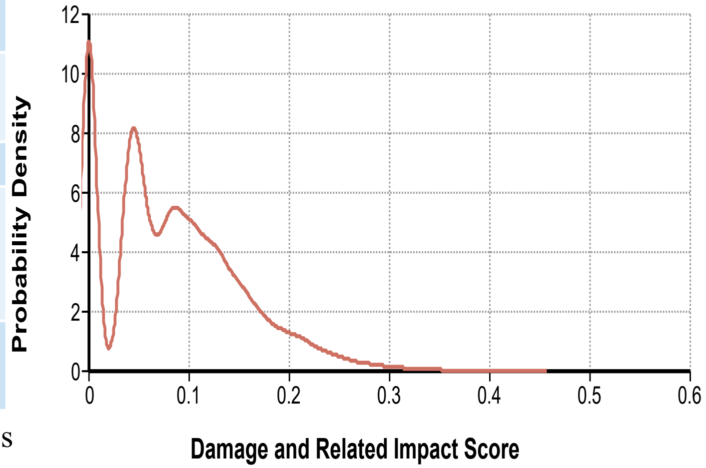
Controllable Input	Verification Scenario 1
AMTs available	22
Aircraft available	56
Active flight students	138
Full-time instructor pilots (Ips)	681

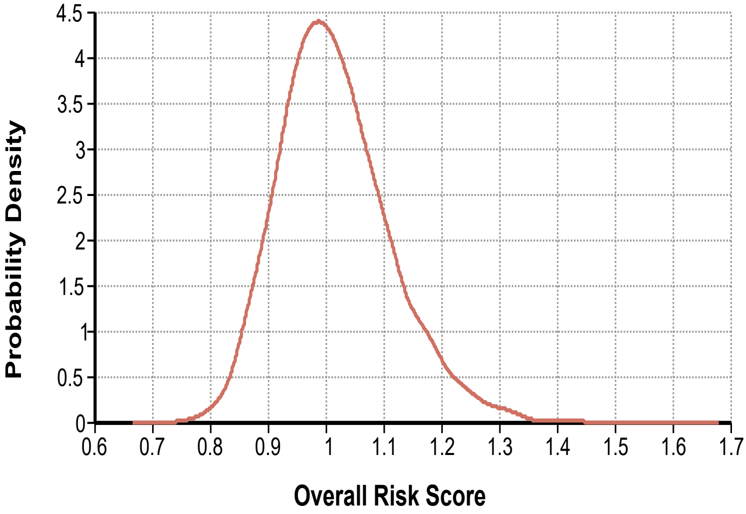




	Output Variable Distributions			Manual Calculation	
Risk Score Output	Min	Max	Mean	Min	Max
Maintenance Score	1.007	2.805	1.49	0.9272	1.7378
Flight Score	1.121	3.466	1.781	1.3347	2.0705
Damage & Related Impact Score	0	0.4197	0.084	0	0.3349
Overall Risk Score	0.7336	1.609	1.015	0.7854	1.1698

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.





Results: Reliability Testing

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

ANOVA assumptions were tested

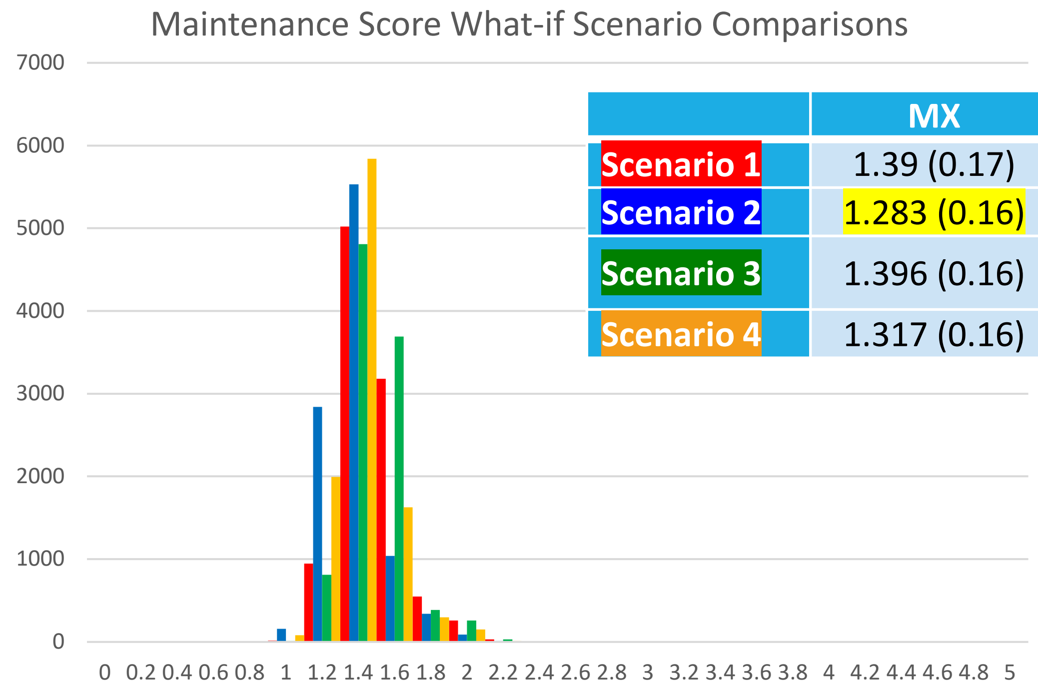
- Large sample size meets normality assumption
- A non-significant Levene's test verified homogeneity of variance

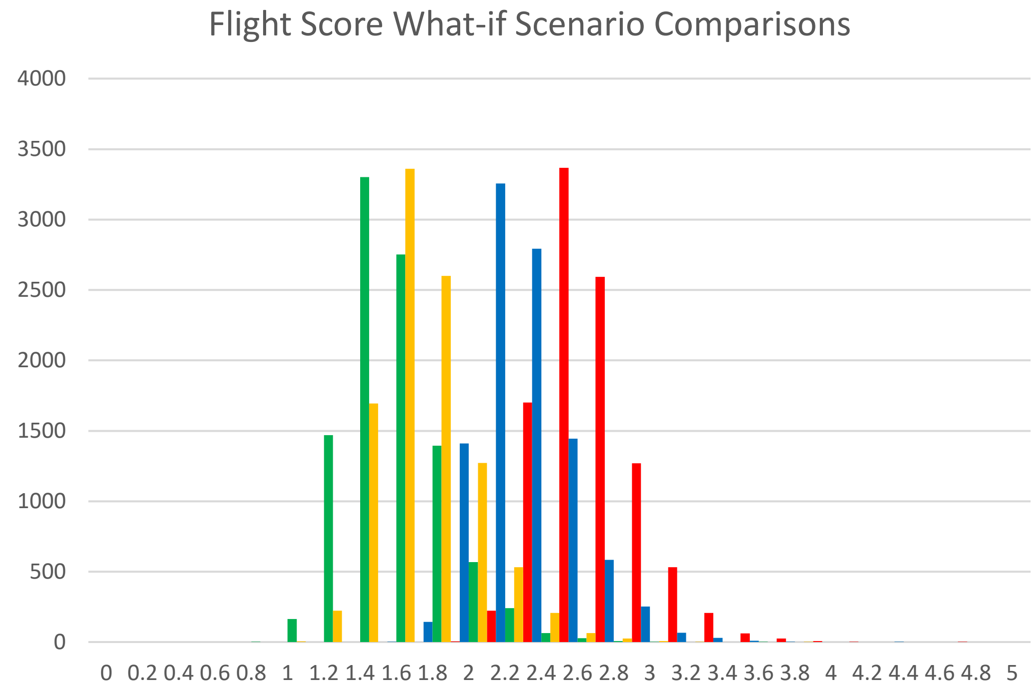
As there are no significant differences among groups, results are statistically reliable

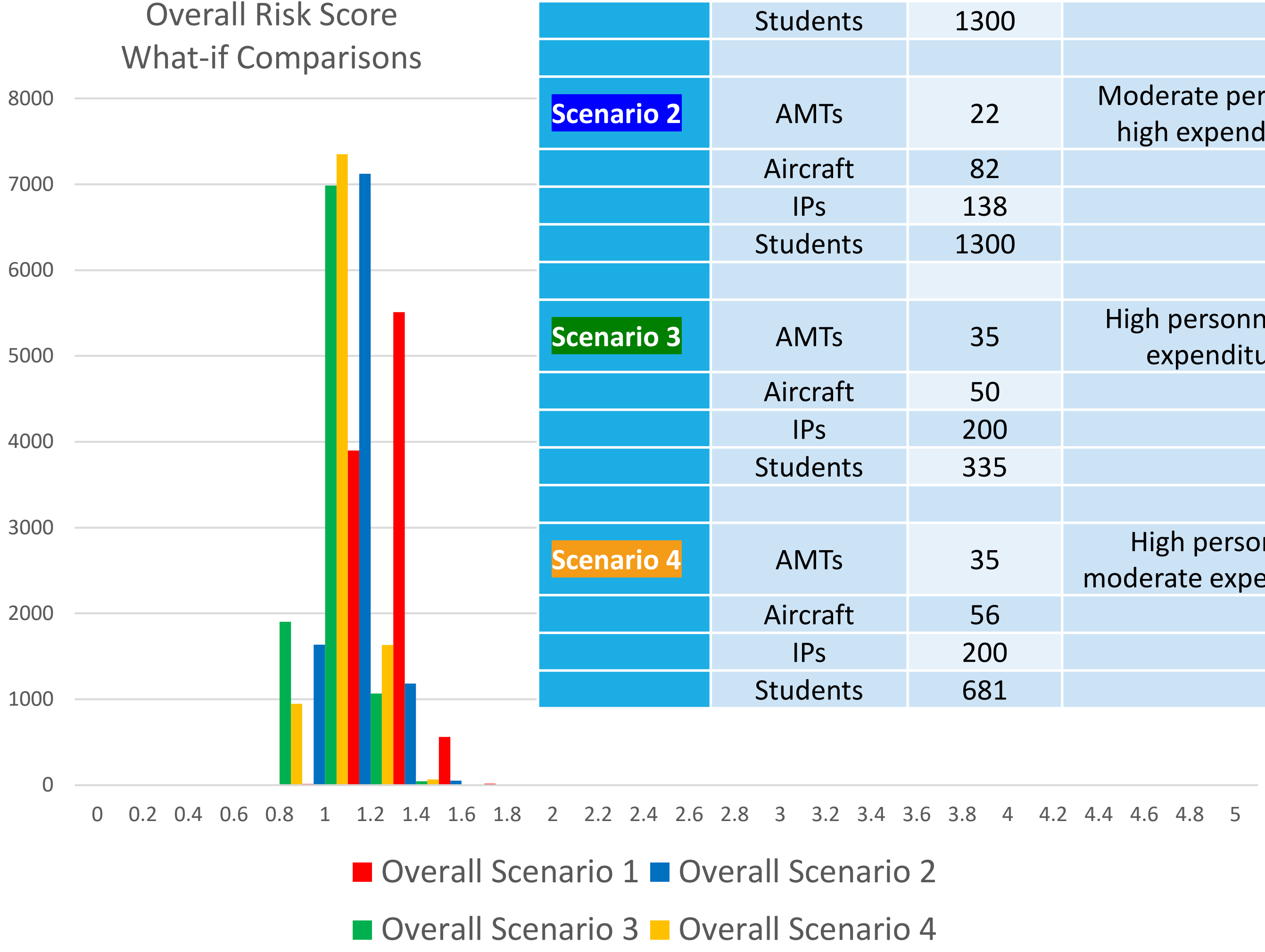
Output	Seed Value	Mean	Standard Deviation	ANOVA F	ANOVA P-value
Maintenance Score	99	1.49	0.1686	3.6446	0.3071
	50	1.491	0.1606		
	10	1.492	0.1638		
Flight Score	99	1.781	0.2627	81	0.0704
	50	1.784	0.2628		
	10	1.792	0.2692		
Damage & Related Impact Score	99	0.0835	0.0687	0.25	0.7048
	50	0.0829	0.0692		
	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		

Results: What-if Scenarios 1-4

	MX	FLT	Overall
Scenario 1	1.39 (0.17)	2.621 (0.26)	1.237 (0.10)
Scenario 2	1.283 (0.16)	2.248 (0.26)	1.092 (0.10)
Scenario 3	1.396 (0.16)	1.441 (0.26)	0.8845 (0.10)
Scenario 4	1.317 (0.16)	1.621 (0.26)	0.9149 (0.09)







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	

Discussion

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	

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)

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?

Conclusions

Theoretical Contributions

Transforms a nonstatistical model into a predictive, safety tool

Demonstrates the utility of Monte Carlo

Resource optimization and improved risk management for CFR Part 141s

Practical Contributions

Shapes understanding of the factors contributing to flight risk in Part 141s

Influence of resources-to-expenditures ratio on operational risk

Enhance the risk management component of the operation’s SMS

Adaptable for use in any CFR Part 141

Recommendations

Large, collegiate 14 CFR Part 141 flight training organizations

- Streamline data collection capabilities/storage
- Utilize a larger sample of raw data
- Include different controllable inputs
- Balance safety with financial resource allocation

Future research

- Explore capabilities and improve accuracy of predictions
- Expand the range of operational values
- Reevaluate the Damage & Related Impact variable
- Additional controllable inputs
- Incorporate human performance variables
- Include indicators (NAC, weather, ATC delay) (Anderson et al., 2020)

Questions?

Additional Slides

Input (22) and Output (4) 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

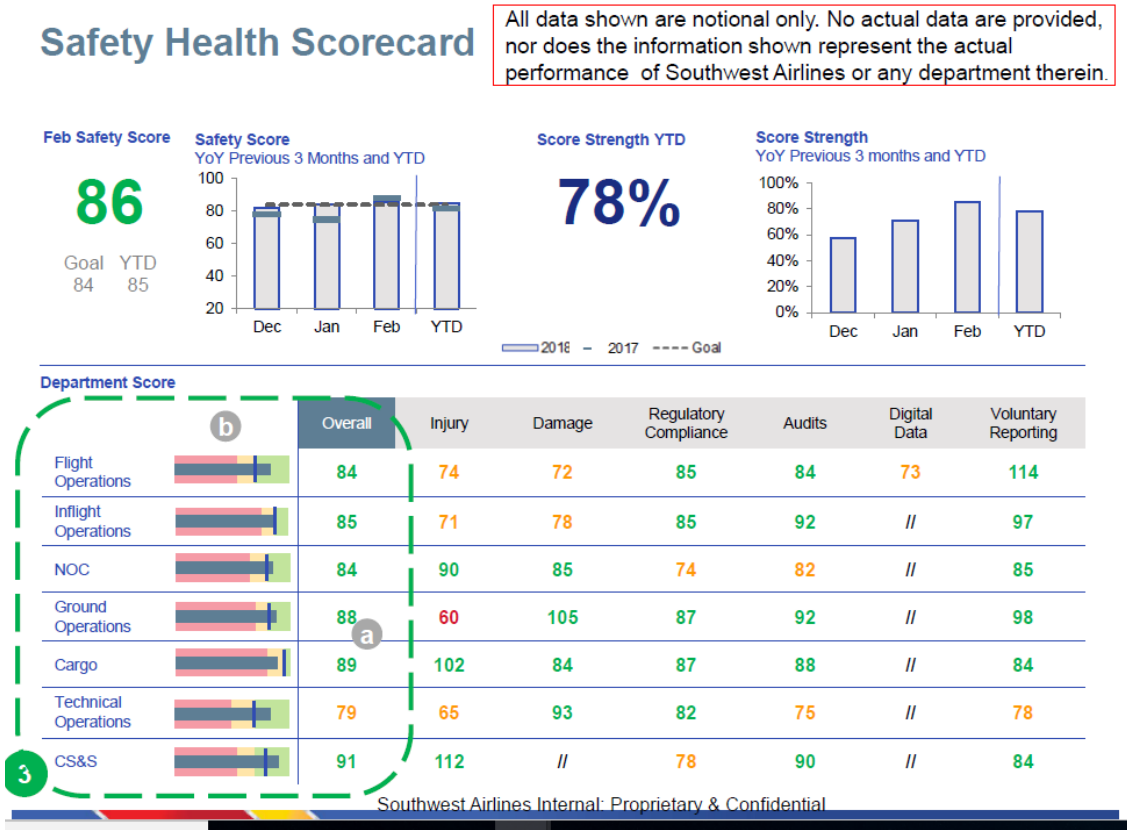
Relevant SPI	Variables	Variable Type
SPI-1 FLT: Occurrences	Number of reported tail strikes	Input Uncontrollable
	Number of hard landings	Input Uncontrollable
	Number of unstable approaches	Input Uncontrollable
	Number of RPM overspeeds	Input Uncontrollable
	Number of G exceedances	Input Uncontrollable
	Number of flap overspeeds	Input Uncontrollable
SPI-2 FLT: Safety Culture	Number of surveys collected	Input Uncontrollable
	Factor Scores	Input Uncontrollable
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

Relevant SPI	Variables	Variable Type
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

Other organizations utilizing safety scores

Southwest Airlines



Safety Health Scorecard

All data shown are notional only. No actual data are provided, nor does the information shown represent the actual performance of Southwest Airlines or any department therein.

Feb Safety Score

86

Goal 84

YTD 85

Safety Score
YoY Previous 3 Months and YTD

Score Strength YTD

78%

Score Strength
YoY Previous 3 months and YTD

Department Score

		Overall	Injury	Damage	Regulatory Compliance	Audits	Digital Data	Voluntary Reporting
Flight Operations		84	74	72	85	84	73	114
Inflight Operations		85	71	78	85	92	//	97
NOC		84	90	85	74	82	//	85
Ground Operations		88	60	105	87	92	//	98
Cargo		89	102	84	87	88	//	84
Technical Operations		79	65	93	82	75	//	78
CS&S		91	112	//	78	90	//	84

Southwest Airlines Internal: Proprietary & Confidential

GOL Airlines

CORPORATE NODSO 2018

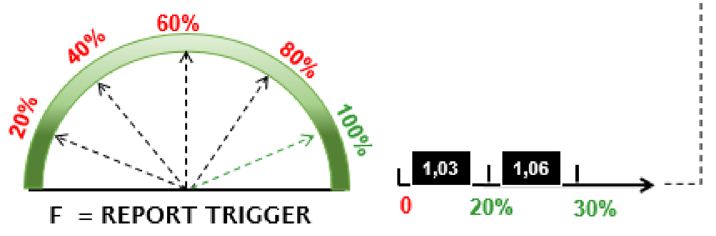
$$\text{NODSO} = \left[0,5 * \underbrace{\left(\frac{(C)}{(C)+\sum FV * \text{Quadrant Weight Index} + \sum FR * \text{Weight}}\right)}_{\text{Verifying safety barriers}} + 0,5 * \underbrace{\left(\frac{(R)}{(R)+\sum FV * \text{Quadrant Weight Index}}\right)}_{\text{Hazards identification}}\right] * F * I$$

Variables

C = the number compliance items | FV = the number of non-compliances (or Findings) resolved | FR = the number of recurrent non-compliances (or Findings) | R = number of Findings received

Quadrant Weight Index				
More than 60% of the time limit	4	8	16	32
30% to 60% of the time limit	2	4	8	16
Until 30% of the time limit	1	2	4	8
▨	Low	Med	High	Serious

Recurrent non-compliances (Weight)				
▨	1	2	4	8
	Low	Med	High	Serious



F = REPORT TRIGGER

I = IOSA results, to be considered only after the data analysis (Dec/18)