Spring 2023

Loss of Control In-Flight (LOC-I): A Mixed Methods Study of Voluntary Versus Mandatory Reports from the United States of America

Roger Lee
Embry-Riddle Aeronautical University, leer25@my.erau.edu

Loss of Control In-Flight (LOC-I): A Mixed Methods Study of Voluntary Versus Mandatory Reports from the United States of America

Roger Chak Man Lee

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

March 2023

Loss of Control In-Flight (LOC-I): A Mixed Methods Study of Voluntary Versus Mandatory Reports from the United States of America

By

Roger Chak Man Lee

This dissertation was prepared under the direction of the candidate's Dissertation Committee Chair, David A. Esser, Ph.D , 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.

David A. Esser Digitally signed by David A. Esser Date: 2023.03.18 07:27:17 -04'00'

David A. Esser, Ph.D.
Committee Chair

Jing Yu Pan Digitally signed by Jing Yu Pan Date: 2023.03.22 15:27:16 -04'00'

Jane Pan, Ph.D.
Committee Member

MIchael O'Toole Digitally signed by MIchael O'Toole Date: 2023.03.22 14:56:11 -04'00'

Mike F. O'Toole, Ph.D.
Committee Member

Committee Member

Kwok Chan Digitally signed by Kwok Chan Date: 2023.03.22 18:36:19 +08'00'

Kwok W. Chan, Ph.D.
Committee Member (External)

Steven Hampton Digitally signed by Steven Hampton Date: 2023.03.29 08:53:23 -04'00'

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

Alan J. Stolzer Digitally signed by Alan J. Stolzer Date: 2023.03.29 11:17:44 -04'00'

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

Lon Moeller Digitally signed by Lon Moeller Date: 2023.03.29 12:52:48 -04'00'

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

February 27, 2023

Signature Page Date

Abstract

Researcher: Roger Chak Man Lee

Title: Loss of Control In-Flight (LOC-I): A Mixed Methods Study of Voluntary Versus Mandatory Reports from the United States of America

Institution: Embry-Riddle Aeronautical University

Degree: Doctor of Philosophy in Aviation

Year: 2023

Loss of control in flight (LOC-I) is one of modern aviation's three most prominent fatal accidents. In the United States, air accidents are mandatorily reported to and investigated by the National Transportation Safety Board (NTSB). Established in 1976, the Air Safety Reporting System (ASRS) is a voluntary safety reporting (VSR) system administered by the National Aeronautics and Space Administration (NASA). Over 1.7 million ASRS reports have been processed to date. While the NTSB system handles LOC-I accidents, less severe incidents may have been reported voluntarily through the ASRS.

Safety reporting has been deemed the most valuable activity and the centerpiece of safety data collection for safety management systems (SMS). Both mandatory and voluntary safety reports (VSRs) are essential sources of SMS for safety assurance and risk management. Based on the age-old Heinrich's common cause hypothesis, mitigating hazards identified in low-severity safety reports, such as voluntary safety reporting (VSR) programs, would prevent more severe events such as fatal accidents.

This mixed methods study aims to determine whether normalized rates of LOC-I hazards identified by NASA, named Belcastro LOC-I Hazards, differ collectively or

individually across mandatory and voluntary safety reports in the United States, represented by NTSB and ASRS reports. The quantitative part dominates this study. LOC-I safety reports were obtained from searches performed on already classified cases by the administrators of the databases, and by augmented search based on the LOC-I precursors keyword search used by Belcastro et al. (2017). A total of 12,432 safety reports from 2004 to 2020 were analyzed.

The research results suggested that the Belcastro LOC-I Hazard rates were statistically different at the multivariate level across the four safety report groups for both commercial and general aviation. Out of the eight Belcastro LOC-I Hazards, five in general aviation and seven in commercial aviation displayed univariate differences. A cursory review of the narratives of the reports also suggested that the textual reports related to the Belcastro LOC-I Hazards were contextually different across the groups. These findings provided insights: firstly, ASRS was a credible source in identifying some, but not all, hazards leading to LOC-I accidents; secondly, the augmented search would enrich intelligence gained from the ASRS database for some LOC-I hazards; and, thirdly, the validity of Heinrich's common cause hypothesis was not generally supported.

While the NTSB system and investigations are more formalized, the research results suggested that ASRS safety reports are still effective in identifying some Belcastro LOC-I Hazards. This point is especially relevant in situations when accident data is limited. This research pointed to the need for a targeted approach, rather than one-size-fits-all, when using safety reporting databases. Before interrogating the data, practitioners should understand the precursors of the hazard to be analyzed, and the strengths and weaknesses of the associated safety reporting system. This awareness will enable safety professionals to calibrate, interpret, and supplement the data appropriately, resulting in more effective safety mitigations.

Keywords: MANOVA, discriminant analysis, quantitative method, qualitative method, loss of control in-flight, safety management system, voluntary safety reporting, Heinrich's theories, mixed methods analysis, multivariate analysis, univariate analysis.

Dedication

This dissertation is an 80th birthday present for my dearest mother, who is living through dementia in the evening of her life. She has served society, her students, and her family with the greatest love, passion, charity, hardship, and righteousness. Mum – there is nothing I can do to reverse your condition, but your life is my greatest lesson. I pray that you will stop being oppressed, suffering from the stigmatism of the so-called “norms”, and be freed into your truly authentic, wonderful, and beautiful self in the arms of God. I will always be there wherever you go.

To the kids, Isabel and Aaron, and my wife, Yvette. You are my pleasure, reason, and treasure. Thank you for standing by me in the dichotomies of life.

Kids – You both sacrificed a part of your childhood for a father overwhelmed by work and study. I have been indulging in seeking those things that are above. I gave up the time being with you. By doing this Ph.D., I wanted to show you the journey of continuously seeking the meaning of life and forever fighting for that. This meaning is my passion for protecting and doing my part for humanity. I hope you both can pursue your dreams in something meaningful and that you enjoy them with passion and resilience, as I did during this Ph.D. journey. Your dad is not perfect, but at least he is a constant pursuer of the road ahead.

It has certainly not been easy; the pain seems constant. I can count on my fingers the days when I am not sleep-deprived. I hope this piece of work can hopefully save some lives.

Acknowledgments

I want to express my deepest gratitude to Dr. David Esser for his never-ending encouragement and motivating words, no matter rain or shine. As a world-class safety professional, an aviator with global experience, and a mentor, Dr. Esser had the insight and personality to guide me through the Ph.D. labyrinth, coordinate with the committee and external experts, and ensure that I stayed on the glide path. I would also like to thank Dr. Antonio Cortes, who offered tremendous support for the earlier part of my dissertation and is a trusted wingman in safety.

My heartfelt thanks to Dr. Pan, a committee member who has worked tirelessly to guide me through the maze of statistics. Gratitude also to Dr. O'Toole and Dr. Chan, who have kindly taken up the challenge of supervising a Ph.D. student's work from the other side of the world. I am thankful to have a committee with such a broad spectrum and depth of expertise.

I want to acknowledge the tremendous support from the teaching and support staff at the Embry-Riddle Aeronautical University College of Aviation, starting with Dean Stolzer. All of you have made my life unanticipatedly hard but tremendously challenging in the past seven years. Special thanks also to the industry subject matter experts who assisted in the taxonomies mapping of this research, all of whom are highly respected friends of mine in the industry: Capt. François Lassale, Capt. Richard Boswell, Mr. C.V. Thian, Capt. Hans von Blucher, Capt. Peter Lawrie, and Capt. Denis Portier.

Lastly, to those pillars, role models, and wingmen of my safety career in the air and on the ground thus far. My sincere gratitude to you for believing in my ability and integrity as a safety practitioner: Uncle and Dr. Wing Tsang, Ms. Janice Tsang, my late

Aunt Mrs. Irene Lam, Mrs. Frances and Mr. Jimmy Ellis, Dr. Ben Lam, Ms. Ker Shien Lee, Ms. Tracy Chow, Ms. Angela Wee, Dr. Philip Cheung, Mr. Don MacKenzie, Mr. Michael Heap, Capt. Mike Webber, the late Dr. Roger Morris, Mr. Steve Tafe, Mr. Tom Perry, Mr. Roger Dillon, Mr. David Hawken, Mr. Russell Batchelor, Sir Michael Kadoorie, Mr. David Tong, Mr. Simon Li, Capt. Victor Liu, Mr. Bjorn Naf, Ms. Sue Ann Law, Mr. S.J. Tan, Mr. John Leigh, Ms. Connie Leong, Mr. Romand Xu, Mr. Geoffrey Yan, the late Ms. Connie Wong, Ms. Sandra Mak, Dr. Y.B. Lee, Mr. Gary Zhan, Mr. Jimmy Ma, Ms. Jessie Lam, Mr. Chen Ning, Mr. Zhu Tao, Mr. Li Dianchun, Mr. Guosong Zhong, Capt. Simon Wu, Mr. Nicolas Wong, Capt. Samson Lau, Capt. Ian Quinn, Capt. Webster Siu, Mr. K.C. Man, Ms. Yamani Chan, Capt. John Richardson, Mr. Andrew Cowen, Mr. Chris Thomas, Mr. Rick Howells, Ms. Zahara Mohideen, Mr. Seong-Jo Choi, Mr. Omer Ozyasli, Mr. Thomas Chan, Mr. Raymond Ng, Ms. Yvonne Tou, Mr. Spencer Mallet, Ms. Elaine Wong, Mr. Terence Chan, Mr. Sam Law, Mr. Eric Cheung, Ms. Eva Li, Mr. C.H. Wong, Dr. Vanessa Forbes, Capt. Michael Cheng, MTR's Corporate Safety teammates, Ms. Kathy Kendall, Mr. Mike Parker, Ms. Gill Meller, Mr. Ronnie Tong, Dr. Jacob Kam, Dr. S.M. Lam, Ms. Joey Leung, and many more.

Finally, thank God for being my rock. It is worth it.

Table of Contents

Page

Signature Page iii

Abstract iv

Dedication vii

Acknowledgments viii

List of Tables xiv

List of Figures xvii

Chapter I: Introduction 1

Statement of the Problem 6

Purpose Statement 7

Research Questions and Hypotheses 7

RQ1 7

RQ2 7

H_A1 8

H_A2 8

H_A3 8

H_A4 8

H_A5 9

H_A6 9

H_A7 9

H_A8 9

H_A9 9

Significance of the Study 9

Delimitations 12
Limitations and Assumptions 13
Definitions of Terms 19
List of Acronyms 22
Chapter II: Review of the Relevant Literature 25
Safety Performance in Modern Aviation 25
LOC-I Events 27
Impetus for the Introduction of Aviation VSR in the United States 37
Development of VSR after ASRSs 40
The Birth of SMS 45
The Rise of SMS in Aviation 47
Relevance of SMS in Managing Critical Hazards in Aviation Such As LOC-I 48
Safety Reporting System for an Airline's SMS 49
Relevance of Heinrich's Theories to SMS and VSR 50
Gaps in the Literature 60
Theoretical Framework 60
Research Model 62
Hypotheses and Support 63
Summary 65
Chapter III: Methodology 67
Research Method Selection 67
Population/Sample 68
Population and Sampling Frame 68

Sample Size ... 69
Sampling Strategy ... 71
Data Collection Process ... 71
Design and Procedures ... 72
Apparatus and Materials ... 75
Sources of the Data ... 75
Ethical Consideration ... 77
Measurement Instrument ... 77
Variables and Scales ... 78
Data Analysis Approach ... 80
Reliability Assessment Method ... 89
Validity Assessment Method ... 90
Summary ... 91
Chapter IV: Results ... 93
Demographics ... 93
Taxonomy Mapping ... 96
Quantitative Analysis – MANOVA ... 102
Discriminant Analysis to Verify the MANOVA Univariate Result ... 142
Qualitative Analysis – A Supplement ... 148
Chapter V: Discussion, Conclusions, and Recommendations ... 157
Discussions of Results ... 157
Conclusions ... 172
Theoretical Contributions ... 174

Practical Contributions 176

Limitations of the Findings 177

Recommendations 178

References 182

Appendix A 193

Appendix B 194

Appendix C 195

Appendix D 197

Appendix E 198

Appendix F 202

Appendix G 206

Appendix H 210

Appendix I 211

Appendix J 216

Appendix K 219

Appendix L 225

Appendix M 234

Appendix N 245

Appendix P 253

List of Tables

Table Page

1 NASA Research Publications on LOC-I 31

2 Primary Causes, Precursors, and Hazards of LOC-I Events 33

3 Research Questions and Alternative Hypotheses of the Current Study 64

4 Sample Frame for Events Identified as LOC-I in the ASRS, AIDS and NTSB Databases Between 2004 and 2020. 70

5 Independent, Dependent Variables (IV & DV) and Covariate for the Research . 78

6 Proportion of Flight Conditions from LOC-I Safety Reports 94

7 Data Elements Extracted from ASRS and NTSB Code Categories for SME Panel Mapping 97

8 Numerical Mapping Codes with Belcastro LOC-I Hazards 100

9 Interrater Reliability Statistics From the NTSB and ASRS Taxonomies Mapping Exercise by SME Panel 102

10 DVs with Zero Frequency in Parts 121 and 135, and Part 91 Datasets 105

11 Descriptive Statistics for DVs and IVs Used in the MANOVA from 2004 to 2020 106

12 Mahalanobis D2 Multivariate Outlier Analysis Results for Parts 121 and 135, and Part 91 Datasets 109

13 Normality Tests Based on Kolmogorov-Smirnova and Shapiro-Wilk for Parts 121& 135 and Part 91 Datasets 111

14 Multivariate Homogeneity Test Results for Parts 121 and 135, and Part 91 Datasets 115

15 Levene's Test Results for Parts 121 and 135, and Part 91 Datasets 116

16 The Differences in Shapiro-Wilk Results Between Original and Transformed Parts 121 and 135, and Part 91 Datasets 119

17 Multivariate Homogeneity Test Results for the Transformed Parts 121 and 135, and Part 91 Datasets 123

18 Levene's Test Results for Transformed Parts 121 and 135, and Part 91 Datasets 126

19 Pillai's Trace Test Result for Multivariate Analysis of Variance Based on IV CODE_TYPE. 127

20 Univariate Tests Results for Parts 121 and 135, and Part 91 Datasets............ 129

21 Post Hoc Comparisons for Individual Group Differences in DVs (TX5 Excluded for Both Datasea constantCODE_TYPE) 133

22 Discriminant Analysis Wilk's Lambda Results for Parts 121 and 135, and Part 91 Datasets 144

23 Discriminant Analysis Eigenvalues for Parts 121 and 135, and Part 91 Datasets 145

24 Percentage of Validated Predicted Membership Results from the Discriminant Analysis for Both Datasets. 147

25 Discriminant Analysis Structure Loadings on Function Results for Parts 121 and 135, and Part 91 Datasets 147

26 Top 10 Frequent Word Comparison for Parts 121 and 135 Dataset from Tree Maps 149

27 Summary of Qualitative Analysis from NTSB, ASRS Narratives.................. 150

28 Recommendations from Answers to RQ1 and RQ2 to Safety Practitioners..179

List of Figures

Figure Page

1 Fatalities Statistics by International Civil Aviation Organization (ICAO) and the Commercial Aviation Safety Team (CAST) Common Taxonomy Team (CICTT) 2

2 ASRS Report Process Flowchart 15

3 Channels Used by ASRS System in Providing Feedback to Industry Stakeholders 16

4 Increasing Severity of Four Groups of LOC-I Safety Report Types (IVs) Deployed in this Study with Illustration of Reason's (2016) Accident Causation Model (2016) 18

5 ICAO Accident Statistic Graphs 26

6 Flight Accident Category Frequency and Fatality Risk 2013–2017 27

7 Generic LOC-I Accident Sequences 34

8 Simplified Generic LOC-I Model 35

9 Evolution of Aviation Safety Reporting Systems in the United States and Its Regulatory Implications 43

10 Heinrich's Triangle—An Application of the Common Cause Hypothesis 51

11 FAA Versions of Heinrich's Triangle 56

12 Figure Summarizing the Four Steps in Data Analysis 74

13 Examples of Interrater Indices (source: Gisev et al., 2013) 82

14 ASRS and NTSB Reports Mapping Table Generation Steps 83

15 Examples of the ASRS Coding Table (Left) and NTSB Air Accident Coding Table (Right) 84

16 Basic Variable Type and Relations MANOVA Model Adopted 85

17 Examples of Word Cloud and Tree Map from Previously Conducted Analysis on Air Safety Reports (Lee, 2017) 89

18 Examples of Flight Phases Data from the Parts 121 and 135 ASRS Dataset 95

19 Illustrations of Codes Selection for Mapping Based on the Pareto Chart for ASRS Part 91 Augmented Search Dataset 98

20 Examples Box Plots Illustrating Univariate Moderate and Extreme Outliers, Parts 121 and 135 DV6 (left) and Part 91 DV4 (right) 108

21 A Segment of the Parts 121 and 135 Dataset Scattered Plots Matrix demonstrating lack of linearity for DV5 and DV7... 113

22 Illustration of Parts 121 and 135 Cube Root Transformation Results on DV1.. 121

23 Examples of Estimated Marginal Means of TX1 in Parts 121 and 135 and TX2 in Part 91 Dataset... 131

24 Visualization of the Post Hoc Comparisons for Individual Group Differences in DVs (TX5 Excluded) and IV (CODE_TYPE), Parts 121 and 135 (top) Part 91 (bottom)... 135

25 Word Clouds from the Classified and Augmented Searched LOC-I Reports Synopsis and Narratives from AIDS, ASRS and NTSB Databases........................ 154

Chapter I: Introduction

This chapter introduces the voluntary safety reporting (VSR) system as a data source for safety management systems (SMSs) in the aviation industry. It describes the problem surrounding the validity of publicly available open-loop VSRs, such as ASRS, where relatively minimal validation, investigation, and feedback have been conducted. It further develops into the purpose statement, research questions, and hypotheses for this research.

Background

Operators' safety reporting has been deemed the most valuable activity and the centerpiece of safety data collection under SMS (Maurino, 2017). Based on established concepts such as Heinrich's triangle and the associated common cause hypothesis, safety practitioners are taught that mitigating hazards identified in low-severity safety reports from VSR programs would prevent more severe events such as fatal accidents (Manuele, 2011).

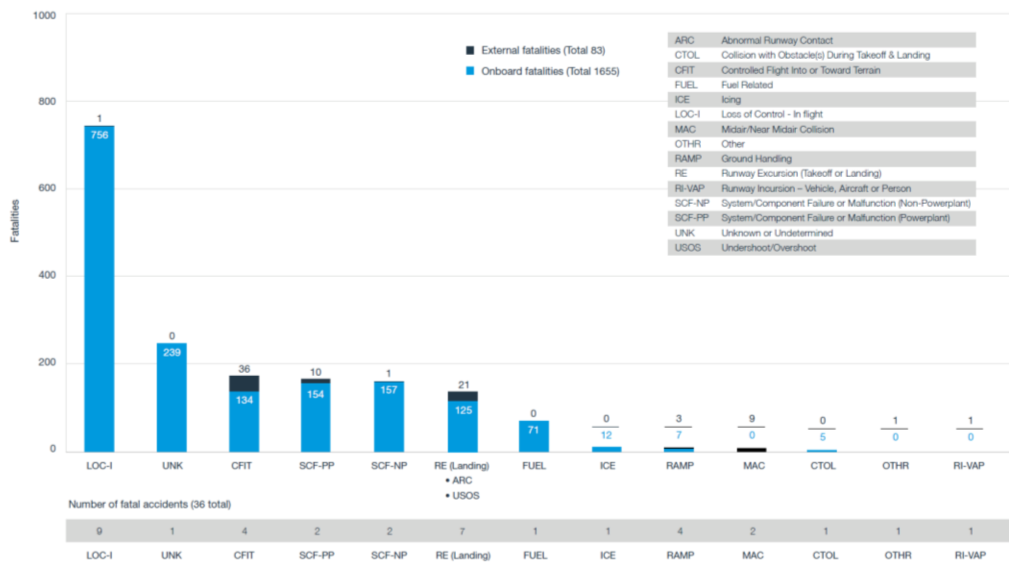
According to NASA (2022a), ASRS is intended "to collect, analyze, and respond to voluntarily submitted aviation safety incident reports in order to lessen the likelihood of aviation accidents" (p. 1). It is unclear whether the nature and quantity of hazards reported in ASRS, typically lower in severity, are similar to the higher severity events found in accident investigations. The existence of this similarity should contribute toward ASRS' intent to reduce the likelihood of aviation accidents.

Air accident statistics published by ICAO state that loss of control in-flight (LOC-I) events are among the three most prominent types of accidents in modern aviation

(ICAO, 2020). Statistically, LOC-I accidents have the highest occurrence and fatality risk among modern commercial aviation accidents (IATA, 2018), as shown in Figure 1.

Figure 1

Fatalities Statistics by International Civil Aviation Organization (ICAO) and the Commercial Aviation Safety Team (CAST) Common Taxonomy Team (CICTT)



Note. Reprinted from "Statistical Summary of Commercial Jet Airplane Accidents Worldwide Operations 1959-2021," by Boeing, 2022, 53rd Edition, p.13.

In 2009, triggered by instrument malfunctions, Air France flight 447 resulted in the loss of 228 lives in a LOC-I accident. In 2018 and 2019, two catastrophic LOC-I accidents involving the newly introduced Boeing 737 MAX 8 airliner led to the loss of 346 lives. Examining the precursors to 122 LOC-I accidents and incidents worldwide from 1996 to 2010, Belcastro et al. (2017) identified a combination of hazards such as vehicle problems, external hazards, inappropriate crew response, and vehicle upset led to a LOC-I event. These hazards were abbreviated as *Belcastro LOC-I Hazards* in this dissertation.

Safety management systems (SMSs) were introduced to the aviation industry in the early 2000s, originating from other safety-critical industries, such as oil and gas. The International Civil Aviation Organization (ICAO, 2013) promulgated the SMS model, which consists of four elements: safety policy, risk management, safety assurance, and safety promotion. Safety assurance incorporates management reviews to ensure safety goals are being achieved. It oversees an organization's effectiveness in managing risks. Stolzer et al. (2017) highlighted the relationship between safety risk management and safety assurance, which relies on an operator's Internal Evaluation Program (IEP) to oversee such effectiveness. The identification and assurance of risks can be performed in the following ways: quality assurance, line operations safety audits (LOSAs), flight operational quality assurance (FOQA), or a non-punitive safety reporting system. A combination of these elements provides the *risk picture* of the organization for the implementation of proactive control measures centered on a risk-based approach (ICAO, 2018; Petitt, 2017; Steckel, 2014). The SMS is a key defense in managing hazards with potentially high-severity consequences in aviation, such as LOC-I (Cacciabue et al., 2015; ICAO, 2020).

Earlier strategies for safety assurance were founded on works by Herbert William Heinrich, an industrial insurer who performed archival data analyses based on insurance claims data in the 1930s. These led to Heinrich's common cause hypothesis and the 300:29:1 accident ratio in Heinrich's triangle (Davies, 2003). This triangle has since been featured in safety science textbooks up to modern times (Dekker, 2019; Friend & Kohn, 2014; Marsh, 2017). While specific inputs of the safety assurance components are quantitative, such as FOQA, the qualitative voluntary safety report continues to serve as a tool for identifying the operational hazards for proactive mitigations. Based on Heinrich's principles, a less severe LOC-I event typically reported via voluntary safety reports resulting in full recovery with an uneventful outcome potentially shares the same hazards as those found in LOC-I events leading to a hull loss. On this assumption, mitigating the hazards that lead to low-severity LOC-I events will reduce the likelihood of high-severity events. However, despite their usefulness as a rule of thumb, Heinrich's theories have been challenged due to their lack of research rigor and verifiable empirical data (Manuele, 2011). Given these uncertainties, the value of publicly available open-loop voluntary reporting systems, such as ASRS, and associated resources deployed to promote such systems, are increasingly being questioned.

By the beginning of 2020 (prior to the emergence of the COVID-19 pandemic), the annual number of commercial aviation flights had grown to 37.8 million flights globally (ICAO, 2019). LOC-I accidents occurred, averaging six yearly, 94% involving passengers or flight crew fatalities. These accidents led to more fatalities than any other accident category (IATA, 2018; IATA, 2019). However, Maurino (2017) cautioned that formulating safety strategies based on limited accident and incident data alone may not be effective safety management. Hence, lower severity LOC-I events would be of interest to be deployed as a supplementary data source for the safety management of LOC-I. This study compared eight Belcastro LOC-I Hazard rates among four independent groups of LOC-I reports, each with different severity levels. The reports were obtained from the publicly available voluntary and mandatory safety reporting systems in the United States.

The eight hazard rates were this study's dependent variables (DVs). NASA Aviation Safety Reporting System (ASRS) represented voluntary safety reports (VSRs), while National Transportation Safety Board (NTSB) accident and incident investigation reports represented mandatory safety reports (MSRs). The LOC-I cases within each reporting system contained two severity levels: those that had been classified as LOC-I by the respective reporting system's administrator, or those that were not classified originally as LOC-Is, but were identified from *keyword search* per Belcastro et al. (2017) as they contained precursors of LOC-I. Therefore, four groups with an increasing severity level of LOC-I reports per dataset were utilized. These groups were represented by this research's independent variable (IV). The comparison was made among the commercial aviation and general aviation reports independently.

Quantitative analyses were performed on data from already-coded voluntary and mandatory aviation safety reports originating from the United States within a 17-year period of 2004 to 2020, supplemented by qualitative analyses. Based on the accident pyramid by Herbert W. Heinrich (1931), also known as the Heinrich Triangle, much research has been carried out to identify the relationships between high-severity and low-severity safety events in a variety of safety-critical industries (e.g., Bellamy, 2015; Gallivan et al., 2008; Marshall et al., 2018; Moore et al., 2020; Yorio & Moore, 2018). Such research attempted to statistically explore the predictability of higher severity events from lower severity events. However, they have not focused on whether the varying severities of events shared common causes.

SMS adopts a data-driven approach in the identification of hazards. Adequate quantity and quality of data are required to describe the *larger mechanism* to generate effective mitigating measures (Stolzer et al., 2017). The number of LOC-I accidents with severe consequences, fortunately, remained low. Hence, the quantity of reactive data was limited. Data must be sought from elsewhere to reduce the probability of LOC-I further. Using proactive voluntary safety reports such as ASRS is a possible option. Secondly, Belcastro et al. (2017) indicated that, in addition to obtaining LOC-I information from reports already classified as LOC-I by the accident database administrators, a precursor keywords search, named as augmented search in this research, yielded more data on events not classified initially as LOC-I, but experienced precursors of LOC-I. These events were later mitigated by measures such as crew action, which resulted in uneventful outcomes. Identification of augmented searched events provided an extra LOC-I dataset for one safety report database, enlarging the sample frame (Belcastro et al., 2017).

From the civil aviation perspective, it was unclear if the same hazards were shared among the four severity levels of LOC-I events within the same operational certification dataset: two (*classified* and *augmented* search) from voluntary ASRS reports and two from incidents or accidents investigated by the NTSB. If the hazards were different, then mitigating hazards identified in lower severity sets might not effectively mitigate the hazards in the higher severity sets. The probability of LOC-I occurrence will stagnate, negating the continuous improvement aim of SMS (Stolzer et al., 2017).

Statement of the Problem

The level of reliance safety practitioners should apply on *open-loop* safety reports such as ASRS to effectively mitigate high-severity LOC-I events is unknown. While there are various publicly available VSR repositories, such as ASRS, an extensive literature search has not identified any assessment to date on the relevance of the information in such reports in being a credible source for mitigating high-severity LOC-I accidents in the United States. The literature review also did not reveal any supplemental information required to compensate for the deficiencies of ASRS reports, if any, for LOC-I mitigations.

Purpose Statement

The purpose of this study was to determine whether there were differences in the eight Belcastro LOC-I Hazard rates (DVs) among four severity groups (IV) of LOC-I safety reports originating from voluntary (ASRS) and mandatory (NTSB) datasets for the commercial and general aviation operating environments in the United States. It also identified the particular Belcastro LOC-I Hazard rates that displayed significant differences or similarities between ASRS and NTSB LOC-I reports.

Research Questions and Hypotheses

This research was based on two fundamental research questions:

RQ1

Do Belcastro LOC-I Hazard rates differ across types of safety reports for commercial and general aviation?

RQ2

Which of the individual Belcastro LOC-I Hazards display(s) significant difference(s) in hazard rates across types of safety reports for commercial and general aviation?

The research questions were founded on Belcastro's (2017) research from accident investigation reports on the Belcastro LOC-I Hazards. The theoretical basis of the hypotheses is detailed in Table 3. These research questions were primarily answered quantitatively, although a cursory qualitative analysis was used to provide additional insights. Hypothesis H_A1 addressed the multivariate comparison related to RQ1, and hypotheses H_A2 to H_A9 addressed the univariate comparison related to RQ2. The four types of safety reports, independent variable groups of this research, were combinations of search types, classified and augmented, and origin types, ASRS and NTSB.

H_A1

The group mean vectors in Belcastro LOC-I Hazard rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A2

The means of adverse onboard conditions - vehicle impairment rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A3

The means of adverse onboard conditions - system and components failure / malfunction rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A4

The means of adverse onboard conditions - crew action / inaction rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A5

The means of external hazards and disturbances - inclement weather atmospheric disturbances rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A6

The means of external hazards and disturbances - poor visibility rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A7

The means of external hazards and disturbances - obstacle rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A8

The means of abnormal vehicle dynamics and upsets - abnormal vehicle dynamics rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A9

The means of abnormal vehicle dynamics and upsets - vehicle upset conditions rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

Significance of the Study

From the theoretical perspective, this research identified that the hazards contained in the lower severity LOC-I reports were not the same as the higher severity reports collectively. This finding provided the theoretical justification to refute Heinrich's hypothesis of common causality in the context of LOC-I. If the relatively lower severity ASRS reports did not contain similar hazards compared with the higher severity reports in normalized quantities, then mitigating hazards identified from ASRS might not directly address the hazards that led to higher severity LOC-I incidents as identified in the NTSB reports. Secondly, from the risk management perspective, the results of this study supported Cooper's (2019) theory that dedicated hazard identification and risk control measures for risks with critical consequences are necessary, regardless of the likelihood of occurrence. This is because critical hazards would not be identified from VSRs, typically lower in consequential severity. Due to the relatively low probability of events with severe consequences, based on the traditional SMS risk tolerability matrices, hazards that may lead to critical consequences may not be assessed as high risks and, therefore, will not attract prioritized attention.

From the practical perspective, the number of LOC-I accidents is not as high as the voluntarily reported low-severity LOC-I events. Therefore, developing preventive measures for LOC-I may use proactive voluntary safety reports such as ASRS and the formal investigation reports conducted by the NTSB. This research highlighted that the means of Belcastro LOC-I hazard rates were not different across the types of safety reports. Hence, operators can make use of ASRS, a publicly available VSR system, to derive preventive measures on some Belcastro LOC-I hazards to prevent high-severity LOC-I events. This research also warned that the Belcastro LOC-I hazard rates differed between ASRS and NTSB reports. In this case, ASRS, at its current state, may be of limited use to support the derivation of high-severity LOC-I safety mitigations. Thirdly, in a world of limited data for critical hazards such as LOC-I, the study informed whether a higher level of efficacy on publicly available VSRs, such as ASRS, can be achieved by the precursor keyword search method named augmented search in this research. (*Augmented search* is further defined within the Definition Section.)

Regarding the groups who could benefit, the study should provide primarily U.S.-based aviation regulators, operators, and front-line staff insight into the relevance of publicly available open-loop VSR in the United States, such as ASRS, in implementing the SMS. For regulators, ICAO Annex 19 requires each member state to exercise its surveillance requirement on operators' SMSs (ICAO, 2019). When exercising this obligation, the research results should inform regulators of the representativeness of the operator's risk profile from examining VSR data. If the representativeness is low, regulators may need to adjust the surveillance strategy by applying more *command-and-control* type safety assurance activities, such as inspections, audits, and monitoring activities, and assessing the effective implementation of VSR (Mills, 2011).

This study should guide operators in setting the strategy to seek the most appropriate data sources from their assurance and accident prevention programs for mitigating high-severity safety events such as LOC-I. Such a strategy should consider the dependency level placed on publicly available open-loop VSR to inform elements of SMS such as safety promotion, risk management, and policy and standards. Operators can apply treatments to relevant data to optimize safety intelligence, especially when VSR is the only option available.

Delimitations

Both IATA (2020) and Boeing (2022) have identified LOC-I events as the type of air accidents resulting in the highest number of lives lost. This research focused on LOC-I reports from one voluntary (ASRS) and one mandatory (NTSB) reporting system in the United States, regardless of whether the event had a successful or severe outcome.

A search on the ASRS database identified 770 reports that were classified as LOC-I for commercial aviation (Parts 121 and 135) and 1,041 reports (named loss of aircraft control in ASRS database) for general aviation (Part 91) between 2004 and 2020. In the same period, 2,791 LOC-I classified reports for commercial aviation and 3,045 reports for general aviation were identified by NTSB. Based on the above datasets, this study was limited to general and commercial aviation fixed-wing operation LOC-I events between 2004 and 2020. Per the NTSB website (NTSB, 2021), the events recorded were civil aviation accidents and selected incidents within the United States, its territories, and possessions, and in international waters.

In addition, instead of using events classified as LOC-I in the relevant databases, Belcastro (2017) conducted an augmented keyword search for precursors to LOC-I for reports that had not been classified as such initially. Keywords used were *loss of control*, *upset*, *unusual attitude*, *stall*, *crash out of control*, and *uncontrolled descent*. An augmented search was conducted for this research and identified 1,732 reports from commercial aviation and 1,028 reports from general aviation in the ASRS database, and 224 and 3,447 from the NTSB database, respectively. The search added one data group to each database, leading to four independent data groups for each certification type. From the FAA Accident and Incident Data System (AIDS) database, an *augmented search* resulted in 52 commercial and 214 general aviation LOC-I reports. As part of the data verification process, the unique case numbers were checked to ensure they did not overlap between the *augmented* and the *classified* groups so that each group was independent.

Due to the presence of coded data for quantitative analysis in the ASRS and NTSB databases, data from the AIDS database was only added during the supplementary qualitative data analysis phases. This qualitative dataset aimed to supplement the quantitative analysis from ASRS and NTSB, informing the research from the perspective of the mid-severity *incidents*. The usage of AIDS data was not designed to support the generalization of the ASRS or NTSB data quantitatively or increase their level of statistical significance. The AIDS analysis was designed to fill the void between low-severity and high-severity events qualitatively. A narrative search on the AIDS database containing *loss of control* highlighted 62 general aviation and 11 commercial aviation LOC-I events for the selected period between January 1, 2004, and December 31, 2020.

The research results and analyses were only valid for the period the data originated (i.e., 2004 to 2020). The aviation industry experienced substantial growth alongside the introduction of SMS during this period. The FAA mandated the full implementation of SMSs by March 2018 (FAA, 2015). Soon after, the industry encountered COVID-19 in 2020, which substantially reduced the number of flights per year, which gradually increased in 2021.

Limitations and Assumptions

The analysis involved in this research may be sensitive to flight hours, as identified by Anderson (2013). The variation of Belcastro LOC-I Hazard frequencies may be affected by exposure in flight hours. Rather than an analysis based on the frequency of LOC-I events, per Anderson (2013), an analysis based on normalized LOC-I hazard occurrence rates was more appropriate to address the possible covariate due to flight hours. For general aviation, the denominators for rate calculations have been provided voluntarily by aircraft owners and operators over the years as part of the FAA General Aviation and Part 135 Activity Surveys (FAA, 2020). The accuracy for corresponding calculations in commercial aviation is expected to be higher, given that data are reported by operators to the centralized Bureau of Transportation Statistics (BTS) database (BTS, 2020).

The basis of this research was analyzing coded and textual data presented in safety reports. It was assumed that mandatory investigation reports were completed in a factual manner, and voluntary reports were submitted truthfully and candidly by their reporters. All coded data used in the study were assumed to have been accurately classified. Chapter IV further explores these assumptions alongside the results obtained in this research. It was acknowledged that the factual content of a major investigation report was of greater detail and rigor than that of a VSR or a low-severity investigation report. NASA does not conduct investigations into the relatively lower severity voluntary safety reports; however, this does not mean a total absence of validation of the submitted report has been conducted through ASRS. As explained in the ASRS Director's program briefing (NASA, 2018), NASA carried out validation on receipt of an ASRS report. This validation might include a callback by an ASRS analyst to clarify the information reported before data de-identification (see Figure 2).

Figure 2

ASRS Report Process Flowchart

Note. Reprinted from the ASRS Director's Program Briefing. Copyright 2022 by NASA (p. 16).

NASA (2022b) stated that, "The ASRS team is composed of experienced pilots, air traffic controllers and mechanics, as well as a management team that possesses aviation and human factors experience" (p.7). As no in-depth investigation would be conducted for ASRS reports, it was argued that if factors were not explicitly identified from the submitted report, it was less likely that such factors would be discovered before the report was closed. By design, ASRS is a publicly available open-loop system with no official follow-up on the individually reported events, unlike Aviation Safety Action Program (ASAP) reports. The significant benefit of ASRS is that the de-identified data is available to the public.

Due to the confidentiality restriction for data from other VSR programs, such as ASAP, this research was designed to focus on publicly available data, such as ASRS, to represent a VSR program. Other VSR programs might have a different rigor of investigation and individual feedback. Therefore, the results of this study cannot be generalized to all VSR programs globally but only to the publicly available programs that provided a generic level of feedback to inform the industry stakeholders via channels stated in Figure 3 instead of individuals related to each case. This specific type is defined as an *open-loop* VSR program in this research.

Figure 3

Channels Used by ASRS System in Providing Feedback to Industry Stakeholders

April 1976 – December 2020

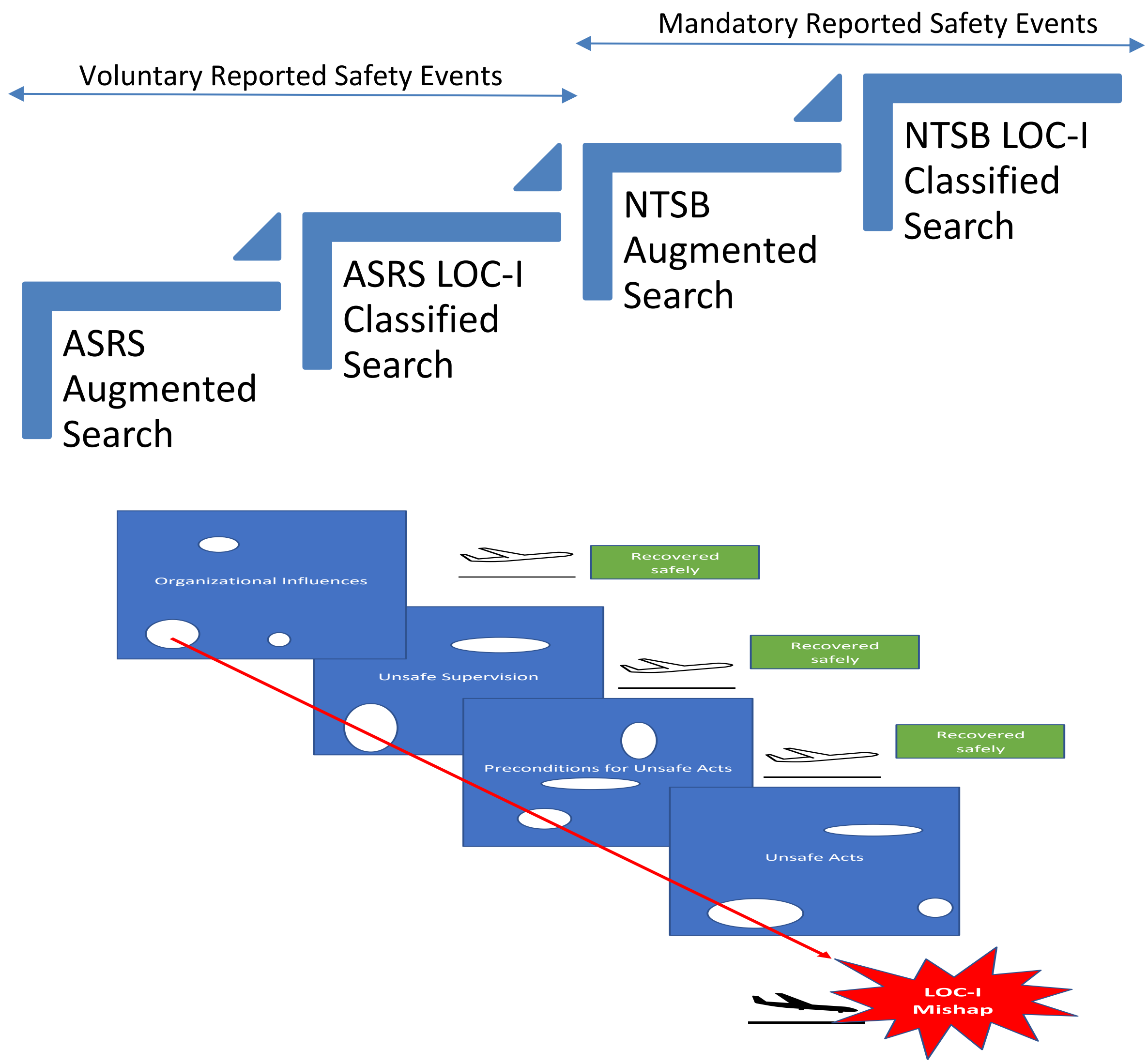
Significant Items	Quantity
Incident Reports Received	1,799,274
Safety Alert Messages	6,795
Quick Responses	144
Search Requests	7,591
CALLBACK Issues	491
ASRS Directline Issues	10
Research Studies	64

Note. Adapted from the NASA ASRS Program Brief (https://asrs.arc.nasa.gov/docs/ASRS_ProgramBriefing.pdf). Copyright 2020 by NASA.

This research used augmented search reports based on keywords deployed by Belcastro et al. (2017). This search method has been published in peer-reviewed journals (Belcastro et al., 2012; Belcastro et al., 2014; Belcastro et al., 2016; Belcastro et al., 2017; Kwatny et al., 2013; Tekles et al., 2017) and was shown to have added granularity and volume of information from the relevant databases. This search method led to increased sample sizes and additional information related to LOC-I. Purely basing research on classified LOC-I reports would forgo the opportunity to obtain the proactive data hidden in the relevant databases. An assumption was made that the augmented-searched cases were less severe than the LOC-I classified cases within the same database. The augmented search identified cases with one or more LOC-I precursors. These were cases that did not lead to a full LOC-I event with a more severe consequence. Otherwise, the safety reporting database administrators would have classified them as LOC-I. Regarding the scale of consequence severity, using the classification of a safety incident and accident in ICAO (2016), NTSB LOC-I Classified group would be the highest severity events, followed by NTSB Augmented, ASRS LOC-I Classified, and ASRS LOC-I Augmented. This was a logical deduction based on the causation chain theory by Reason (1990) that would require further validation in this context for future research (see Figure 4).

Figure 4

Increasing Severity of Four Groups of LOC-I Safety Report Types (IVs) Deployed in this Study with Illustration of Reason's (2016) Accident Causation Model



Note. Adapted from *Managing the Risks of Organizational Accidents*, by James Reason. Copyright 2016 by James Reason.

It was not the purpose of this study to identify why a difference in hazard rate exists between voluntary and mandatory reports. However, the analysis has identified such differences, and recommendations have been made to verify the rationale behind them. On the qualitative analysis supplement, the information in the narrative sections could vary within an individual safety database, as reporters might include varying levels of detail due to reasons and biases mentioned in the literature review. However, the supplement provides an opportunity to unveil factors related to LOC-I events that were embedded in the narratives but have not yet been coded, providing insights into the quantitative results.

Lastly, it was not the purpose of this research to assess the accuracy of each case and consistency of factors classification with the safety reports databases' administrators. The purpose was to identify if the Belcastro LOC-I Hazards were the same across the four types of safety reports. The result expands the body of knowledge in the practical and theoretical contributions highlighted in this chapter.

Definitions of Terms

Accident	ICAO defines an accident as an occurrence associated with the operation of an aircraft that takes place between the time any person boards the aircraft with the intention of flight until such time as all such persons have disembarked, in which: i. A person is fatally or seriously injured ii. The aircraft sustains damage or structural failure iii. The aircraft is missing or is completely inaccessible (ICAO, 2016).
Augmented searched LOC-I report	Reports not classified as LOC-I originally in the ASRS or NTSB database but contained LOC-I precursors per

	Belcastro et al. (2018) and have been identified by a text search.
Belcastro LOC-I Hazards	Eight hazards identified by Belcastro (2017) that lead to LOC-I events: • Adverse onboard conditions - Vehicle Impairment • Adverse onboard conditions - System and components failure / malfunction • Adverse onboard conditions - Crew action / inaction • External hazards and disturbances - Inclement weather atmospheric disturbances • External hazards and disturbances - Poor visibility • External hazards and disturbances - Obstacle • Abnormal vehicle dynamics and upsets - Abnormal vehicle dynamics • Abnormal vehicle dynamics and upsets - Vehicle upset conditions
Classified LOC-I reports	LOC-I events already classified by the ASRS or NTSB administrators, which are searchable from the respective databases.
Flight	The operation of an aircraft on a stage from taxi to landing or number of flight stages with the same flight number (ICAO, 2009).

Hazard	A condition or an object with the potential to cause death, injuries to personnel, damage to equipment or structures, loss of material, or reduction of the ability to perform a prescribed function (ICAO, 2013).
Hazard Rate	Particular Belcastro LOC-I Hazard Count over one calendar year divided by the number of flight hours flown for the particular operational certification for that particular year.
Incident	An occurrence, other than an accident, associated with the operation of an aircraft that affects or could affect the safety of operation (ICAO, 2016).
Loss of Control In-Flight	An event which may become unrecoverable if no intervention is made that fulfills at least one of the following criteria: • Outside normal envelopes (adjusted for flight phases) • Not predictably altered by pilot control inputs (i.e., aircraft response is no longer predictable to the pilot) • Characterized by nonlinear effects that degrade handling qualities • Kinematic/inertia coupling • Disproportionately large responses to small state variable changes • Oscillatory/divergent behavior • Likely to result in high angular rates/displacements

	• Characterized by the inability to maintain heading, altitude, and wings-level flight • The flight path is outside acceptable tracking tolerances and cannot be predictably controlled by the pilot (or auto-flight system inputs).
Open-Loop Voluntary Safety Report	A safety reporting system that has comparatively little investigation, verification, and feedback to the originators compared with a closed-loop system.
Serious Incident	An incident involving circumstances indicating that an accident nearly occurred. The difference between an accident and a serious incident lies only in the result (ICAO, 2016). Examples of serious incidents are listed in Appendix C.

List of Acronyms

AC	Advisory Circular
AD	Airworthiness Directive
AIDS	FAA Accident and Incident Data System
ALARP	As Low As Reasonably Practicable
AOA	Angle of Attack
ASAP	Aviation Safety Action Program
ASIAS	Aviation Safety Information Analysis and Sharing
ASRP	Aviation Safety Reporting Program
ASRS	Aviation Safety Reporting System

BTS	Bureau of Transportation Statistics
B737-8	Boeing 737-8 Airliner (formerly branded as 737 MAX 8)
CAA	Civil Aviation Authority
DV	Dependent Variable
EAIB	Ethiopian Airplane Accident Investigation Bureau
EASA	European Aviation Safety Agency
FAA	Federal Aviation Administration
FAR	Federal Aviation Regulation
FOIA	Freedom of Information Act
GADM	Global Aviation Data Management
GASP	Global Aviation Safety Plan
ICAO	International Civil Aviation Organization
IV	Independent Variable
LOC-I	Loss of Control In-Flight
MANOVA	Multivariate Analysis of Variance
MOU	Memorandum of Understanding
MSR	Mandatory Safety Report
NAA	National Aviation Authority
NASA	National Aeronautics and Space Administration
NTSB	National Transportation Safety Board
OEM	Original Equipment Manufacturer
SAR	Special Administrative Region
SARP	Standards and Recommended Practices

SME	Subject Matter Expert
SMS	Safety Management System
VSR	Voluntary Safety Reporting or Voluntary Safety Report

Chapter II: Review of the Relevant Literature

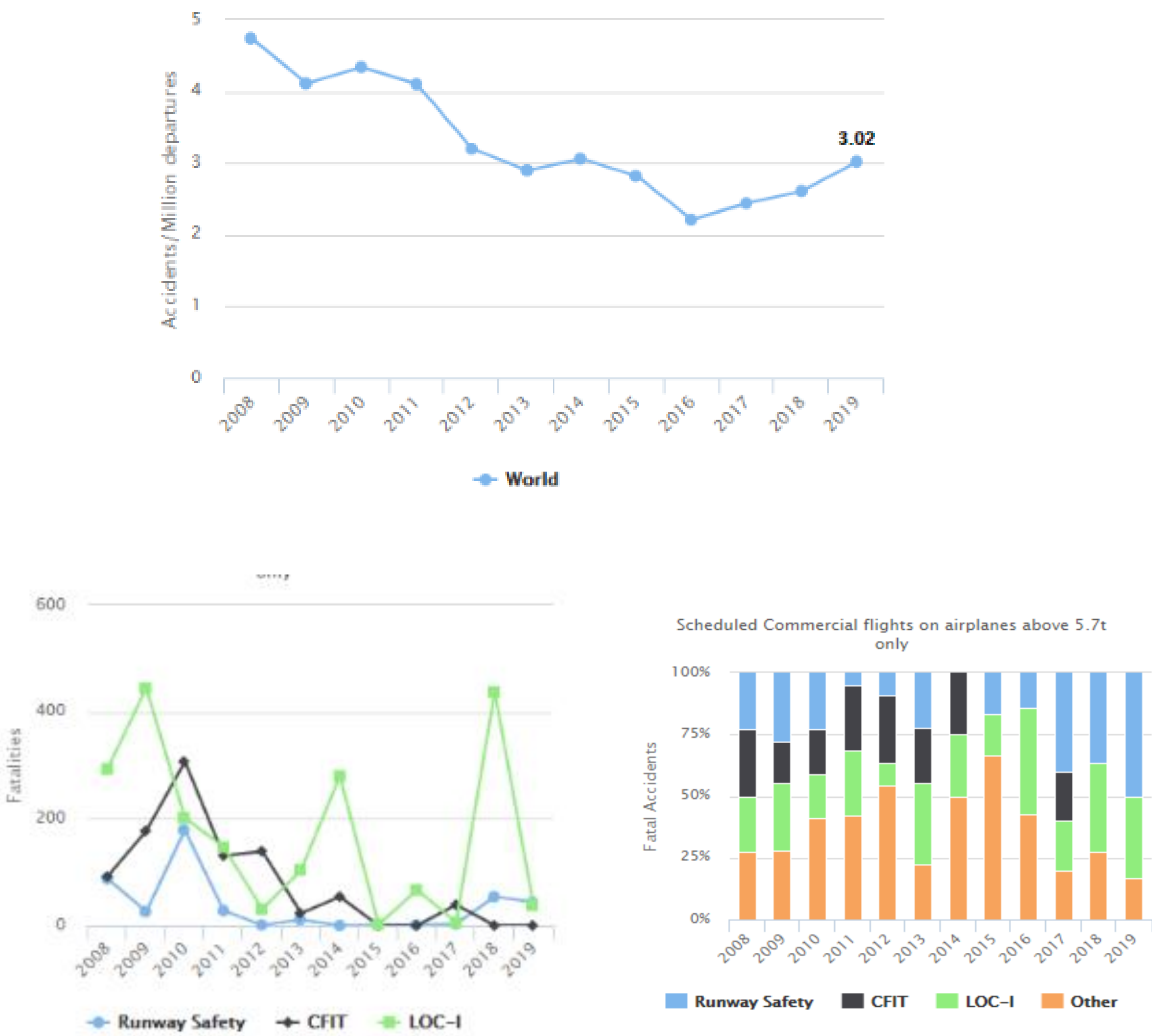
This chapter identifies the extant research and literature relevant to voluntary safety reporting as a key input to aviation SMS. The relevance of safety reporting to safety performance, critical hazards, the influences from Heinrich's theories on accident causation, and SMS strategies for mitigating identified hazards are discussed. Also explored are the opposing views on the relevance of Heinrich's principles in safety reporting, the caution against reliance on lower severity hazard mitigation to prevent events of high severity in other safety-critical industries, and the relation of safety reporting to reduce critical aviation safety hazards leading to LOC-I. Finally, gaps in the literature leading to the research questions are identified.

Safety Performance in Modern Aviation

Modern commercial aviation is arguably the safest form of transport (Lower et al., 2016; Valdes, 2011). In commercial aviation history, 2017 was a record year with zero LOC-I fatal accidents or hull losses reported among member airlines of the International Air Transport Association (IATA, 2018). That year, the global accident rate was 1.08 accidents per million departures, only half the rate recorded in 2015 (ICAO, 2016a). However, despite the low fatality rate in 2017, the general accident rate has risen since 2016, reaching 3.02 accidents per million departures in 2019 (see Figure 5). Moreover, the world marked four fatal commercial aviation accidents in 2018 due to LOC-I, controlled flight into terrain, and runway safety events (ICAO, 2020). Despite focused accident prevention efforts, one hull loss due to LOC-I was almost a yearly occurrence in commercial aviation worldwide. This frequency was further exacerbated by the introduction of the Boeing 737 MAX 8 in 2018, resulting in the loss of two hulls and 438 lives in LOC-I accidents. It is unclear whether the current reductions in accident rates will continue, or the industry's safety performance has plateaued (see Figure 5).

Figure 5

ICAO Accident Statistic Graphs



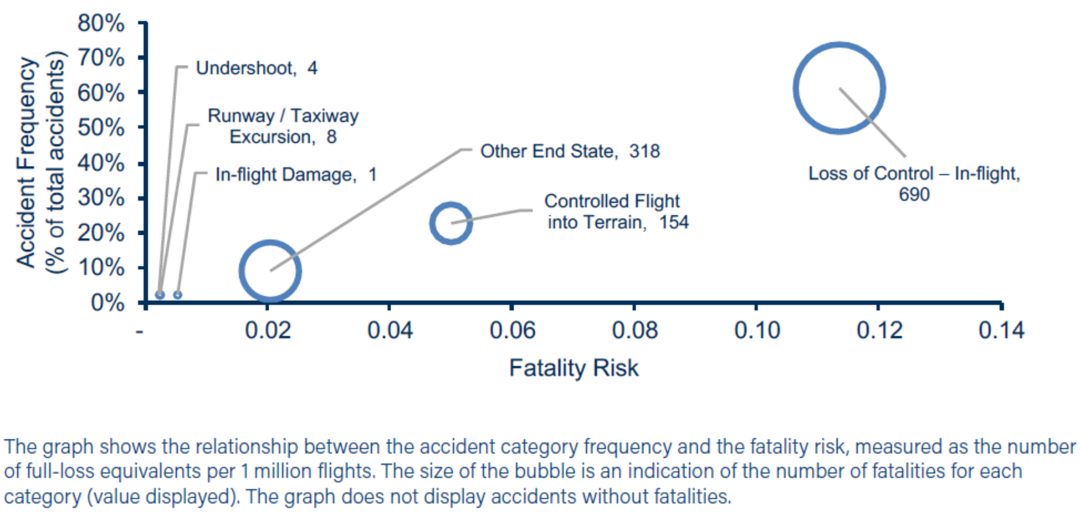
Note. Counterclockwise from top: graphs for Global Accident Rate, Fatalities by Risk Category, and Share of Fatal Accidents by Risk Category. Reprinted from ICAO Accident Statistics (https://www.icao.int/safety/iStars/Pages/Accident-Statistics.aspx). Copyright 2020 by ICAO.

LOC-I Events

IATA has described LOC-I events as one of three aviation accident categories that accounted for all the deaths in aviation catastrophes (IATA, 2018), the other two being controlled flight into terrain and runway undershoot events (see Figure 6). According to the 2018 IATA Safety Report, LOC-I accidents resulted in 926 fatalities from 2014 to 2018, of which 372 occurred in 2018 alone. In the same year, while LOC-I events represented only 6% of accidents, they accounted for 71% of onboard fatalities (IATA, 2019).

Figure 6

Flight Accident Category Frequency and Fatality Risk 2013–2017



Note. Reprinted from the IATA 2018 Safety Report (https://www.iata.org/en/publications/safety-report/). Copyright 2018 by IATA (p. 44).

The most recent notable cases of LOC-I involved two Boeing 737 MAX 8 aircraft, operated by Lion Air and Ethiopian Airlines, both of which crashed during their initial climbs on scheduled flights (IATA, 2019). The related accident investigation reports have been published (EAIB, 2022; KNKT, 2019). The findings regarding the Lion Air case suggested that the accident was caused by a miscalibrated angle of attack (AOA) sensor, which triggered an augmentation function similar to a stick shaker. The function was embedded in the maneuvering characteristics augmentation system (MCAS). It forced the aircraft to pitch down constantly to prevent an anticipated stall. What transpired was not the design intention, as the MCAS was supposed to command pitch down once. However, because of the false signal of one AOA, the aircraft was commanded to pitch down again. The onboard response was complicated by the first officer's unfamiliarity with the procedure for disengaging this erroneously-activated feature. The design, maintenance, training, and certification of the B737-8 were identified as contributing factors to the event. The findings from the Ethiopian Airlines investigation were similar.

Definition of LOC-I

Belcastro et al. (2017) define LOC-I as:

> Motion that is outside the normal operating flight envelopes; not predictably altered by pilot control inputs; characterized by nonlinear effects, such as kinematic/inertial coupling; disproportionately large responses to small state variable changes or oscillatory/divergent behavior; and likely to result in high angular rates and displacements: it is characterized by the inability to maintain heading, altitude, and wings-level flight. LOC-I also includes situations in which the flight path is outside of acceptable tracking tolerances and cannot be predictably controlled by pilot (or auto-flight system). (p. 737)

Similarly, the ICAO Commercial Aircraft Safety Team (CAST) defines LOC-I as a significant deviation of the aircraft from the intended flight path or operational envelope (Russell & Pardee, 2000).

Much research has been conducted on the hazards that lead to the occurrence of a LOC-I event, as well as associated mitigation strategies. Belcastro et al. (2017) further characterize LOC-I as an event that is not necessarily unrecoverable but can become unrecoverable if no appropriate intervention is made. A LOC-I event thus fulfills *at least one* of the following criteria (Belcastro et al., 2017, p. 737):

- Outside normal envelopes (adjusted for flight phases)
- Not predictably altered by pilot control inputs (i.e., aircraft response is no longer predictable to the pilot)
- Characterized by nonlinear effects that degrade handling qualities:
    - Kinematic/inertia coupling
    - Disproportionately large responses to small rate variable changes
    - Oscillatory/divergent behavior
- Likely to result in high angular rates/displacements
- Characterized by the inability to maintain heading, altitude, and wings-level flight
- The flight path is outside of acceptable tracking tolerances and cannot be predictably controlled by pilot (or auto-flight system inputs)

Factors Contributing to LOC-I: A 15-year NASA Study

Extensive research has been conducted to identify the factors leading to the onset of LOC-I accidents. The International Committee on Aviation Training in Extended Envelopes (ICATEE, n.d.) identified aerodynamic stall, flight control system failures, spatial disorientation, icing, and atmospheric disturbance as major contributing factors. One of the most significant LOC-I accidents in the United States was the Colgan 3407 accident, where an aerodynamic stall occurred. The accident resulted in the death of all 49 passengers and flight crew on board, as well as an individual in a house into which the aircraft crashed.

In the late 2000s, a team of NASA, NTSB, and industry experts formed the LOC-I Research Working Group. Examining a total of 278 LOC-I mishaps, accidents, and incidents from 1996 to 2010 documented by seven air accident investigation authorities and four aviation safety databases, the group identified a series of precursors and hazards from the dynamics and control perspective that led to the onset of LOC-I (Belcastro et al., 2010, 2012, 2014, 2016, & 2017). Table 1 details the research papers on LOC-I published by NASA during the 15 years from 2004 to 2017.

Table 1

NASA Research Publications on LOC-I

Reference	Title	Summary
Wilborn and Foster, 2004	Defining Commercial Transport Loss-of Control: A Quantitative Approach	Development of a set of metrics for defining LOC-I. Covers airplane flight dynamics, aerodynamics, structural integrity, and flight control use.
Belcastro and Foster, 2010	Aircraft Loss of Control Accident Analysis	Review of 126 LOC-I accidents from 1979 to 2009. Identification of worst-case combinations of causal and contributing factors. A detailed compilation of 52 LOC-I sequences.
Belcastro and Jacobson, 2010	Future Integrated Systems Concept for Preventing Aircraft Loss of Control Accidents	Presentation of future system concepts and research directions for preventing LOC-I accidents. Based on a generalized LOC-I accident sequence, the S-Factor concept on the stability matrix is discussed. A holistic aircraft-integrated resilient safety assurance and failsafe enhancement (AIRSAFE) system is proposed.
Belcastro, 2012	Validation of Safety-Critical Systems for Aircraft Loss of Control Prevention and Recovery	Based on previous research on LOC-I sequences, causal and contributing factors, provision of NASA's validation methods and tools within the Vehicle Systems Safety project, and detailing a preliminary set of test

Reference	Title	Summary
		scenarios for validation of technologies for LOC prevention and recovery.
Belcastro, Goff, Newman, Foster, Crider, Klyde, and Huston, 2014	Preliminary Analysis of Aircraft Loss of Control Accidents: Worst Case Precursor Combinations and Temporal Sequencing	Defines a comprehensive set of LOC-I accidents and incidents from 1996 to 2010. Presents a preliminary analysis of worst-case combinations of causal and contributing factors and their temporal sequences.
Belcastro, Foster, Shah, Gregory, Cox, Crider, Groff, Newman, and Klyde, 2017	Aircraft Loss of Control Problem Analysis and Research Toward a Holistic Solution	Summary of the body of research conducted by NASA to develop a holistic solution for LOC-I hazards. Captures the identification of accident precursors and sequences using a team approach, and analyzes individual precursor contributions, worst-case hazard combinations, and worst-case sequences relative to the resulting number of accidents and fatalities. Provides scenarios for testing technological mitigation strategies such as onboard systems.

Table 2 presents the primary causes of LOC-I, precursor, or hazard categories leading to LOC-I events.

Table 2

Primary Causes, Precursors, and Hazards of LOC-I Events

Primary Causes	**Precursor/Hazard Categories and Subcategories**
• Entry into vehicle upset condition (e.g., stall) • Reduction or loss of control effectiveness • Changes to vehicle dynamic response and handling/flying qualities (including asymmetric effects) • Combinations of the above	Adverse onboard conditions: Vehicle impairment • System faults, failures, and errors • Inappropriate crew action/inaction External hazards and disturbances: • Inclement weather and atmospheric disturbances • Poor visibility • Obstacle Abnormal dynamics and vehicle upsets: • Abnormal vehicle dynamics and control response • Abnormal attitude, airspeed, angular rates, asymmetric forces, or flight trajectory • Uncontrolled descent (including spiral dive) • Stall/departure from controlled flight

Note. Adapted from Aircraft Loss of Control Problem Analysis and Research Toward a Holistic Solution by Belcastro et al., 2017, *Journal of Guidance, Control, and Dynamics* American Institute of Aeronautics and Astronautics. Copyright 2017 by AIAA.

Belcastro et al. (2010) further summarized the various LOC-I temporal sequences into ten generic ones, emphasizing the level of complexity and the importance of the temporal sequence to the onset of a LOC-I event (see Figure 7).

Figure 7

Generic LOC-I Accident Sequences

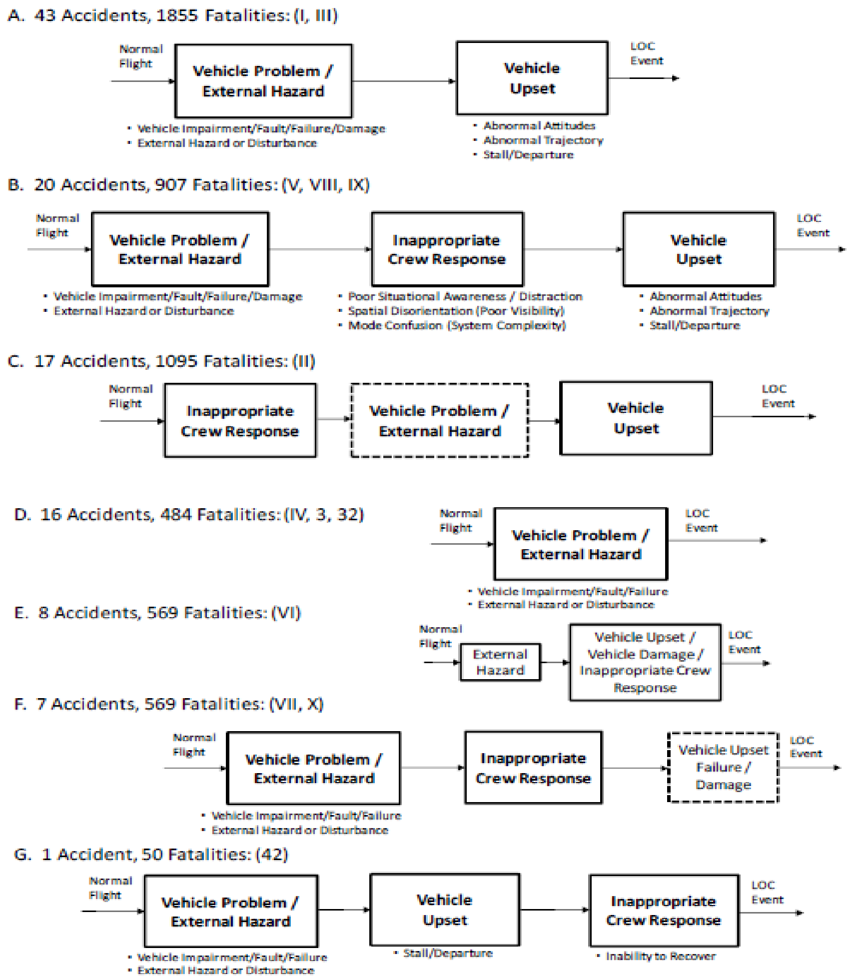


Figure 18. Generalized LOC Accident Sequences.

Note. Adapted from Aircraft Loss-of-Control Accident Analysis (p. 11) by C. Belcastro and J. Foster from *American Institute of Aeronautics and Astronautics*, p. 11. Copyright 2010 by NASA. Reprinted with permission.

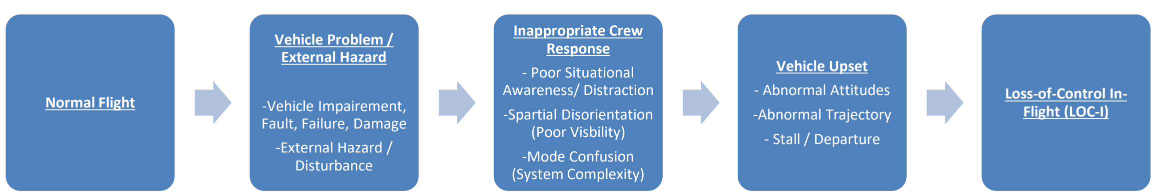
The sequences above were simplified in Belcastro et al. (2017) into one generic sequence, as illustrated in Figure 6, except for Sequence C in Figure 7, which begins with an *inappropriate crew response*, such as incorrectly setting the automation. The findings by Belcastro et al. (2017) indicated that a LOC-I event was typically preceded by three generic precursors/hazard categories, namely (see Figure 8):

1. Vehicle problem/external hazard
2. Inappropriate crew response
3. Vehicle upset

The NASA LOC-I study by Belcastro and her research team has been published progressively in various scholarly forums and peer-reviewed papers, including the Guidance, Navigation, and Control Conference, the Atmospheric Flight Mechanics Conference, and the Modeling and Simulation Technologies Conference in 2005.

Figure 8

Simplified Generic LOC-I Model



Note. Adapted from Aircraft loss of control problem analysis and research toward a holistic solution by Belcastro et al. from *Journal of Guidance, Control, and Dynamics, 40(4), 733-775.*Copyright 2017 by NASA. Reprinted with permission.

Mitigation Strategies for LOC-I

The aviation industry has undertaken various efforts to mitigate the onset of LOC-I events. Addressing the general aviation sector, Balogh (2006) conducted a LOC-I study based on flight data, highlighting the importance of AOA monitoring in preventing aerodynamic stalls. From the perspective of organizational management and aircraft design and manufacturing, IATA has published its guidance on LOC-I mitigation (IATA, 2015), addressing vehicle problems and inappropriate crew responses.

Summarizing the 15-year NASA study analyzing the causal and contributing factors of LOC-I, the paper authored by Belcastro et al. (2017) represents a collaborative approach between industry, government, and academia to guide the industry toward mitigating LOC-I events in the short, medium, and long term. The approach focuses on detecting vehicle problems and external hazards, mitigating inappropriate crew responses, and recovering from vehicle upsets. Preventive mitigation has also been applied through improving crew training under LOC-I precursor conditions to elevate their awareness of LOC-I. This work is being used as one of the blueprints for implementing NextGen (Petitt, 2017). Based on their precursors and hazards analysis, NASA projected the need to build a comprehensive set of LOC-I test scenarios to evaluate the resilience of the deployed mitigation technologies. Three technology development areas have been identified (Belcastro et al., 2017, pp. 744-755):

a. Dynamic vehicle modeling and simulations for LOC-I effects characterization
b. Onboard systems for LOC-I prevention and recovery
c. Validation of mitigation technologies under realistic LOC-I conditions

It is to be noted that Belcastro's team at NASA used mandatory safety reports from accident investigations to provide the data required for the research. While accidents or serious LOC-I incidents do not occur regularly, it is not known if LOC-I events of less severity from a VSR system can obtain similar results.

The impetus for the Introduction of Aviation VSR in the United States

To observe the evolution of VSR employed in the U.S., Mills (2011) analyzed the macro- and micro-level aspects of the country's civil aviation regulatory environment throughout its modern aviation history. In this research, he identified a shift from a command-and-control regulatory style to an industry-regulator partnership supported by a voluntary reporting system.

Traditional Command-and-Control Approach Adopted by the FAA

Traditionally, regulatory authorities adopted a command-and-control approach to managing airline safety. Under this regulatory approach, the development of rules, standards, penalties, and enforcement mechanisms shapes the behavior of firms and individuals alike. Standards were typically implemented by granting government licenses, permits, or certificates. Once these standards were established, regulators developed penalties, such as fines and suspensions, to deter companies from violating rules and standards. The strength of this regulatory approach is that expected behaviors are clearly defined, making it easy to enforce laws and identify breaches in legal standards (Gunningham & Grabosky, 1998).

The regulatory approach adopted by the FAA prior to the 1970s was primarily dependent on enforcement. This approach included conducting inspections, issuing mandatory Advisory Circulars (ACs) and Airworthiness Directives (ADs), and releasing instructions requiring inspections of any modifications to previously certified aircraft. The data collected from such inspections informed reactive enforcement actions based on Federal Aviation Regulations (FARs) established in Aeronautics and Space (2012), which is Title 14 of the Code of Federal Regulations (14 CFR).

The Birth of ASRS

The command-and-control regulatory approach adopted by the FAA was not without weaknesses. Mills (2011, p. 28) summarized them as follows:

a. No inspection program can detect all violations at all times because inspection resources are always limited (Iannuzzi, 2002). Regulatory programs are generally considered to have extensive enforcement systems involving an army of inspectors. In reality, enforcement relies heavily on voluntary reporting by regulated entities and infrequent inspections (May, 2002).

b. Regulated entities often engage in calculated compliance, weighing the costs and risks of getting caught against the benefits of compliance (Salamon, 2002).

c. Inspection programs require regulators to have comprehensive and accurate knowledge of the operations and capacities of the industry.

d. Compliance-based oversight lacks incentives for firms to go beyond minimum standards and may ultimately result in reduced compliance with rules (Gunningham & Grabosky, 1998).

e. Increasing administrative complexity vis-à-vis the sheer volume of statutes and regulations, makes it difficult for regulators and industry personnel alike to comply with the law.

The crash of TWA Flight 514 on December 1, 1974, outside Mount Weather, VA, marked a turning point in the FAA's regulatory approach. Due to a misinterpretation of an approach chart, the inbound flight to Dulles Airport descended below the minimum safe altitude. It collided with a Virginia mountaintop, killing 85 passengers and seven crew members on board (Reynard et al., 1986). Ironically, the same hazard was reported and disseminated within United Airlines through its *Flight Safety Awareness Program* safety sharing platform; however, the system was not made available to the rest of the industry and the federal government. As a result of the crash, the FAA implemented the Aviation Safety Reporting Program (ASRP)—a confidential, voluntary, and non-punitive reporting system—in May 1975 (FAA, 2011), offering waivers of sanctions and anonymity to those who made reports through the program. To reinforce trust, the FAA also signed a Memorandum of Agreement (MOA) with NASA that delegated the administration of its ASRP reporting system (ASRS) to NASA as an independent broker. The result was the first nationwide, government-sponsored aviation VSR system in the United States

ASRS is still operating to this date. Referring to the Program Briefing document issued by NASA and posted on the ASRS website, the purpose of the program is to "collect, analyze, and respond to voluntarily submitted aviation safety incident reports in order to [emphasis added] *lessen the likelihood of aviation accidents" (p. 1).* (NASA, 2022). The program has received 1.7 million reports from January 1981 to December 2019. In 2019 alone, 107,879 VSR reports were received. After the report validation process documented in Figure 2, short of individual follow-up, findings from ASRS reports would be fed back to the industry by the following means:

a. Alert Messages – Safety information issued to organizations in positions of authority for evaluation and possible corrective actions.

b. Quick Responses – Rapid data analysis by ASRS staff of safety issues with immediate operational importance generally limited to government agencies.

c. ASRS Database – The public ASRS database online and data available in Database Report Sets or Search Requests fulfilled by ASRS staff.

d. Callback Newsletter – Monthly newsletter with a lessons-learned format, available via website and email.

e. Focused Studies – Studies / Research conducted on safety topics of interest in cooperation with aviation organizations.

ASRS is a repository based on crowdsourcing of voluntarily submitted safety reports (Schnittker et al., 2020). It is to be noted that no formal investigation will be carried out upon submission of ASRS reports. If necessary, the administrator will telephone the originator to clarify the information provided (NASA, 2022b).

Development of VSR after ASRSs

Since the introduction of ASRS in the 1970s, VSR systems in the United States have undergone various stages of development. Following serious accidents in the mid-1990s, such as USAir Flight 427 and ValuJet Flight 537, the effectiveness of the FAA's reactive and mandatory enforcement approach was questioned (Gore, 1997). In 1996, President Clinton established the White House Commission on Aviation Safety and Security, intending to reduce aviation fatalities. The work of the Commission led to the birth of the FAA Air Transportation Oversight System (ATOS), which fundamentally shifted aviation regulation toward a systems-based approach. Under ATOS, each airline is required to establish a surveillance plan based on data analysis and risk assessments (GAO 2006, as cited in Mills, 2011), reinforcing the data-driven focus of the regulatory approach.

Voluntary Disclosure Reporting Program (VDRP) in the Mid-1990s

In the mid-1990s, in responding to calls from air carriers to ease enforcement actions and allow the voluntary disclosure of violations in exchange for reduced penalties, the FAA established a VDRP system under the direction of Admiral James Busey (Mills, 2011). The VDRP offers reduced regulatory enforcement actions for certificate-holding air carriers if they voluntarily report systemic problems within their operations and work collaboratively with their local FAA Certificate Holding District Offices (CHDO) on designing the resolutions to those issues. For companies that self-disclose apparent violations through the VDRP scheme and fully implement resolutions agreed upon by their local CHDO, any enforcement is carried out through administrative action, such as letters of correction, instead of legal action, such as civil penalty fines. Furthermore, all data released in the VDRP scheme per 14 CFR Part 193 is protected from exposure to the public under the Freedom of Information Act (FOIA). Since December 2006, the FAA has been operating a web-based system for VDRP submissions by major air carriers (Mills, 2011).

Aviation Safety Action Program (ASAP) in the 2000s

As the next evolutionary step in safety management, the aviation industry introduced the concept of a risk-based approach to managing aviation safety through the implementation of SMS in the early 2000s (Stolzer, 2017). Although VSR programs are one of the primary sources for risk and hazard identification, as Mills (2011) suggested, one disadvantage of the ASRS system is that the de-identified nature of the data recorded in ASRS cannot support risk-based inspections for specific air carriers. To address this, the FAA has implemented the Aviation Safety Action Program (ASAP) to partner with participating air carriers. Such a system provides a regulatory incentive for air carriers and other industry employees to submit reports of violations voluntarily. ASAP involves a partnership between three entities, namely the FAA, individual air carriers, and employee unions, codified through a memorandum of understanding (MOU). The FAA first published guidance on the ASAP program, particularly for its data protection elements, in 2002 through the release of AC 120-66B. As stipulated in the circular, each ASAP report is reviewed by an event review committee (ERC) to decide whether it should be accepted by the program and what corrective actions must be taken.

After the Colgan Air Flight 3407 accident outside of Buffalo, NY, the FAA encouraged carriers to implement ASAP and FOQA programs. This call demonstrated the administration’s increased reliance on information collected from VSR systems. Given the rapid advances in the National Airspace System and its associated spectrum of technologies, it is inevitable that the FAA will not be adequately equipped with the range of SMEs and safety information sources necessary countrywide to continue safeguarding safety using a directive approach, without first acquiring data from operational communities. Figure 9 summarizes the evolution of the civil aviation regulatory environment in the United States.

Figure 9

Evolution of Aviation Safety Reporting Systems in the United States and Its Regulatory Implications

Note. Adapted from Collaborating with Industry to Ensure Regulatory Oversight: The Use of Voluntary Safety Reporting Programs by R. Mills. Copyright 2011 by Kent State University.

Representativeness of ASRS in VSR

Mills (2011) indicated that a benefit of ASRS is the duplicates of many de-identified ASAP reports that it contains. As ASRS is a public database, the FAA can commission NASA to conduct database analyses without requiring approval from an external board. Since establishing ASRS in 1976, key stakeholders such as the FAA, industry, NASA, the Government Accounting Office (GAO), and Congress have regularly requested ASRS to conduct analyses based on de-identified data. In academic literature, a search for research dissertations with the keywords *aviation safety reporting system* or *ASRS* identified 40 dissertations and theses published in the past five years. Table 1 lists the relevant research publications that have used ASRS as a dataset, all of which have successfully passed the validity and reliability requirements for their research purposes, as documented in Appendix A. Among VDRP and ASAP programs, ASRS is designed with minimal individual follow-up and investigation. It is therefore described as an open-loop publicly available VSR in this regard.

Aviation Safety Incident, Serious Incident, and Accident Classifications

To define the classification of an aviation safety event according to its severity, ICAO published Annex 13, a Standards and Recommended Practices (SARP) document related to aircraft accident and incident investigations was issued (ICAO, 2016). Annex 13 defines three event classifications in ascending severity: *incident, serious incident*, and *accident*. Exact definitions for these classifications are documented in the Definitions of

Terms section and reproduced in Appendix B. Furthermore, the SARPs listed under Annex 13 clearly state that an investigation's sole objective is the prevention of accidents and incidents and not to apportion blame or liability (ICAO, 2016). Under Annex 13 protocols, the state of occurrence is responsible for launching an investigation into an accident or serious incident; however, the state of occurrence may delegate, wholly or in part, such an investigation to another state or a regional accident and incident investigation organization. For example, following the October 3, 2017, incident in which the fourth engine of an Airbus A380 failed while flying over Greenland, the Danish Accident Investigation Board delegated the conduct of the investigation to the French Air Accident Investigation Authority (Bureau d’Enquêtes et d’Analyses pour la Sécurité de l’Aviation Civile) (BEA, 2020). Most investigation authorities publish preliminary and final reports to share safety information.

The Birth of SMS

SMS was introduced to safety management in modern aviation during the early 2000s. ICAO (2013) described the accurate and timely reporting of relevant information related to hazards, incidents, or accidents as a “fundamental activity of safety management” (pp. 2-16). It also recognized direct reporting by front-line personnel as the best data source, given that this group of personnel observes hazards as part of their daily activities; consequently, such personnel should be trained and encouraged to submit safety reports (ICAO, 2013). ICAO classifies safety reporting into hazard reporting and occurrence reporting; both support the safety risk management (SRM) and safety assurance (SA) processes of the SMS.

On a global scale, ICAO has established the Global Aviation Safety Plan (GASP; ICAO, 2016) that states the requirements for the implementation of SMSs by service providers—including aircraft, airport, air traffic management, and maintenance providers—that are overseen by the state safety programs (SSPs) of each member state. GASP emphasizes a strong safety reporting culture alongside effective safety oversight.

A mature safety management approach, such as the one established in GASP, requires the collection and application of data for predictive risk management. A drive to implement SMSs and associated safety reporting systems globally has occurred in response to GASP. For example, the FAA issued a mandate for the implementation of an SMS in the United States aviation industry by 2018 (FAA, 2017). In Europe, the European Aviation Safety Agency (EASA) established the European Plan for Aviation Safety (EPAS) to set up an aviation SMS for the European industry (EASA, 2017), identifying better EU-wide occurrence reporting data for NAAs as a deliverable for 2017. A review of the EASA website confirms that EASA has since established the European Aviation Reporting portal (http://www.aviationreporting.eu), as well as issued guidance on safety reporting for organizations and individuals through facilitating an internal occurrence reporting (IOR) system. Reports are submitted through the portal on mandatory and voluntary bases (EASA, 2017). Although regulatory immunity obtained from submission is not explicitly stated, EASA (n.d.) has stated that "the reported occurrence data will not be held against the reporting parties and will be used for the interest of aviation safety" (para.3). EASA also assures data protection for both internal and external parties handling the data, which is covered by various European regulations, including (EC) No. 1049/2001, Article 72 of (EC) No. 2018/1139, and (EC) No.

379/2014. Corresponding manifestations of VSR in aviation were also found in the United Kingdom, Australia, the Hong Kong SAR, and New Zealand through further research, summarized in Appendix C.

The Rise of SMS in Aviation

In parallel with the work conducted by Belcastro et al. (2017) on understanding and mitigating LOC-I, the concept of SMS continued to develop in the early 2000s. It is described as a "systematic approach to managing safety, including the necessary organizational structures, accountabilities, policies, and procedures" (ICAO, 2013, p.12). SMS transformed aviation safety management from a compliance-based approach to a performance-based one (Maurino, 2017). The introduction of SMS required airline management to monitor its operations and safety performance as an entire *system* consisting of people, hardware, software, and the environment (Stolzer, 2017). Hence, rather than a piecemeal approach to safety, SMS offers a management system based on the foundation of a quality management system.

In 2006, ICAO published Doc. 9859, its first guidance document for the aviation industry on SMS (ICAO, 2013). The guidance provided was based on the four-pillar philosophy for an SMS: safety policy, risk management, safety assurance, and safety promotion. Further guidance followed in 2013 in the form of a dedicated Annex to the Convention on International Civil Aviation, Annex 19 (ICAO, 2013). Many ICAO member states and entities have since ratified the SARPs in local legislation urging operators to implement SMSs, including the European Commission (European Commission, 2015) and the United States (FAA, 2015). Since 2018, SMS has become a mandatory safety requirement for U.S.-based airlines, regional air carriers, and cargo carriers operating under 14 CFR Part 121 (FAA, 2015). The FAA also encourages voluntary implementation of SMS for non-regularly scheduled air carriers, maintenance and repair organizations (MROs), and training organizations.

Under an SMS, an operator obtains knowledge of safety hazards and their associated risks through risk assessments. Risk mitigation measures are then applied to reduce risks to levels as low as reasonably practicable *ALARP* (Stolzer, 2017). As part of the quality loop, the organization's safety performance is measured by safety objectives and performance indicators. This information is typically obtained through safety assurance activities that form part of the SMS, including audits, inspections, and mandatory and voluntary safety reporting (Maurino, 2017; Stolzer et al., 2018).

Relevance of SMS in Managing Critical Hazards in Aviation Such As LOC-I

SMS provides the framework for operators to identify hazards, assess, and proactively mitigate risks. When harmonizing the European norms and standards on SMS, EASA has established a three-tier approach among the *SMSs* of operators, State Safety Program (SSP) and the State plan for Aviation Safety (SPAs) at the member states level, and the European Plan for Aviation Safety (EPAS) at European Level (EASA, 2023). EASA emphasized that each operator is responsible for the safety of its operation. Each operator’s SMS should address relevant EPAs or SSP / SPAs topics and the risks of their unique operating environment. In terms of managing critical hazards, EASA member states and their operators are required to focus on using SMS to manage five critical safety hazards in aviation below, as well as addressing the hazards unique to their environments (EASA, 2021). The critical hazards detailed in the EPAS also aligned with ICAO’s Global Aviation Safety Plan (GASP) (ICAO, 2022):

i. Runway excursion

ii. Mid-air collision

iii. Controlled flight into terrain

iv. Loss of control in flight (LOC-I), and

v. Runway incursion

Risk management is a key element of SMS (Stolzer et al., 2011). The COVID-19 pandemic led to a significant impact on aviation demand (Truong, 2022). A study conducted by Cranfield University identified an association between the COVID-19 pandemic and flight data monitoring exceedances (Li et al., 2022). Some of such exceedances are related to precursors of critical hazards, including LOC-I. The study highlighted risks of manual flying skill decay, lack of practice effects on using standard operating procedures, and reduced knowledge of flight deck automation should be further assessed, monitored, and mitigated by operators' SMSs.

Safety Reporting System for an Airline's SMS

In a discussion paper presented at the International Transport Forum in 2017, Maurino (2017) described that "effective safety reporting relies to a large degree on the voluntary reporting of experiences by people who *operate the system*" (p. 46). The paper continued to describe safety reporting as the centerpiece of SMS data collection processes informing management decisions, in addition to evaluating employee safety reporting as "the single most valuable activity for safety data collection under SMS" (p. 56).

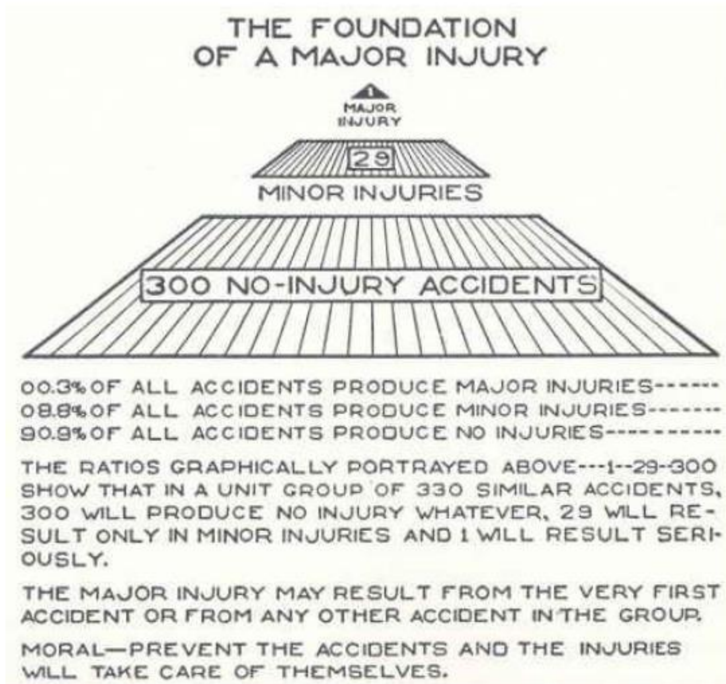
The VSR system forms part of the risk management and safety assurance elements of an SMS. Airlines administer VSR databases to collect hazard data from relatively low-severity events, expecting higher severity events to be prevented, per Heinrich's common cause hypothesis.

Relevance of Heinrich's Theories to SMS and VSR

Among the cornerstones of safety management, the theories attributed to Heinrich include the domino theory, Heinrich's triangle (or Heinrich's pyramid), and the common cause hypothesis (Davies et al., 2017). In particular, the common cause hypothesis suggested that safety events with more severe consequences shared the same causes as those with less severe consequences. Heinrich's triangle, an application of the hypothesis, postulates that a reduction in no-injury incidents leads to reductions in minor and major injury incidents (see Figure 10). Heinrich supported the notion that mitigating less severe events would prevent more severe events from occurring (Davies et al., 2017).

Figure 10

Heinrich's Triangle—An Application of the Common Cause Hypothesis



Note. Reprinted from H.W. Heinrich, 1931, *Industrial accident prevention: A scientific approach*, McGraw-Hill. Copyright 1931. Reprinted with permission.

Among the theories in his book detailing his research on insurance claims data in the 1930s, the common cause hypothesis behind Heinrich's triangle was significant for suggesting that mitigating less severe safety events, typically reported in VSRs, could mitigate more severe events, and vice versa. This hypothesis propelled the development of behavior-based safety (Basford, 2017), which focuses on identifying and treating front-line safety behavior discrepancies. Many safety initiatives in the occupational safety and health domain are based on this hypothesis, given the strong emphasis on identifying hazards of any severity level in the field, as well as collecting and analyzing reports on near-miss events with minor consequences (Davies et al., 2003). Heinrich's theories are mentioned in textbooks for prospective and practicing safety practitioners (Davies & Ebrary, 2003; Jeelani et al., 2018; McKinnon, 2017; Stolzer et al., 2017). They have also played a guiding role in shaping the thinking on obtaining an organization's risk profile through implementing SMSs.

Mounting Challenges to Heinrich's Theories

Despite the significance of Heinrich's triangle as a rule of thumb, occupational safety and health professionals have raised concerns about whether the theory (Heinrich, 1931) is still relevant to the modern world (Manuele, 2011; Marshall et al., 2018). A cohort of scholars challenged the basis of Heinrich's triangle and the associated common cause hypothesis by questioning the validity of the claimed causal relationship between occurrences with minor consequences and occurrences with more severe outcomes (Manuele, 2011; Yorio & Moore, 2018).

In the first edition of his book, *Industrial Accident Prevention: A Scientific Approach*, based on his analysis of industrial insurance data in the 1930s, Heinrich (1931) expressed the relationship between the occurrences of no-injury, minor-injury, and major-injury accidents as a ratio of 300:29:1 (see Figure 10). Substantial research has since been conducted in occupational health and safety, as well as process safety, challenging whether Heinrich's works still apply to modern industries (Basford, 2017; Manuele, 2018; Marsden, 2018). For instance, as part of his attempt to validate the applicability of Heinrich's triangle, Basford (2017) analyzed occupational injury statistics from the U.S. Bureau of Labor Statistics. He compared injury and fatality rates and observed that the industrial sectors whose accident ratios closely aligned with Heinrich's triangle were construction, manufacturing, trade, transportation, and utilities; however, he also concluded that nine other industries displayed little or no alignment with Heinrich's model.

The key challenges that scholars have mounted toward Heinrich's work concern the following issues:

a. Heinrich's ratio was calculated based on accident numbers reported to insurance companies, which may not have represented the actual figures, particularly for those of lesser severity (Manuele, 2011).

b. It is unclear whether Heinrich's ratios are consistent across industries (Bellamy et al., 2008; Gallivan et al., 2008).

c. The oversimplification of accidents amid the desire to pinpoint the *unsafe act* at the worker's level neglects systemic workplace issues. This approach focuses too much on workers rather than management, leading to overemphasizing behavioral safety programs (Manuele, 2011).

d. Company management may become preoccupied with searching for and measuring low-severity events as their safety performance indicators, based on the statistically unsubstantiated myth that reducing casual factors for such events will reduce the probability of more severe events occurring (Manuele, 2018; Marsden, 2011).

e. The simplistic linear causation model may not apply to modern, complex organizational accidents such as the Deepwater Horizon oil rig accident (Barstow et al., 2010).

f. The premise that reducing the frequency of occurrences will reduce the severity of occurrences has not been statistically substantiated (Manuele, 2011; Marsden, 2018).

g. Based on insurance classifications in the 1930s, the definitions for each severity class of safety events in Heinrich's triangle differ from those adopted in modern occupational safety and health settings, including the aviation industry (Manuele, 2011).

Support for Heinrich's Theories by Modern Safety Practitioners

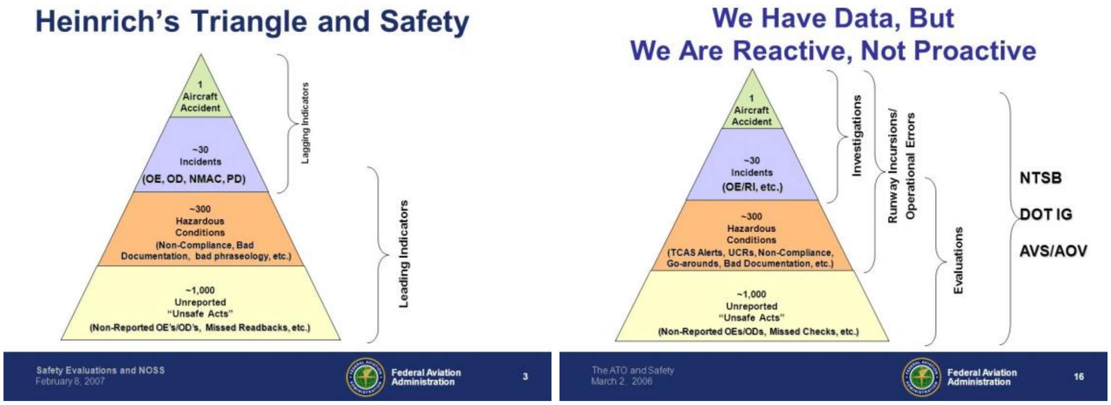
Other contemporary researchers have supported Heinrich's common cause hypotheses and Heinrich's triangle despite theoretical challenges. First, research conducted by Alamgir et al. (2009), a team of occupational health professionals who analyzed the causal factors for three levels of occupational injuries across three regions in Canada, found similar causal factors across the three severity levels. Second, a similar congruency of causal factors was observed in the rail industry by Wright (2002), who analyzed 250 railway incidents and identified only three out of twenty-one causal factors (knowledge-based errors, training, and procedures) significantly different across the three severity levels. Third, in their survey of 1,069 health professionals and research on various significant mishaps in the medical profession, such as sharps injuries and bodily fluid exposure, Kim et al. (2010) identified similar frequencies in risk factors for those events as well as their less severe near-miss cases. Finally, when comparing the safety reporting systems in aviation and medicine based on research by Reason (2016) and Heinrich (1931), Merry et al. (2017) claimed that "the chain of events that leads to a near miss is often the same as the chain of events that leads to a serious accident, and the underlying cause may often also be the same" (p. 291).

Concerns Reflected upon Aviation VSR Systems

The literature review on SMS indicated that VSR is integral to SMS's risk management and safety assurance elements. Hazard reports originating from VSRs are expected to provide data for proactive safety management. In addition, Heinrich's principles are widely manifested in present-day safety management, particularly in modern aviation, which relies on VSR as one source of safety performance data (ICAO, 2016). For instance, the FAA has used Heinrich's triangle to explain the relationships between various safety reporting systems (see Figure 11). Likewise, a cursory internet search showed Heinrich's triangle is used in various safety training programs, particularly aviation SMS training.

Figure 11

FAA Versions of Heinrich's Triangle



Note. Reprinted from Presentations to Second ICAO Global Symposium on TEM / NOSS In Air Traffic Control and Aerospace Control and Guidance Systems Committee by FAA. Copyright 2006 and 2007 by FAA.

Can VSRs Effectively Identify and Mitigate the Hazards Behind High-Severity Events?

Despite the benefits of VSR, concerns have been raised regarding the effectiveness of VSR programs in mitigating high-severity events. In reviewing the history of safety management strategies in the United Kingdom through industrial safety performance, Cooper (2019) found that safety management strategies maximized efforts to identify and mitigate through VSR and other means, and that the number of events resulting in temporary disability had been reduced by 66% over the past 32 years. Nevertheless, the rate of decline in serious injuries and fatalities (SIFs) for the region has been negligible, stagnating over the same period. The findings made by Cooper (2019) are analogous to those for the aviation industry (see Figures 5 and 6); whereas overall accident rates have been reduced significantly, and fatal accident rates have stagnated at the same order of magnitude for decades, with LOC-I events continuing to be a key contributor to such figures.

Resources Spent Not Commensurate with Risks Mitigated

As the implementation of SMS has become a mandatory requirement for civil aviation regulators worldwide (European Commission, 2015; FAA, 2015; ICAO, 2016), significant resources have been, and will continue to be, invested in their establishment and implementation, including VSR programs. While the literature review for the present research did not result in any study to date that focuses on the financial costs of implementing VSR programs, the FAA predicted that the implementation of SMSs in the U.S. aviation industry would cost around $135.1 million from 2015 to 2025 (Okwera, 2016). In the case of the United States, Okwera (2016) identified that the estimated total annual and maintenance costs for SMSs would depend on the size and complexity of the business; however, since such costs are not directly proportional to organizational size, Okwera (2016) argued that most small- and medium-sized companies lacked the means to implement extant safety programs that larger companies have already put in place. Okwera (2016) placed the annual cost for an air operator to implement an SMS in the United States at $483,500–$1,267,000.

In human resources terms, taking the example of a regional low-cost carrier based in Hong Kong with 1,000 staff, 24 aircraft, and an average of 70 regional flights daily, implementing an SMS would require a team of five full-time staff involved in administering and facilitating risk assessments, as well as investigating VSR reports (Lee, n.d.). Having assessed the low-risk events using the operator’s risk matrix by the team, the safety focal points in each operational department are responsible for executing, tracking, and lobbying line departments to implement identified mitigation actions. To ensure financial viability in commercial aviation, airline management must frequently scrutinize business performance, return on investment (ROI), and cost controls (Moss & Ryan, 2016). Given that the literature review has not identified research evaluating the effectiveness of VSR programs, notably those publicly available such as ASRS, it is argued that the resources spent on VSR may be better utilized on directly addressing hazards leading to significant risks.

Reporting Bias Leading to Actual Hazards Being Unidentified

As VSR systems are being implemented in aviation organizations worldwide, organizational culture may affect the information being reported and, thereby, the overall effectiveness of a VSR program. Research has found that organizations exhibit various safety culture maturity levels (Hudson et al., 2006) and national cultures, in which the willingness to report and the quality of VSR reports vary significantly (Flynn et al., 2018; Noort et al., 2016). Jausan et al. (2017) identified individual, organizational, and environmental factors that can affect the performance of a safety reporting system.

A parallel can be drawn with the medical industry. Using a survey of approximately 800 healthcare professionals and follow-up questionnaires to 315, Noble and Pronovost (2010) highlighted the epidemiological problems in voluntary safety reporting. The three areas are underreporting, leading to a systematic bias, lack of generalizability to whole patient populations, and participation bias. The barriers in reporting are structural, process-based, outcome-oriented, and fear and attitude related. Similar research was carried out by Spigelman and Swan (2005) in Australia, focusing on the Australian Incident Monitoring System (AIMS). While underreporting and bias were still identified, most respondents (83%) reported that AIMS investigations resulted in significant changes to equipment usage, medication prescribing or administration, clinical protocols, training programs, and fall risk assessment tools.

Another particularly notable factor leading to the challenge of the relevance of VSR is the COVID-19 pandemic. A search in the ASRS database revealed that while the exact number of reports coded as LOC-I had fallen in 2020, the number of reports did not fall at the same rate as the air traffic in the United States for 2020 decreased to 41.7% of the volume in 2019 for commercial aviation (BTS, 2020). This result is to be contrasted with research by Anderson (2013), where accident rates for general aviation remained consistent regardless of flight hours over ASRS data spanning eight years. There has been no research to date on the impact of COVID-19, such as lower flight hours to commercial aviation LOC-I VSR reporting. As highlighted by Noble and Pronovost (2010), the level of underreporting or reporting bias in a relatively less intense, lower flight hours environment is unknown.

Study Involving Interrater Reliability Analysis

Based on Human Factors Analysis and Classification (HFACS), Yesilbas (2014) coded 272 Uncrewed Air Vehicle (UAV) accident records from the U.S. Navy. They validated them against various accident models using a Structural Equation Modelling (SEM) technique. Four raters were deployed for the study. Yesilbas (2014) raised the agreement rate between raters from the defaulted 50%for untrained raters, as stated in O’Connor et al. (2010). With the training and retraining regime on the coding with the raters, Yesilbas (2014) obtained the confidence level of $\alpha < 0.05$ sampling resolution and misclassification rate required for the study.

Gaps in the Literature

The literature review conducted above indicates gaps in the following areas:

a. Lack of archival or empirical assessment in the relevance of publicly available aviation *open-loop* VSR such as ASRS as one of the sources to support defining safety mitigations to reduce the likelihood of severe or catastrophic LOC-I events;

b. Lack of archival or empirical assessment of the strengths and weaknesses in the quantity and context of the safety reports from publicly available VSR databases such as ASRS for the provision of proactive risk mitigation information in an SMS;

c. Lack of sensitivity analysis on the LOC-I VSR reporting rate. Despite differences in flight hours and the number of accidents, Anderson (2013) found a constant reporting rate of accidents in general aviation with eight years of accident data. This is not the case in the commercial aviation LOC-I VSR data over the 2020 COVID-19 Pandemic period based on a preliminary ASRS search by the researcher. Therefore, the effect of this possible covariate is not known;

d. The validity of Heinrich's theories, including the common cause hypothesis, in modern aviation; and

e. Whether publicly available VSR such as ASRS should be viewed as a priority or dependable tool for safety assurance in the resource-limiting environment of modern aviation.

Theoretical Framework

The theoretical framework of this research originates from the modeling of LOC-I events conducted by Belcastro et al. (2017), which identified the hazards leading to the occurrence of LOC-I. The present research is thus centered on whether VSR systems represented by ASRS can provide relevant and adequate information for preventing severe LOC-I events. Consequently, the hazards identified in VSR reports are compared with those listed in high-severity LOC-I event reports. Based on the ICAO classification for safety events, such high-severity events are processed as accident investigations.

The study is centered on whether the hazards identified in VSR reports are identical or equivalent to those listed in accident reports, providing an opportunity to proactively execute preventive measures before accidents or more consequential events manifest. The study was conducted using a quantitative approach, supplemented by a qualitative approach with the following rationale.

A related theory is Heinrich's common cause hypothesis. The literature review demonstrated the significance of this hypothesis and its related theories to modern aviation SMSs. Heinrich's theories suggest that low-severity events share the same causes as their high-severity counterparts. Per ICAO requirements, more severe incidents and accidents are to be investigated by the state's investigation authority, such as the NTSB for the United States, giving light to the capture of less severe events by VSR systems such as ASRS. Although this study does not analyze the causal relationships between the hazards identified from each report dataset, the absence of similarity in hazard distribution will refute Heinrich's theories in the context of this research.

The application of MANOVA is widely used in safety science research. The theories behind the MANOVA methodology were covered by Hair et al. (2019), who addressed the main effect of the independent variable (IV) on the dependent variables (IVs), as well as identified the magnitude and significance of the univariate differences.

Research Framework

The study references research conducted by Anderson (2013), who explored the relationship between certificate types and types of general aviation accidents using a quantitative supplemented by qualitative approach. To assess whether a publicly available VSR such as ASRS is a relevant tool, a technical analysis on the occurrence rates of hazards alone may not provide a complete picture, as relevance is a subject as well as a dichotomy of quantitative supplemented by qualitative measures (Teddie & Tashkakori, 2009). Traditionally, a safety report consists of an assessment of the findings or hazards associated with the case, a narrative description of the sequence of events, the actions taken, and recommendations to prevent another occurrence (ICAO, 2001). While identifying the coded hazards provides the statistical data required, the richness of the narrative descriptions also needs to be explored due to the contextual and emerging information that may be concealed, thereby justifying the deployment of qualitative techniques to supplement the quantitative research. Teddie and Tashkakori (2009) described this as a pragmatist paradigm focusing on *what works*, which is the exact purpose of this study. Anderson (2013) also adopted this approach when researching the impact of certifications on accident rates for various types of aviation accidents, canvassing quantitative and qualitative data.

To answer the research questions, an exercise was conducted to identify and compare the Belcastro LOC-I Hazards reported in the VSR (ASRS) and accident (NTSB) reports. The already available coded data was beneficial to data collection. Anderson (2013) successfully used coded data to reach her study's reliability and validity requirements. As the taxonomies differed between ASRS and NTSB reporting systems, analyzing the coded hazards mapped to the Belcastro LOC-I Hazards provided the universal instrument for comparison.

Hypotheses and Support

The hypotheses generated relate to the quantitative part of the study and provide the statistical basis for answering the research questions, as explained in Table 3.

Table 3

Research Questions and Alternative Hypotheses of the Current Study

Research Question	Alternative Hypotheses	Theoretical Background
RQ1: Do Belcastro LOC-I Hazard rates differ across types of safety reports for commercial and general aviation?	H_A1 The group mean vectors in Belcastro LOC-I Hazard rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.	Belcastro (2017) identified eight factors that led to the onset of a LOC-I. Heinrich (1931) described the common causation hypothesis and Heinrich's triangle.
RQ2: Which individual Belcastro LOC-I Hazards display(s) significant difference(s) in hazard rates across types of safety reports for commercial and general aviation?	H_A2 The means of adverse onboard conditions - vehicle impairment rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020. H_A3 The means of adverse onboard conditions - system and components failure / malfunction rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020. H_A4 The means of adverse onboard conditions - crew action / inaction rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020. H_A5 The means of external hazards and disturbances - inclement weather atmospheric disturbances rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020. H_A6 The means of external hazards and disturbances - poor visibility rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020. H_A7 The means of external hazards and disturbances - obstacle rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020. H_A8 The means of abnormal vehicle dynamics and upsets - abnormal vehicle dynamics rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020. H_A9 The means of abnormal vehicle dynamics and upsets - vehicle upset conditions rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.	Hair et al. (2019) described the methodology for a one-way MANOVA and associated post hoc techniques such as discriminant analysis. Anderson (2013) used accident rates to research the impact of certifications for various types of aviation accidents, canvassing quantitative and qualitative data. H_A1 tests if Belcastro LOC-I Hazard rates differ with commercial and general aviation in ASRS and NTSB reports at multivariate levels. H_A2 to H_A9 test if Belcastro LOC-I Hazard rates differ in ASRS and NTSB LOC-I reports for commercial and general aviation at univariate levels.

The quantitative data analyzed provided the core materials to answer the research questions. The researcher attempted to take a quantitative approach, supplemented by a qualitative view. Cursory qualitative data analysis would identify patterns in textual clusters and contextual information, providing additional insights.

Summary

The literature review presented in Chapter II highlighted the widespread application of VSR systems in modern aviation safety management. While VSR is officially supported by regulators worldwide as part of SMS solutions, its application may be susceptible to underreporting, biases, and the reporting rate sensitivity to exposure levels, such as flight hours, is unknown.

The lack of empirical research on the relevance of open-loop VSRs in aviation, particularly those publicly available VSRs such as ASRS, has been identified. The need to scrutinize the relevance of VSRs as a credible source in reducing the likelihood of severe accidents in modern aviation was highlighted. The common assumption further compounded this scrutiny that reporting low-severity, near-miss events, typically through VSR, can reveal the hazards causing high-severity events, providing organizations with the information and early intervention opportunities to prevent accidents. This assumption aligns with Heinrich’s common cause hypothesis, the basis of Heinrich’s triangle, which again has not been validated in any context, nor has the source data been disclosed. This assumption might bias the consideration of modern aviation critical risk events such as LOC-I, highlighting the need for further validation. The use of MANOVA as a technique to analyze the main effect of an IV on DVs in a multivariate and univariate setting for safety topics was documented in various research and showed acceptability in peer-reviewed works. The MANOVA methodology was based on guidance by Hair et al. (2019). The information discussed in this chapter forms the theoretical basis of the research.

Chapter III: Methodology

The academic foundation for the research methodology and design has been examined in the literature review. The content of Chapter III details and justifies the steps taken in this research, answering research questions one and two by testing their associated hypotheses. The information documented is sufficiently detailed to enable other scholars to replicate this research, increasing internal validity.

Research Method Selection

This research adopted a quantitative-dominated mixed research method based on archival data from existing coded safety reports. Apart from the safety reporting systems taxonomy mapping, the research was conducted by a single researcher. The research questions required representative samples nationally, and using already coded data from established databases such as ASRS and NTSB were deemed appropriate per Vogt et al. (2012). The research questions were formally answered, and associated hypotheses were tested by quantitative analyses results, with additional insights provided from cursory qualitative analysis. The research was conducted with the rigor in assumptions testing and data analyses necessary for multivariate quantitative research (Hair et al., 2019). With the availability of textual data from each safety report, cursory qualitative data analysis was conducted on the original dataset with the addition of an un-coded source, AIDS, to provide insights into the reasons behind the results obtained.

The research consisted of four phases (Teddie & Tashakkori, 2009). The first phase involved collecting the classified and augmented searched LOC-I reports from the ASRS and NTSB systems for general and commercial aviation. The second phase involved mapping the code taxonomies from the ASRS and NTSB databases to Belcastro et al.'s (2017) LOC-I precursors / hazards model, referred to as Belcastro's LOC-I Hazards Model hereafter. The mapping was performed by a team of four aviation safety practitioners to provide a common instrument for measurement across the datasets. The third phase involved operationalizing the collected reports into hazard rates and performing quantitative analyses. Descriptive and inferential statistical analyses were performed to test the nine hypotheses using MANOVA and discriminant analyses. The fourth phase consisted of cursory qualitative data using narrative texts of accident and incident investigation reports from ASRS, AIDS, and the NTSB databases. Techniques such as tree maps, hierarchy charts, and word clouds were deployed. Qualitative data analysis provided insights into the rationale behind the quantitative results but was not at the same level of rigor as the quantitative analysis.

Population/Sample

The data sources for this study originated from the United States; this study is primarily focused on fixed-wing commercial and general aviation (FAR Parts 91, 121, and 135) operational certifications under the U.S. regulatory environment.

Population and Sampling Frame

The population for the study consists of all fixed-wing flights registered in the United States operating under commercial aviation (FAR Parts 121 and 135) and general aviation (FAR Part 91) operational certifications. The sampling frame in terms of time is the period between 2004 and 2020. For the quantitative part of the research, the sampling frame includes flights that were involved in the following:

a. A LOC-I event, voluntarily reported through the ASRS system or under mandatory investigation by the NTSB, which is classified in the relevant databases as LOC-I. The sample consists of LOC-I events reported between 2004 to 2020. This type of report is known as the *classified search* report in this study.

b. Events not classified as LOC-I in the relevant database but identified by augmented search based on LOC-I precursors' keywords prescribed by Belcastro et al. (2017). This type of report is known as the *augmented search* report in this study.

For the qualitative part of the study, the synopsis and narratives on LOC-I reports from ASRS and NTSB databases were supplemented by AIDS LOC-I reports to enhance the qualitative data for medium-severity incidents.

Sample Size

Out of the population of LOC-I events, a search in the ASRS and NTSB databases for reports classified *as loss of control* between 2004 to 2020 provided the sample frame of LOC-I safety reports outlined in Table 4, a total of 7,681 cases. To ensure independence among sets of data, the unique case numbers in each group were checked, and any duplicates were removed from the supplementary groups. The data analysis section justified this sample size based on analysis using GPower®.

Table 4

Sample Frame for Events Identified as LOC-I in the ASRS, AIDS and NTSB Databases Between 2004 and 2020.

Operation Type	Number of Events		
	ASRS	**NTSB**	**AIDS[a]**
General aviation (Part 91) classified search	1041	3045	62
General aviation (Part 91) augmented keyword search	770	2282	51
Commercial aviation (Parts 121 and 135) classified search	804	2791	11
Commercial aviation (Parts 121 and 135) augmented keyword search	1502	197	40

[a]Data used in the qualitative analysis only.

It was anticipated that some safety events had not been directly classified as LOC-I but contained LOC-I precursors. The presence of such precursors was highlighted in Belcastro et al.'s (2017) research. The detection of such precursors highlighted the onset of LOC-I, which was synonymous with the earlier part of the accident causation chain (Reason, 2016). Such events might result in a less severe, or uneventful, consequence, and hence were not classified as a LOC-I initially. The research has therefore been extended by covering LOC-I cases selected by the augmented keyword search based on precursors identified by Belcastro et al. (2017). This research compared if there was a difference in LOC-I hazard rates between the classified LOC-I and the augmented search reports. This search method provided an additional group of events that contained LOC-I precursors but with less severe consequences for analysis. Table 4 shows that an additional 4,751 cases were identified from the augmented search. To provide more comprehensive answers to the research questions, the quantitative results were supplemented by insights from cursory qualitative analysis using NVivo®.

Sampling Strategy

All available fixed-wing safety reports from the ASRS and NTSB databases that had either been classified as LOC-I or fulfilled the augmented keyword search criteria based on LOC-I precursor keywords search per Belcastro et al. (2017) within the 2004 to 2020 sample frame were used. The numbers of relevant reports are indicated in Table 4. AIDS LOC-I reports were added to the supplementary qualitative analysis.

Data Collection Process

For taxonomies alignment, the original coding taxonomies have been obtained from the ASRS and NTSB accident investigation webpages. For the quantitative analyses, the datasets required were downloaded from the ASRS and NTSB databases using the search functions provided. Where the augmented keyword search was used, the reports were reviewed by the researcher to ensure the validity of the selected reports. For example, if the word *upset* was identified in the report, the researcher verified if this was related to the in-flight attitude upset rather than a human psychological state of being upset to avoid irrelevant data being analyzed. Also, special effort was made to ensure that no rotary-wing LOC-I reports were included in the research data, and that no duplication of cases between the classified and augmented groups.

A Microsoft Excel® spreadsheet was created to capture the coded and mapped data from the reports. The spreadsheet was then exported to the statistical analysis software, IBM Statistical Product and Service Solutions (SPSS®), for the data to be analyzed. The qualitative data for the study was extracted from the synopsis and narrative sections of the NTSB accident investigation reports and ASRS reports for the LOC-I events. Qualitative data analysis was performed to analyze the qualitative portion using the NVivo® tool.

Design and Procedures

This research was centered on the application of MANOVA analysis. Multivariate statistical techniques, such as MANOVA, have been successfully deployed in modern research related to flight safety. For example, Wang et al. (2020) used MANOVA as a statistical method to assess pilot workload from four dimensions: cognitive activity, control activity, stress, and flight performance. Balaj et al. (2018) used MANOVA to analyze pilots' gaze behavior (gaze time at areas of interest) and pilot groups (IV) on 20 pilots and no-pilots using a flight simulator.

Five steps were designed for this research. Firstly, noting the differences in the coding systems between ASRS and NTSB, a team of four experts aligned the taxonomies from each reporting database by mapping them onto the eight Belcastro LOC-I Hazards (see Table 5). These eight hazards were highlighted in Belcastro et al.'s (2017) research as hazards leading to LOC-I. Secondly, LOC-I reports were obtained from ASRS and NTSB databases through LOC-I classification or augmented search based on LOC-I precursors' keywords (Belcastro et al., 2017).

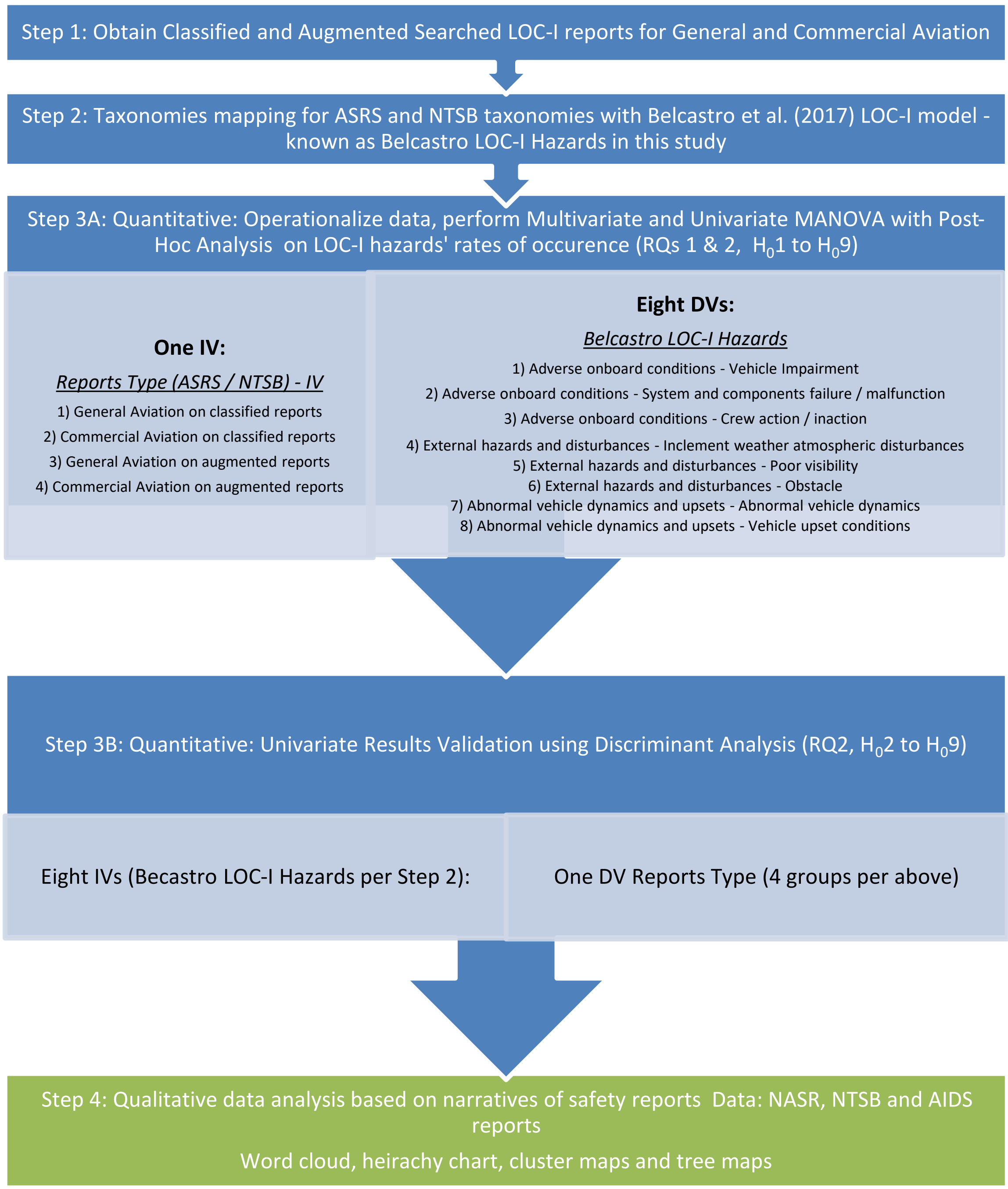
Thirdly, MANOVA was performed to identify the differences between the eight Belcastro LOC-I Hazard rates among four groups of ASRS and NTSB LOC-I reports from the multivariate and univariate perspectives for commercial and general aviation. The MANOVA analysis was based on the normalized annual rates of eight Belcastro LOC-I Hazards. The flight hours data used for normalization to obtain annual hazard rates were obtained from the relevant agencies in the U.S. government, such as the Bureau of Transportation Statistics (BTS) and the FAA. A literature review suggests a collection of factors influencing the VSR reporting rate. An initial search of the ASRS database suggested a variation between the VSR LOC-I reporting rate in commercial aviation and the reduction in flight hours over the COVID-19 pandemic. This was contrary to the findings by Anderson (2013) that the overall accident report rate remained relatively consistent with general aviation accident data over an eight-year period. Hence, measuring hazard rates normalized by flying hour addressed this potential covariate to any VSR reporting rate analysis.

Once the multivariate and univariate results from the MANOVA and related post hoc analyses were obtained, the univariate results were further validated using discriminant analysis. This analysis assessed the individual outcome variables' differences across the treatment variable. As the objective was to profile the outcome variables in terms of their differences between groups of treatment variables, Hair et al. (2018) stated that discriminant analysis was particularly insightful when the treatment variable had three or more levels, as in this research.

Although the primary focus of this research was quantitative based on MANOVA, with the vast textual data available, on an opportunity basis, the research attempted to use the textual data and performed cursory qualitative analysis to identify insights that could explain the quantitative results. See Figure 12 for the summary of the steps involved.

Figure 12

Figure Summarizing the Four Steps in Data Analysis



Apparatus and Materials

ASRS and NTSB aircraft accident databases were used for the entire study. The ASRS database provided the data source for the voluntary and comparatively low-severity reports. In contrast, the NTSB accident database provided the data for the higher severity and mandatorily reported incidents and accidents. The AIDS reports, which were positioned with the medium severity level between ASRS and NTSB reports, were used to provide the textual narrative data for the qualitative portion of the research only due to the lack of coded data in that database. IBM SPSS® and NVivo® were used for the quantitative and qualitative parts of the analysis.

Sources of the Data

This study explored the efficacy of the VSR system in the LOC-I context. As stated in the literature review, the ASRS is a fountain of resources for LOC-I voluntary safety reports administered by a professional organization, NASA. The level of rigor of the investigation also increases with the AIDS and NTSB accident investigation reports, which follow the ICAO protocol in the investigation. All the databases adopted are publicly available online in Microsoft Access® and Excel® formats. The flight hours data for normalization of the MANOVA datasets were obtained from the FAA General Aviation Survey and the Bureau of Transportation Statistics (BTS), respectively. Both are publicly available governmental sources based on operators’ data. The assumption on the accuracy of the self-reported flight hour FAA data for general aviation and Part 135 operation has already been detailed in the assumptions section. Both data sources are suitable for archival research, as Anderson (2013) demonstrated.

LOC-I safety reports with four severity levels were extracted from the NTSB and ASRS databases for this study, forming the four independent groups represented by one IV for this research. The unique case identification numbers of each group were compared among the other groups to ensure no duplication of cases, ensuring independence, as follows:

a. NTSB classified search LOC-I reports

b. NTSB augmented search LOC-I reports

c. ASRS classified LOC-I reports

d. ASRS augmented search LOC-I reports

It is acknowledged that ASRS and NTSB were not the only VSR and Mandatory Safety Reports (MSR) safety reporting systems and other data sources that could be used. The rationale for deploying ASRS and NTSB datasets for covering the required demographics for the quantitative MANOVA and discriminant analyses in answering RQ1 and RQ2 is as follows:

a. availability of the data in the public domain that covered the demographic of the U.S. aviation community

b. data was already coded by the database administrators

c. ability to contrast the two reporting systems with ASRS being an open-loop voluntary safety reporting system with comparatively little investigation, verification, and feedback to the originators, and NTSB being an air accident safety reporting system involving high rigor investigation by investigators

d. research conducted by Belcastro et al. (2018) suggested that augmented keyword search of cases such as *loss of control*, *upset*, *unusual attitude, stall*, *crash out of control*, and *uncontrolled descent* yielded a selection of LOC-I cases not classified previously, which enriched the relevant research

The study was extended to include the AIDS voluntary safety reporting system, which had a more enhanced closed-loop structure than ASRS. Introducing AIDS supplemented the overall quantitative analysis result. The extended research was only performed in a qualitative manner, as publicly available AIDS safety reports were not coded by the database administrator.

Ethical Considerations

This study was archival research based on available data published in the public domain. The involvement of human participants in generating research data was not part of the research plan. Therefore, no application to the Internal Review Board (IRB) at Embry-Riddle Aeronautical University was necessary. The pool of SME raters who supported the researcher's mapping of the taxonomy codes team was given anonymity and privacy statements, as part of a workshop provided. Their expressed consent to participate on a voluntary basis, as well as to include their career resumés in Appendix F, was obtained.

Measurement Instrument

This research was an archival study with the data already coded from the data sources for the quantitative part from relevant Microsoft Access® file downloads. One instrument, a mapping table, was used to ensure the NTSB and the ASRS codes were mapped toward the Belcastro et al. (2017) model for LOC-I events. This mapping table aligns the coding taxonomies deployed by ASRS and NTSB, so the content between the two reporting systems can be analyzed concurrently. The process of using Microsoft Forms® and subsequent workshops to collate the raters' assessment to achieve a congruent set of mapped codes are discussed later in this chapter.

In addition, a qualitative approach to supplement the quantitative analysis using NVivo® was also deployed, whereby the instruments of word frequency, text search, hierarchy chart, and cluster analysis were conducted.

Variables and Scales

The operationalized variables, definitions, and scales are summarized in Table 5. The independent variable containing four groups is based on the severity level of safety reports, categorical in nature, with the dependent variables as the normalized rate of the eight mapped Belcastro LOC-I Hazards contained in the coded ASRS and NTSB reports, continuous in nature. The normalized rates for such analysis were successfully used by Anderson (2013), who conducted similar research.

Table 5

Independent, Dependent Variables (IV & DV) and Covariate for the Research

IV / DV / Covariate	Definition	Scale	Addresses RQ(s)
IV	Type of report: mandatorily reported (NSTB) investigation report or VSR (ASRS) 1 – NTSB Classified, 2- NTSB Augmented, 3 – ASRS Classified, 4 – ARS Augmented	Categorical	1 & 2
Covariant	Hours flown per certification type per year (used as normalization denominator)	Metric	1 & 2
DV1	Hazard rate per report obtained – Vehicle Impairment: • Improper maintenance action/inaction/procedure • Inappropriate vehicle configuration • Contaminated airfoil	Metric	1 & 2

IV / DV / Covariate	Definition	Scale	Addresses RQ(s)
	• Smoke/fire/explosion • Improper loading: weight/balance/CG • Airframe structural damage • Engine damage/foreign object damage (FOD)		
DV2	Hazard rate per report obtained – System & Components Failure/Malfunction: • System design/validation error/system inadequacy • System software (SW) design/verification error/software inadequacy • Control component failure/inadequacy • Engine failure/malfunction (F/M) • Sensor system F/M • Flight-deck instrumentation malfunction/inadequacy • System F/M (non-control component)	Metric	1 & 2
DV3	Hazard rate per report obtained – Crew Action/Inaction: • Loss of attitude state awareness/spatial disorientation • Loss of energy state awareness • Lack of aircraft/system state awareness • Aggressive maneuver • Abnormal/inadvertent control input • Improper/ineffective recovery • Inadequate crew resource monitoring/management • Improper/incorrect/inappropriate procedure/action • Fatigue/impairment/incapacitation	Metric	1 & 2
DV4	Hazard rate per report per year– Inclement weather and atmospheric disturbances: Thunderstorms/rain: • Wind shear • Wind/turbulence • Wake vortex • Snow/icing	Metric	1 & 2
DV5	Hazard rate per report per year – Poor visibility: • Fog, haze • Night	Metric	1 & 2
DV6	Hazard rate per report per year – Obstacle: • Fixed obstacle • Moving obstacle	Metric	1 & 2
DV7	Hazard rate per report per year – Abnormal vehicle dynamics: • Uncommanded motions • Oscillatory response/pilot-induced oscillation • Abnormal control for trim/flight and/or control asymmetry • Abnormal/counterintuitive control response	Metric	1 & 2
DV8	Hazard rate per report per year – Vehicle upset conditions: • Abnormal attitude • Abnormal airspeed/energy • Abnormal angular rates • Undesired abrupt response • Abnormal flight trajectory • Minimum control speed (Vmc)/departure • Stall/departure	Metric	1 & 2
Total	MANOVA 1 IV, 4 groups	8DVs	

Note. Technique deployed was MANOVA and Discriminant Analysis. Sources were NTSB Accident Investigation Reports and ASRS reports database, and BTS data

Data Analysis Approach

Before the commencement of the data analysis, the adequacy of the sample size was explored. As stated, the population for this research was 5,836 classified higher severity LOC-I events, with 1,845 classified low-severity events from VSRs in the ASRS database. To test hypotheses H_A1 to H_A9, a one-way MANOVA with four groups (ASRS classified, ASRS augmented, NTSB classified, NTSB augmented) was performed on the commercial and general aviation datasets. This MANOVA was based on normalized hazard rates of the eight Belcastro LOC-I Hazards per year over the 17-year timespan.

The MANOVA study involved four independent groups with eight dependent variables. GPower® was used to ascertain the total sample size required. Based on a large effect size of f2 of 0.2, power of 0.8, and an alpha value of 0.05, GPower® calculated that for MANOVA global effects analysis, the total sample size required was 44, which achieved actual power of 0.80. With this calculation, as there were four independent groups, it would require a minimum of 11 samples per group. Hence, the hazard rates for a minimum of 11 years per hazard were required. The study covered 17 years of hazard rate data from 2004 to 2020. Therefore, the sample size surpassed the requirement by one year. The GPower® calculations are documented in Appendix D.

As mentioned above, having obtained the datasets, there were three steps to the data analysis: step two generated a LOC-I taxonomy mapping table, step three involved MANOVA (multivariate, univariate, and post hoc), followed by validation of the univariate MANOVA result by discriminant analysis, and the final step involved qualitative data analysis to supplement the quantitative analysis. The last part of the research assimilated the quantitative results with qualitative data analysis from the reports' narratives in ASRS, NTSB, and AIDS databases for identified LOC-I reports. The qualitative step supplemented the quantitative results and compensated for any statistical inadequacies, such as assumptions met partially. The steps are summarized in Figure 12.

Step 2: Taxonomies Mapping

The research was grounded on the generic LOC-I model developed by Belcastro et al. (2017). Eight Belcastro LOC-I Hazards were applied as DVs for the quantitative MANOVA analysis (see Table 5). As the ASRS and NTSB taxonomies were not identical (see Figure 15), a common set of taxonomies based on the Belcastro LOC-I Hazards was necessary to facilitate the MANOVA analysis. A mapping table was developed by four aviation subject matter experts (SMEs) to map the ASRS and NTSB taxonomy codes with the Belcastro LOC-I Hazards. A validation process was in place to achieve the required inter-rater reliability. The mapped codes were used as the theoretical basis for the DVs in this study to quantitively assess the hazards coded in the ASRS and NTSB reports, as identified in Table 5.

Inter-Rater Reliability Assurance

Four SME raters, having ten or more years of experience in aviation safety management or flight operations as commercial pilots, were presented with the taxonomy codes from the ASRS and NTSB investigation reports (see Figure 15). Online workshops were held for the raters to discuss and arrive at a mapping table that mapped the eight Belcastro LOC-I Hazards. Due to the vast number of codes (over one thousand) and the time available for this research, for the ASRS and NTSB taxonomies, the codes that SMEs mapped were the codes that covered 95% or more of the classified and augmented searched LOC-I cases. This method aligned with Pareto's principle (Grami, 2020), the adoption of which will be further explained. In order to ensure the mapping is reliable, the finalized mapping should achieve interrater reliability with ICC kappa of 0.7 or higher (Gisev et al., 2013) among the raters. The justification of the kappa and the value to be used can be found in Figure 13. The summarized steps for deriving the mapping table can be found in Figure 14.

Figure 13

Examples of Interrater Indices (source: Gisev et al., 2013)

Examples of interrater indices suitable for use for various types of data[a]

	Level of measurement					
	Nominal/categorical		Ordinal		Interval and ratio	
	2 raters	>2 raters	2 raters	>2 raters	2 raters	>2 raters
Interrater indices	Cohen's kappa	Fleiss' kappa	Weighted kappa	Kendall coefficient of concordance	Bland-Altman plots	ICC
	ICC	ICC	ICC	ICC	ICC	
	Weighted kappa					

[a] Table is not exhaustive and represents a summary of some of the indices and the contexts in which they can be used only.

Note: Reprinted from *Interrater Agreement and Interrater Reliability: Key Concepts, Approaches, and Applications* by N. Gisev et al., Research in Social and Administrative Pharmacy. Copyright 2013 by ScienceDirect, p. 333.

Figure 14

ASRS and NTSB Reports Mapping Table Generation Steps

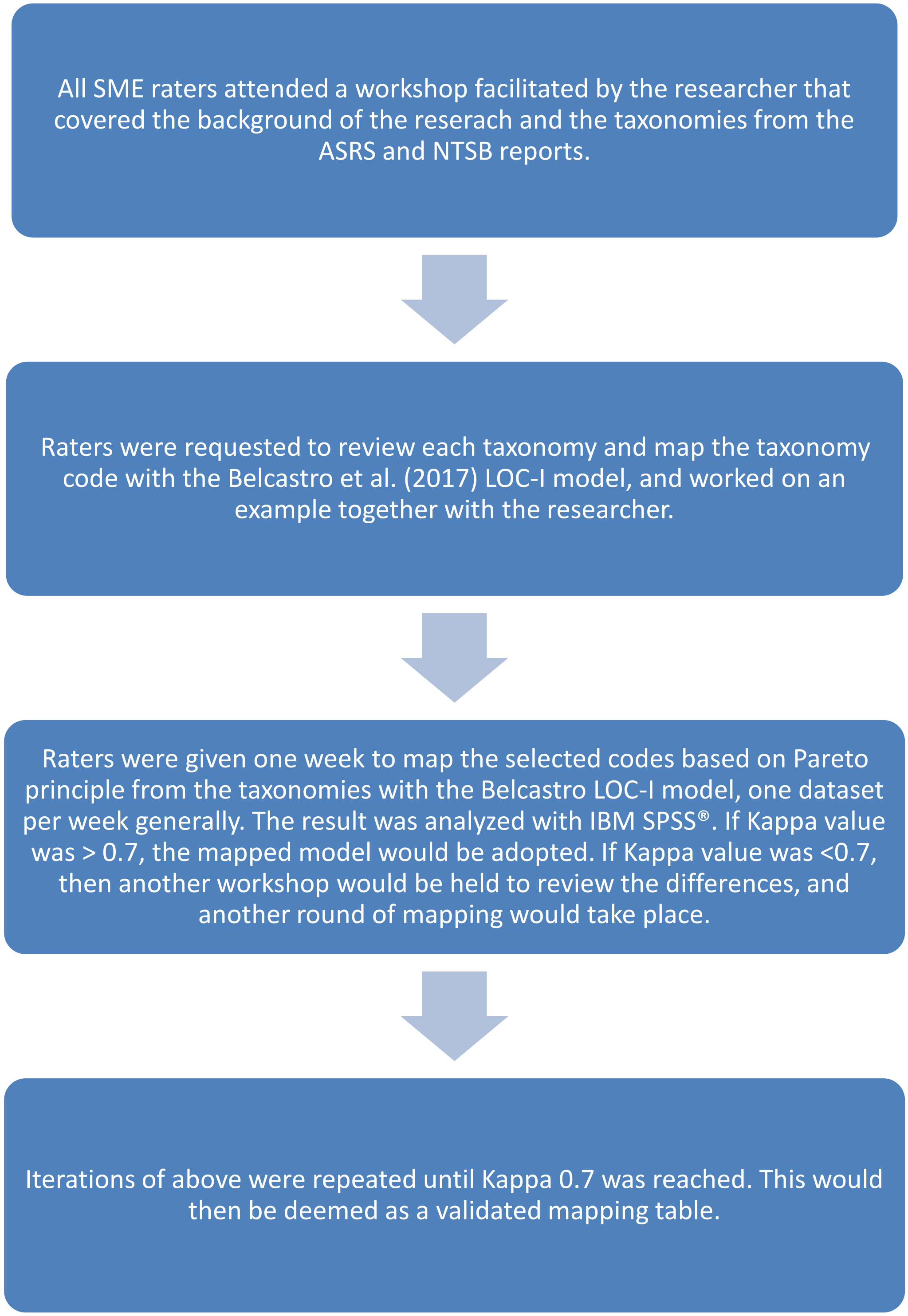


Figure 15

Examples of the ASRS Coding Table (Left) and NTSB Air Accident Coding Table (Right)

Aviation Safety Reporting System
Database Fields

Assessments	
Contributing Factors / Situations	Q1
5273 - Aircraft	
5274 - Airport	
5275 - Airspace Structure	
5276 - ATC Equip / Nav Facility / Buildings	
5277- Chart or Publication	
5278 – Company Policy	
5279 – Equipment / Tooling	
5280 - Environment – Non Weather Related	
5281 - Human Factors	
5282 - Incorrect / Not Installed / Unavailable Part	
5283 - Logbook Entry	
5284 - Manuals	
5285 - MEL	
5286 - Procedure	
5287 - Staffing	
5288 - Weather	
Primary Problem	Q1
5289 – Aircraft	
5290 – Airport	
5291 - Airspace Structure	
5292 - ATC Equipment / Nav Facility / Buildings	
5293 - Chart or Publication	
5294 – Company Policy	
5295 – Equipment / Tooling	
5296 - Environment – Non-Weather Related	
5297 - Human Factors	
5298 - Incorrect / Not Installed / Unavailable Part	
5299 - Logbook Entry	
5300 - Manuals	
5301 - MEL	
5302 - Procedure	
5303 - Staffing	
5304 - Weather	
5305 - Ambiguous	

CONTENTS

I. INSTRUCTIONS 1
II. DEFINITIONS 7
III. SEQUENCE-OF-EVENTS MATRIX/WORKSHEET 13
IV. CODES FOR OCCURRENCES (100-430) 14
V. CODES FOR PHASES OF OPERATION (500-610) 15
VI. CODES FOR CAUSES, FACTORS, AND OTHER EVENTS:

CODES FOR SECTION IA

AIRCRAFT STRUCTURE SUBJECTS (10000-11306 & 13000-13014)

DOOR 16
FLIGHT CONTROL SURFACES/ATTACHMENTS 16
FLIGHT CONTROL SYSTEM 17
FUSELAGE 17
LANDING GEAR 18
MISCELLANEOUS, AIRFRAME 18
MISCELLANEOUS, ROTORCRAFT 19
NACELLE/PYLON 19
ROTOR DRIVE SYSTEM 19
ROTOR SYSTEM 20
ROTORCRAFT FLIGHT CONTROL SYSTEM 20
ROTORCRAFT FLIGHT CONTROL 21
STABILIZER 21
WINDOW 21
WING 21

AIRCRAFT SYSTEM SUBJECTS (12000-13110, except 13000-13014)

AIR CONDITION/HEATING/PRESSURIZATION 22
ANTI-ICE/DEICE SYSTEM 22
AUTOPILOT/FLIGHT DIRECTOR 22
COMM/NAV EQUIPMENT 22
ELECTRICAL SYSTEM 23
FIRE EXTINGUISHER 23
FIRE WARNING SYSTEM 23
FLIGHT/NAV INSTRUMENTS 24
HYDRAULIC SYSTEM 24
OTHER SYSTEM(s) 24
OXYGEN SYSTEM 24

i

Revised 12/98

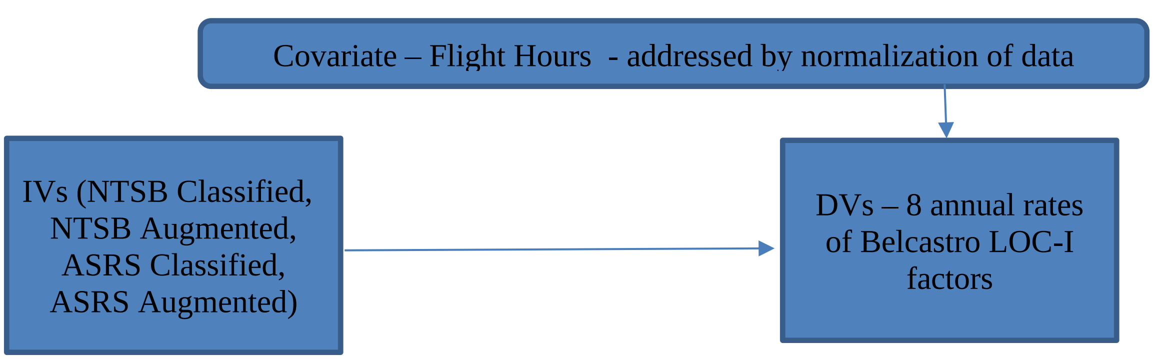
Step 3: Quantitative Data Analysis

Once the taxonomy mapping table was obtained, the already coded data elements from the NTSB and ASRS classified and augmented searched LOC-I reports were mapped against the Belcastro LOC-I Hazards (DVs), the DVs. The frequencies of the DVs were taken from Belcastro LOC-I Hazard counts directly from the LOC-I reports. The frequencies were subsequently normalized by annual hours flown per certification type, known as the *hazard rate* in this study. MANOVA was performed based on the Dependent Variables (DVs) and Independent Variables (IV) listed in Table 5. The dependent variables for the MANOVA were grounded on the eight Belcastro LOC-I Hazards, per Belcastro et al. (2018). The usage of rates, rather than frequency data, for the DVs, was grounded in the works performed by Anderson (2013) that drew the relationship in occurrence rates between various certification types of air accidents, the purpose of which was similar to this research.

The objective of the MANOVA was to answer multivariate and univariate questions (Hair et al., 2019) generated from RQs 1 and 2. The research questions were addressed by analyzing the differences in the means of eight DVs over four groups in the one IV. The DVs were the normalized rates of eight Belcastro LOC-I Hazards from each data group. Apart from analyzing the DVs in a multivariate manner, such DVs needed to be analyzed in a univariate manner; the MANOVA supported this by exerting control over the error rate (Hair et al., 2019). The IV consisted of four LOC-I safety report groups for each operational certification type: two from ASRS and two from NTSB databases. The already classified LOC-I cases and augmented searched LOC-I cases were extracted within each database. The MANOVA was run based on statistical relationships specified in Hair et al. (2019). This is captured in Figure 16. In terms of the covariate, Anderson (2013) highlighted that the number of flight hours per year for each certification category could be a possible covariate for this research. This was addressed by the normalization of the data into hazard rates. Hence, the number of flight hours was not explicitly identified as a covariate.

Figure 16

Basic Variable Type and Relations MANOVA Model Adopted



Quantitative MANOVA Hazards Rates Analysis

Four SMEs generated a mapping table with the support of the researcher. The procedure for generating this table is in Figure 14. Subsequent to this, the identified ASRS and NTSB LOC-I safety reports from LOC-I classification search or augmented search were coded using the mapping table as per Table 5. The Belcastro LOC-I Hazard coding was performed such that count data was obtained using the mapped code from each report. The data was captured in a Microsoft Excel® spreadsheet. Normalization took place by dividing the hazard count by the number of flight hours conducted for the operational certification (commercial or general aviation) for the designated year. For example, for an ASRS case, if two counts of external hazards and disturbances had been coded for a particular year, and the number of hours flown for the year was 100, then the rate for this hazard would be calculated to be 2 divided by 100 (i.e., 0.05).

In terms of tools, Microsoft Excel® was used for initial data gathering, clean-up, and rate calculations, and IBM SPSS® was used for the descriptive statistics and quantitative analysis on a year-to-year basis. Before inputting into SPSS® for analysis, guidance from Chapters 2, 6, and 7 of Hair et al. (2019) was followed when performing the data analysis. Firstly, descriptive statistical analysis was performed to compare the results with Belcastro et al. (2017) on the distribution of hazards. Then, following De Veauz et al. (2013) and Hair et al. (2019), the following generic assumptions were verified before continuing the multivariate data analysis:

a. Linearity – by scatterplots

b. Independence – by checking regression residuals

c. Equal variance – by checking the scatterplots are not thickening

d. Normality – by checking the histogram of residuals and normal probability plot

e. Homoscedasticity of variance-covariance matrices among groups –by conducting Levene's and Box's M tests.

f. Correlation and normality of dependent variables – by conducting Bartlett's test for sphericity to determine whether the dependent measures were significantly correlated.

g. Outliers – by identifying extreme points from Box plots for each group.

Transformations of the datasets needed to be considered if any of the assumptions had not been made. For MANOVA, Hair et al. (2019) stated that the assumption of homogeneity of variance-covariance matrixes across the groups is important. Once the assumptions were verified, the data were analyzed using IBM SPSS®.

Discriminant Analysis

Subsequent to the MANOVA post hoc analysis, the last part of the quantitative analysis involved discriminant analysis, which validated the univariate results. The analysis assessed the individual outcome variables (Belcastro LOC-I Hazards) in terms of their differences across the treatment variables (safety report type), per Hair et al. (2019). The objective was to *profile* the outcome variables in terms of their differences between groups of the treatment variable. Hair et al. (2019) stated that this analysis is particularly insightful when the treatment variable has three or more levels, as in this study. The results from the discriminant analysis were used to validate the univariate MANOVA results as an alternative to repeating the analysis using Tukey's and Scheffe's methods

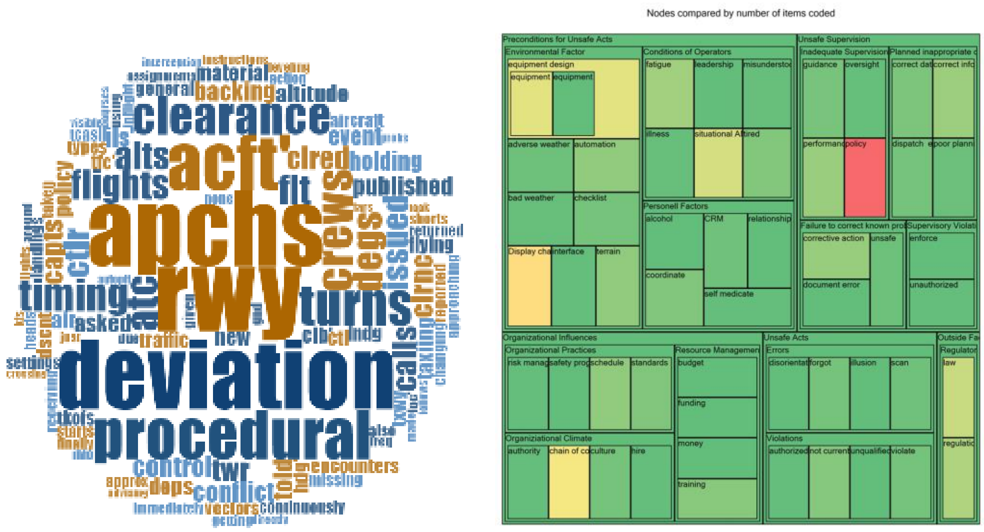
(Hair et al., 2019). This research used discriminant analysis to validate the post hoc univariate analysis.

Step 4: Qualitative Data Analysis

With the abundance of textual data obtained when identifying the coded data for the MANOVA analysis, a limited qualitative data analysis was attempted to identify possible contextual insights to explain the findings made in the quantitative assessment for validation in future research. The textual synopses and narratives of the ASRS, NTSB, and AIDS reports were imported into NVivo® to supplement the quantitative results. Qualitative data analysis techniques were used to analyze the narrative cause descriptions within the ASRS and NTSB reports. Word clouds, cluster maps, hierarchy charts, and tree maps (Bazeley, 2013) were created to represent the dominant factors and node clusters graphically; examples can be found in Figure 17. Upon the results, the qualitative data analysis provided sets of related words from the narrative portion of the report and identified clusters of similar circumstances, possible patterns, relationships among type reports, and LOC-I severities.

Figure 17

Examples of Word Cloud and Tree Map from Previously Conducted Analysis on Air Safety Reports (Lee, 2017)



Note. Reprinted from Commercial Aviation Air Safety Reports Human Factors Analysis by R. Lee. *Ph.D. in Aviation DAV 726 Assignment One.* Copyright 2015 by Embry Riddle Aeronautical University. Reprinted with permission.

Reliability Assessment Method

The instrument that required reliability assessment is the mapping table used to map ASRS, and NTSB taxonomies with the Belcastro et al. (2018) adapted LOC-I model. As this research aimed to identify if the means of the hazard rates between the two reporting systems are the same, it is not necessary to explore the temporal relationship. The design of interrater reliability methodology for the taxonomy mapping table references the attribute agreement analysis developed by Yesilbas (2014) and other studies documented in the literature review (Anderson, 2013). Four raters were deployed to ensure the coding was performed to minimize the level of biases and to increase the level of consistency, per Figure 14.

Validity Assessment Method

Internal validity refers to the degree to which the research design and evidence gathered will answer the research questions (Vogt, 2012). The research questions centered on whether ASRS effectively identified the same Belcastro LOC-I Hazards as NTSB. If this is valid, the likelihood of a higher severity LOC-I event could be reduced based on ASRS data alone. To answer the questions, quantitative supplemented by qualitative methods were used to compare the LOC-I hazard rates identified in the ASRS and NTSB datasets. ICAO’s classification system of *incident*, *serious incident*, and *accident*, which is deployed globally (ICAO, 2016), has been used to ascertain the event's severity level. The selection criteria of LOC-I events benefitted from the already classified LOC-I events in the respective databases as well as the augmented precursors’ keyword search to identify LOC-I events of varying severity not classified as LOC-I, as performed previously by Belcastro et al. (2017). Therefore, it is argued that the selection and assessment of LOC-I reports are valid and comprehensive. Secondly, high internal validity means the changes in DVs are caused solely by the manipulation of the IV. One covariate, the annual flight hours, had already been addressed through the normalization of the data. Thirdly, an interrater reliability test was conducted to test the reliability of the taxonomy mapping per the scholar-reviewed methodology published in Yesilbas (2014). Fourthly, extracted ASRS and NTSB data were coded by professionals in ASRS and NTSB, which provided confidence in the validity.

External validity refers to whether the research can be generalized to other contexts (Leedy & Ormand, 2013). IATA has identified LOC-I as one of the top accident categories in modern aviation. Also, much research has already been undertaken on this topic, such as by Belcastro et al. (2014, 2016). A real-life setting of actual LOC-I accidents and low-severity LOC-I safety report data over seventeen years was used in this study. The ASRS and NTSB samples are representative of United States registered commercial and general aviation operations, which is a matured aviation market. It was noted that the samples might not be representative of a less mature market; however, this was not a concern for this research, as the purpose was to explore the efficacy of an open-loop VSR system in a LOC-I context. In a mature market like the United States, there is a relatively stable market with fewer confounding variables such as language barriers and regulatory differences. It is acknowledged that other VSRs with more rigor in the investigation could be deployed, such as ASAP. This type of VSR could be used in future research using the same methodology should coded data be readily available. Secondly, the methodology adopted would be equally applicable to other safety events modeling, such as Controlled Flight into Terrain (CFIT) or runway excursion, as the datasets are readily available.

Summary

This chapter explained that the research was primarily a quantitative analysis based on MANOVA, with the results supplemented by discriminant analysis and qualitative analysis. The chapter described a four-step process with the associated rationale: starting with the taxonomies mapping by four SMEs to map the ASRS and NTSB taxonomies with Belcastro LOC-I Hazards, conducting the MANOVA with post hoc analysis to obtain the multivariate and univariate results, completing discriminant analysis to verify the univariate results, and lastly engaging in qualitative analysis using NVivo® to generate insights in the reasons behind the quantitative analysis.

The methodology answered RQ1 and RQ2 with their prescribed hypotheses. The results will help to assess the strengths, weaknesses, and relevance of ASRS as one of the credible sources in the safety management of high-severity LOC-I events. The methodology will also assess Heinrich's common cause principles in the context of LOC-I in modern commercial and general aviation. With the estimated sample sizes, the quantitative analysis was expected to have adequate statistical power for a reasonable effect size for the multivariate and univariate analysis.

Chapter IV: Results

This chapter documents the results of the analyses. The research questions were answered by testing associated hypotheses with a one-way MANOVA using IBM SPSS®. The MANOVA examined multivariate and univariate differences in the means of eight Belcastro's LOC-I Hazard rates (DVs) among four types of safety reports, represented by the IV (NTSB Classified, NTSB Augmented, ASRS Classified, ASRS Augmented). Preliminary checks assessed normality, outliers, linearity, homogeneity of variance-covariance matrices, and multicollinearity. The need for the transformation of datasets was assessed. The MANOVA was performed in the order of multivariate, univariate, post hoc analyses, and hypotheses testing. The univariate analysis was validated by discriminatory analysis. Insights into the quantitative results were identified from word clouds, tree maps, and hierarchy charts analyses using NVivo®. The narrative information was analyzed to assess contextual differences across safety report types.

Demographics

Regarding demographics, the 17 years of data covered general aviation and commercial flight LOC-I events within the United States, its territories, and possessions, and in international waters that were reported to the NTSB and ASRS databases (NTSB, 2021). The case numbers for each dataset are documented in Table 4. In her study using safety reporting data to assess the effects of aircraft certification rules on general aviation accidents, Anderson (2013) highlighted the potential threat of Visual Flight Rules (VFR) flights into Instrument Meteorological Conditions (IMC). She detected a cluster of take-off and landing phases of the flights for LOC-I cases. Therefore, two demographic characteristics were also reviewed: phase of flight and flight condition. For each type of safety report, the distributions of the two attributes were analyzed with the results presented in Appendix E. Pie charts represented the distributions of flight conditions and are summarized in Table 6. On examination, VMC was the dominant flight condition from all types of reports occupying a minimum 62% for the ASRS Parts 121 and 135 classified group and a maximum of 97% for the NTSB Part 91 augmented group.

Table 6

Proportion of Flight Conditions from LOC-I Safety Reports

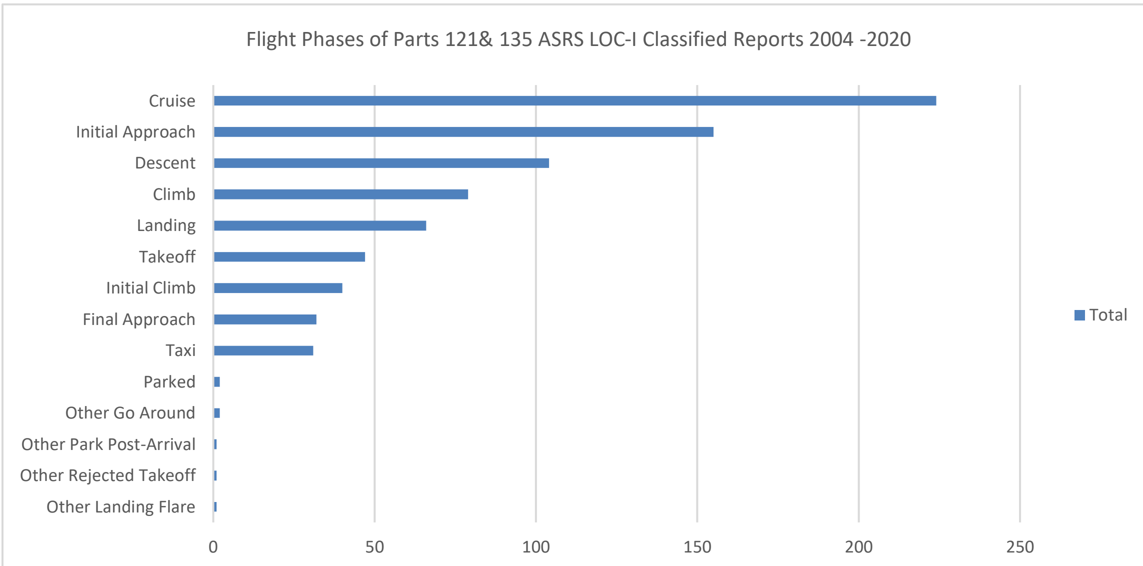
		Part(s)	VMC	IMC	Marginal	Mixed
Classified	ASRS	91	71%	23%	1%	5%
		121 and 135	62%	28%	5%	5%
	NTSB	91	93%	7%	-	-
		121 and 135	96%	4%	-	-
Augmented	ASRS	91	85%	10%	3%	2%
		121 and 135	75%	19%	2%	4%
	NTSB	91	97%	3%	-	-
		121 and 135	83%	17%	-	-

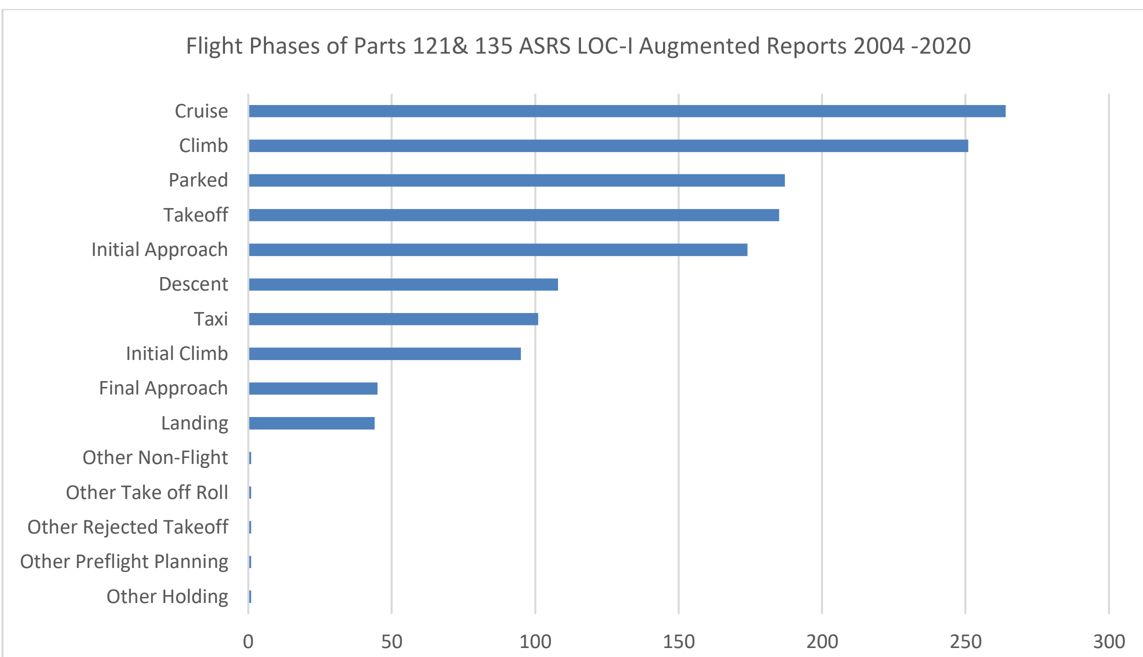
Bar charts were used to show the distribution of the flight phases for each LOC-I safety report type. A full presentation of the demographics is documented in Appendix E, with examples from the Parts 121 and 135 dataset in Figure 18. The top five flight phases of the safety report types are listed in Table E.1 in Appendix E. NTSB reports tended to be focused on the descent and landing phases. The cruise was a typical phase reported with

LOC-I for all types of safety reports, apart from Parts 121 and 135 NTSB augmented group. ASRS reports covered most flight phases, from takeoff to descent and landing.

Figure 18

Examples of Flight Phases Data from the Parts 121 and 135 ASRS Dataset





Taxonomy Mapping

This section describes the process of creating a mapping table that maps the ASRS and NTSB LOC-I reports' coded data elements with the Belcastro LOC-I Hazards for subsequent quantitative analysis. As this study involved two different safety reporting system codes using different taxonomies, the analysis depended upon having a common taxonomy for LOC-I events or potential LOC-I events based on Belcastro LOC-I Hazards. Measurements of these hazards formed the MANOVA DVs set.

Identification of Data Elements

Preparations were made to identify categories of codes that resembled Belcastro LOC-Hazards, known as *data elements* in this research, from ASRS and NTSB databases. This required assessment by the researcher as the structure of ASRS and NTSB databases are fundamentally different. For the ASRS database, the Primary and Contributory Factors and Anomaly codes categories were selected. For the NTSB database, the Subject Code and Sub-Category Codes were selected for post-2008 reports, and Subject Code and Modifier Code were selected for pre-2008 reports due to the eADMS system change. The researcher has made the best attempt to identify the relevant data elements (see Table 7). This might not be exhaustive due to the volume of categories of data in each safety reporting database. This research focused on the impact of IV (safety report types) on already mapped DVs (Belcastro LOC-I Hazards). The relative change across the IV groups on available mapped Belcastro LOC-Hazard was being assessed, not the number of data elements mapped. Hence, the risk of non-exhaustive identification of data elements is mitigated.

Table 7

Data Elements Extracted from ASRS and NTSB Code Categories for SME Panel Mapping

ASRS	NTSB Database
Primary Factor	findings_category_no
Contributory Factor	findings_subcategory_no
Anomaly	findings_section_no
	findings_subsection_no
	finding_modifier_no

After obtaining the LOC-I reports from the relevant databases, the codes under each data element were extracted from the ASRS and NTSB databases for the identified LOC-I events. These codes were inputted into the online Microsoft Forms® platform. The platform was used for taxonomy mapping exercises by an SME panel.

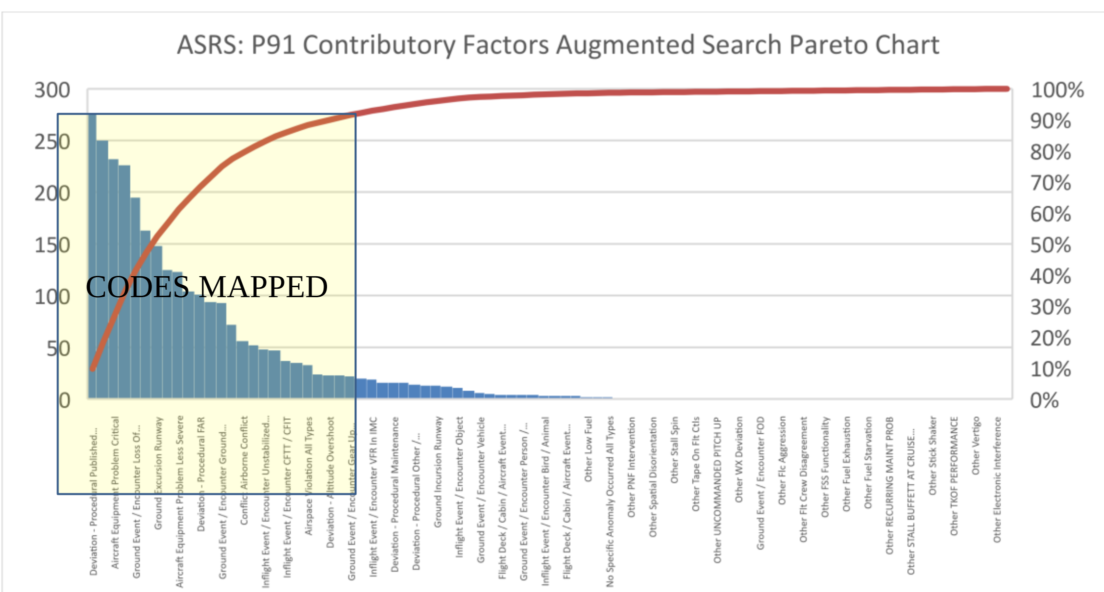
Preparations of Codes To Be Mapped

There were over one thousand unique codes from twelve (four ASRS, eight NTSB) LOC-I sub-datasets. These sub-datasets were consolidated into four ASRS and four NTSB datasets. Each dataset represented a unique severity type, denoted by search method *classified* and *augmented*, safety report type, *ASRS* or *NTSB*, and operational certification type, *Parts 121 and 135* or *Part 91*, respectively. Datasets obtained were labeled as the four groups of the IV: NTSB Classified, NTSB Augmented, ASRS Classified, and ASRS Augmented. A large number of unique codes was partially due to the change of NTSB's eADMS coding system in 2008, necessitating this research into two initial sets of codes to process for NTSB cases, one set pre- and one set post-2008. Due to the finite time allocated to the mapping exercise with the SME panel, it was unrealistic to map all codes without a structured and prioritized manner. Therefore, an enhanced Pareto approach (Grami, 2020) in mapping data element codes from the ASRS and NTSB databases that covered a minimum of 95% of all identified LOC-I code counts (classified and augmented) was adopted, as explained in Chapter III.

An example of the Pareto approach is illustrated in Figure 19. The Y-axis on the right shows the percentage of the overall number of coded data elements accumulatively from the ASRS Part 91 Augmented Search dataset. The Y-axis on the left shows the code count. The X-axis is the labels of the data elements for one IV group. Due to resolution issues, only a portion of the code labels is shown on the X-axis. From the graph, the factors that contributed to 95% or above of the total number of data elements codes for Part 91 ASRS augmented search LOC-I dataset had been selected for mapping by the SME panel to form the Belcastro LOC-I Hazards DVs mapping table.

Figure 19

Illustrations of Codes Selection for Mapping Based on the Pareto Chart for ASRS Part 91 Augmented Search Dataset



Taxonomy Mapping Arrangements

The codes originated from the data elements from each safety database and were transferred to the SME panel for assessment. The panel assessed if mapping the code to Belcastro LOC-I Hazards was warranted and ensured interrater reliability was reached in this decision. Per the planned methodology, an SME panel gathered eight times from September 25, 2021, to November 6, 2021, to conduct the taxonomy mapping exercise online. Seven industry SMEs had initially agreed to support the exercise when invited. The SMEs are experienced aviation professionals and have held leadership positions in aviation safety management, piloting, and / or have been professional flight crew members of international flights. Brief biographies of the SMEs are documented in Appendix F. Due to the SMEs' operational challenges imposed by COVID-19, only four of the seven SMEs attended the scheduled workshops and performed all the required mapping. The Kappa result was calculated based on inputs from the four SMEs who attended all the workshops and performed all mapping. The other three SMEs who only attended some of the workshops provided valuable opinions and advice during the workshops but, for consistency, did not contribute towards the mapping and Kappa calculations.

Before each workshop, the SMEs were requested to perform pre-reading and online mapping using Microsoft Forms®. The researcher then presented the results, and the differences were discussed in the subsequent workshop. The SMEs typically performed the coding again until the value of Kappa was more than 0.7, as calculated by IBM SPSS®. A Microsoft Teams® group was also set up for general communications and support for the SMEs during the mapping exercises. Appendix G contains screenshots of the Microsoft Teams® group setup, online workshop footage, Microsoft Forms®, and a chronological summary of the SMEs mapping activities.

Mapping Results

17,750 LOC-I reports have been identified using classified and augmented searches from the ASRS and NTSB databases. The codes extracted from the LOC-I cases that had undergone Pareto analyses were entered into Microsoft® Forms, one form per dataset. Four members of the SME panel followed the mapping program above consistently. The SME Panel mapped the data element codes against the eight Belcastro LOC-I Hazard counts, normalized by the annual flight hours for the operational certification to convert into *rates*. These rates were the dependent variables (DVs) of this study. Each Belcastro LOC-I Hazard was coded by a numerical DV code, as indicated in Table 8.

Table 8

Numerical Mapping Codes with Belcastro LOC-I Hazards

Mapping Code (DV) Identification	Belcastro et al. (2018) Description
1	Adverse onboard conditions - Vehicle Impairment
2	Adverse onboard conditions - System and components failure / malfunction
3	Adverse onboard conditions - Crew action / inaction
4	External hazards and disturbances - Inclement weather atmospheric disturbances
5	External hazards and disturbances - Poor visibility
6	External hazards and disturbances - Obstacle
7	Abnormal vehicle dynamics and upsets - Abnormal vehicle dynamics
8	Abnormal vehicle dynamics and upsets - Vehicle upset conditions

The LOC-I reports within the sample frame contained 35,500 counts of codes from the respective ASRS and NTSB databases. Some of those codes appeared repeatedly. The SME Panel successfully mapped 422 unique codes from the ASRS and NTSB databases with the eight Belcastro LOC-I Hazards, which were the DVs of this research (see Table 8). The identities of the codes and the mapped Belcastro et al. (2018) DVs are listed in Appendices H and I. For the quantitative analysis, the normalized rates based on annual flight hours per certification type of the mapped DVs would be made for each dataset.

During the mapping exercise, SMEs found it challenging to code between DV (1) and (2) per Table 7 as both DVs were in the same main category, which described failures on an aircraft, though the severity was different. However, the source database codes did not refer to the severity of the failure. Particularly for ASRS codes, the SMEs had to extrapolate the causes of the symptoms rather than the symptoms themselves for more accurate mapping with the eight Belcastro LOC-I Hazards DVs.

Reliability Testing

As described, a panel of four SMEs was involved in the taxonomy mapping exercise. Interrater reliability Kappa value of > 0.7 was to be met before the iterations of mapping were deemed complete. Table 9 highlights the Kappa value results for each database and the number of codes the experts successfully mapped to the Belcastro LOC-I Hazards. In total, 422 unique codes retrieved were mapped. These codes originated from 35,500 repeat appearances within the selected data elements (see Table 7) in ASRS and NTSB databases based on classified and augmented searched LOC-I events. For the mapped codes, Kappa values were more than 0.7, with an alpha value of less than 0.05. See Table 9.

Table 9

Interrater Reliability Statistics From the NTSB and ASRS Taxonomies Mapping Exercise by SME Panel

	ASRS		NTSB			
	Primary and Contributory Factors	Anomaly Code	2008 to 2020 Subject Code	2008 to 2020 Subcategory Code	2004 to 2007 Subject	2004 to 2007 Modifier
Data Elements Code Counts	4583	4583	6289	6289	6878	6878
Number of codes mapped	16	47	4	20	128	207
Concluding Kappa	0.90	0.71	0.82	0.86	0.72	0.81
Standard Error	0.02	0.01	0.07	0.03	0.01	0.01
Significance Level	0.000	0.000	0.000	0.000	0.000	0.000

Quantitative Analysis – MANOVA

The primary purpose of this study was to answer the research questions, which analyzed the similarities or differences in the means of hazard rates between NTSB and ASRS databases for classified and augmented searched LOC-I. The core of this analysis was supported by a MANOVA analysis with one IV containing four groups and eight DVs. The IV represented four types of LOC-I safety reports differing in the source database or case identification method: NTSB classified, NTSB augmented, ASRS classified, and ASRS augmented. Classified search reports were reports from relevant databases already classified as LOC-I events. Augmented search reports were identified from the LOC-I keyword search used in Belcastro et al. (2017). The DVs collected were the normalized annual rates of eight mapped Belcastro LOC-I Hazards, as introduced in Chapter III. These were retrieved from data elements (see Table 7) in identified LOC-I reports. The normalization factors used were the annual flight hours of the relevant operational certification. The normalization calculation is demonstrated in the descriptive statistics section.

The MANOVA was based on the guidance given by Hair et al. (2019) and Field (2020). Firstly, the collected data was examined. The graphical method was adopted to examine the characteristics of the data or relationships of interest. Then, the potential impact of missing data was assessed. Subsequently, univariate and multivariate outliers were examined, and assumptions of normality, linearity, multicollinearity, and homogeneity were tested. Due to the violations of some assumptions, data transformation was performed and explained in this chapter. Subsequently, the estimation of the MANOVA model and assessment of the overall fit was carried out using statistical significance testing. The results were interpreted by assessing the effects of the IV with multivariate, univariate, and post hoc tests. Discriminant analysis was carried out to validate the univariate results.

Descriptive Statistics

As described in Chapter III, it was unknown if the number of flight hours influenced the variation of the eight Belcastro LOC-I Hazard rates, which were the DVs in this study. As the number of flight hours differed each year, the hazard rate counts were normalized to hazard rates per year for each DV to eliminate the effect of this possible covariate. It was impossible to formally add the flight hours as a covariate in the MANOVA analysis as Hair et al. (2019) stated a condition where covariates could be added, as follows:

$$\textit{maximum number of covariates} = (0.10 \times \textit{Sample size}) - (\textit{Number of groups} - 1)$$

For this study, the sample size per cell was seventeen, and the number of groups was four. Therefore, no covariate could be deployed in the MANOVA analysis. Thus, normalized data were the most appropriate method to treat the possible effect of the flight hours covariate.

The normalization factor was derived from the flight hours' data from the BTS database for the Parts 121 and 135 operational certification and the FAA GOA survey for the Part 91 operational certification. The mapped Belcastro LOC-I code counts, flight hours per type of operation obtained from BTS and FAA GOA databases, and the normalized rates are documented in Appendix J. The appendix shows 17 years of data (2004 to 2020) for each type of report, denoted as *CODE_TYPE*. The definition of this categorical variable (IV) can be found in Table 5. The formula of the normalization adopted was

$$N = C / H$$

where N was the normalized rate, C was the count of the DV occurrence, and H was the flight hours per calendar for the operational certification, denoted by the IV, CODE_TYPE. For example, from the Parts 121 and 135 dataset, external increment weather (DV4) had occurred 16 times in 2004 (C), the annual flight hour for this year was 21,338,088 hours (H), and the normalized rate (N) was therefore calculated to be 7.49E-07 counts per flight hour.

Two data points for *H* were missing from the FAA GOA database. The FAA's 2011 GOA data were unavailable, and FAA recommended that 2011 data be taken as an extrapolation from the forecast (FAA, 2021). When this analysis was performed (i.e., December 2020), the FAA had yet to publish the 2020 GOA survey results. Hair et al. (2019) indicated that if there are less than ten percent missing values, any missing value imputation methods could be applied. Therefore, an estimation was made based on NTSB data that, due to the emergence of COVID-19, the levels of commuter aviation (Part 135) and general aviation (Part 91) had been reduced by 46% and 11%, respectively (NTSB, 2021). An asterisk annotated these extrapolations in Appendix J1.

The normalized eight Belcastro LOC-I Hazard rates (DVs) for all four datasets (four groups in one IV) for commercial (Parts 121 and 135) and general aviation (Part 91) were analyzed. It was of note that some DVs had zero coded data for specific IV types. These were not missing data but indicated that Belcastro LOC-I Hazard was not found for the specific type of safety report and Belcastro LOC-I Hazards, summarized in Table 10.

Table 10

DVs with Zero Frequency in Parts 121 and 135 and Part 91 Datasets

Dataset	DV	IV	Frequency
Parts 121 and 135	DV5 - External hazards and disturbances - Poor visibility	NTSB AUGMENTED	0
		ASRS CLASSIFIED	0
		ASRS AUGMENTED	0
	DV7 - Abnormal vehicle dynamics and upsets - Abnormal vehicle dynamics	NTSB AUGMENTED	0
		ASRS AUGMENTED	0
Part 91	DV5 - External hazards and disturbances - Poor visibility	ASRS CLASSIFIED	0
		ASRS AUGMENTED	0
	DV8 - Abnormal vehicle dynamics and upsets - Vehicle upset conditions	NTSB AUGMENTED	0

The descriptive statistics of the two datasets are presented in Table 11. Due to the magnitude of the normalization factor, the descriptive statistics for the hazard rates were distributed with means and standard deviations between five to nine decimal places in value. The number of samples, *N* value, remained constant at seventeen as seventeen years of hazard reporting rates were obtained for this analysis.

Table 11

Descriptive Statistics for DVs and IVs Used in the MANOVA from 2004 to 2020

DV	IV Group[a]	Parts 121 and 135 Mean($\times 10^{-8}$)	Parts 121 and 135 Standard Deviation($\times 10^{-8}$)	Part 91 Mean ($\times 10^{-8}$)	Part 91 Standard Deviation($\times 10^{-8}$)
DV1 - Adverse onboard conditions - Vehicle Impairment	N-CFD	97.5	68.5	756.0	1383.0
	N-AUG	9.5	11.1	95.2	103.5
	A-CFD	39.3	46.9	74.3	45.1
	A-AUG	177.3	73.9	69.1	26.8
DV2 - Adverse onboard conditions - System and components failure / malfunction	N-CFD	7386.7	4921.9	1650.0	662.5
	N-AUG	333.2	310.4	5130.0	3665.0
	A-CFD	122.4	139.7	227.0	120.2
	A-AUG	469.3	100.2	184.0	76.1
DV3 - Adverse onboard conditions - Crew action / inaction	N-CFD	14766.0	9373.5	8020.0	5338.0
	N-AUG	9946.6	137.2	7100.0	4993.0
	A-CFD	19585.4	23274.0	909.0	383.3
	A-AUG	15106.0	23136.8	750.0	371.8
DV4 - External hazards and disturbances - Inclement weather atmospheric disturbances	N-CFD	1535.5	957.3	1570.0	1830.0
	N-AUG	1043.3	59.6	570.0	392.8
	A-CFD	2027.7	2819.5	333.0	101.5
	A-AUG	1546.2	2759.9	126.0	67.9
DV5 - External hazards and disturbances - Poor visibility	N-CFD	3.5	7.2	183.0	349.8
	N-AUG	0	0.0	0.6	2.4
	A-CFD	0	0.0	0.0	0.0
	A-AUG	0	0.0	0.0	0.0
DV6 - External hazards and disturbances - Obstacle	N-CFD	1496.0	967.9	3320.0	5834.0
	N-AUG	26.3	23.7	425.0	344.8
	A-CFD	107.2	119.4	91.3	34.4
	A-AUG	96.0	46.3	37.8	18.5
DV7 - Abnormal vehicle dynamics and upsets - Abnormal vehicle dynamics	N-CFD	0.8	2.4	58.4	110.3
	N-AUG	0.0	0.0	0.6	2.4
	A-CFD	235.7	52.5	344.0	154.3
	A-AUG	0.0	0.0	0.3	1.4
DV8 - Abnormal vehicle dynamics and upsets - Vehicle upset conditions	N-CFD	0.3	1.1	186.0	347.3
	N-AUG	0.3	0.0	0.0	0.0
	A-CFD	56.3	4.7	23.1	11.2
	A-AUG	27.4	4.7	3.4	4.1

Note. N = 17 for each IV group

[a]N-CFD is NTSB Classified, N-AUG is NTSB Augmented, A-CFD is ASRS Classified, A-AUG is ASRS Augmented

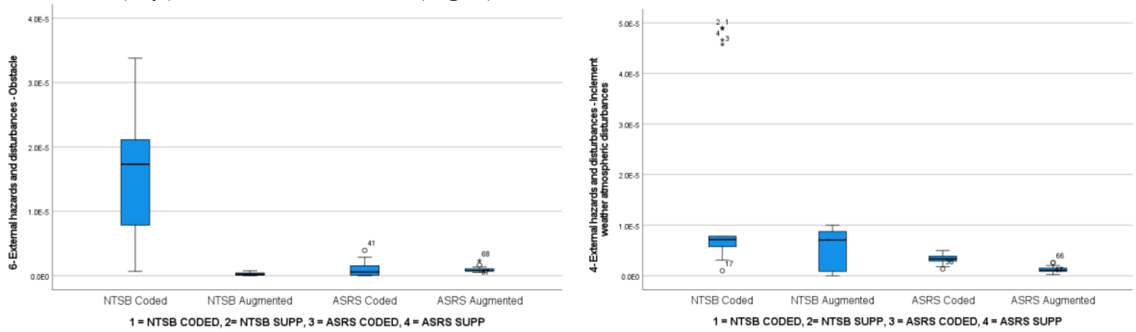
Data Assumptions

Before conducting the MANOVA analysis, the assumptions on normality (univariate and multivariate), outliers (univariate and multivariate), linearity, homogeneity of variance-covariance matrices, and multicollinearity were tested. This testing aligns with the assumptions testing requirement specified by Hair et al. (2019). Scattered plots and histograms of DVs using normalized rates data were produced for the Parts 121 and 135, and Part 91 datasets, as documented in Appendix K.

Outliers. First, univariate outliers were tested. Initially, the datasets were visually examined using the univariate method. Box plots were produced using IBM SPSS®. Outliers were found. Some outliers were within the moderate outlier range (i.e., third quartile plus three interquartile ranges and first quartile minus three interquartile). These were annotated with circles. Some outliers, annotated by an asterisk, were outside this range and were extreme outliers, which provided cause for concern (Hair et al., 2019). Box plots of the Parts 121 and 135 dataset were captured in Appendix K1. From the univariate perspective, six moderate and nine extreme outliers were found in all DVs apart from DV2. For the Part 91 dataset, univariate outliers were found in all eight DVs, as indicated in Appendix K4. Examples of the box plots with univariate outliers are documented in Figure 20.

Figure 20

Examples Box Plots Illustrating Univariate Moderate and Extreme Outliers, Part 121 & 135 DV6 (left) and Part 91 DV4 (right)



As MANOVA is a multivariate analysis, per Hair et al. (2019), in addition to the univariate review using box plots, a multivariate outlier analysis based on Mahalanobis D^2 measurement would be required. The linear regression function of IBM SPSS® was used to perform the Mahalanobis D^2 analysis, with the dependent variable as years and the independent variables as DV1 to DV8. The Mahalanobis D^2 values were added to the IBM SPSS® data file. With eight DVs, Tabachnik and Fidell (1996) stated that the critical value of Mahalanobis D^2 was 26.13, the maximum value permitted for multivariate normality. For Parts 121 and 135 dataset, two multivariate outliers were detected for 2006 data, and four were detected for Part 91. The Mahalanobis D^2 results with the data points that were beyond the critical values are documented in Table 12.

Table 12

Mahalanobis D^2 Multivariate Outlier Analysis Results for Parts 121& 135 and Part 91 Datasets

Dataset – Part(s)	IV Group	Year	Mahalanobis D^2
121 and 135	1	2004	34.60031
121 and 135	1	2008	51.75192
91	1	2007	54.41121
91	1	2005	47.02656
91	1	2004	63.94161

Hair et al. (2019) defined outliers as observations with a “unique combination of characteristics identifiable as distinctly different from what is ‘normal’” (p.85). Outliers could be problematic or beneficial as beneficial outliers would indicate population characteristics that would not be discovered in the normal course of analysis. In contrast, problematic outliers would counter the objectives of the analysis and could seriously distort statistical tests (Hair et al., 2019). The univariate and multivariate outliers were considered and retained as they were not aberrant. The outliers represented the observations in the data recorded for the years concerned. The data collection and normalization process had been verified. The outliers were confirmed not to be error outliers originating from procedural errors, as the procedure had been rechecked for such data points. No observations were extreme on a sufficient number of variables to be considered unrepresentative of the population (Hair et al., 2019). Also, no transformation had taken place to reduce the impact of the outliers at this stage. Instead, they were *interesting outliers* that were different such that they may bring new insight into the analysis (Hair et al., 2019) as they reflected the actual results. However, due to the presence of univariate extreme outliers, although the outliers were retained at this stage, the transformation of the model was required. This is covered later in this chapter.

Normality. Univariate normality was tested using statistical methods. DVs with normalized Belcastro's LOC-I Hazard rates were obtained from the Parts 121 and 135, and Part 91 datasets. Kolmogorov-Smirnova and Shapiro-Wilk tests for normality were conducted using IBM SPSS®. The results are documented in Table 13. Based on the Shapiro-Wilk test results, 17 variates in the Parts 121 and 135 dataset and 19 variates in the Part 91 dataset demonstrated *p* values were less than .05. This meant the null hypotheses that the variates were normally distributed were rejected. Moreover, due to the absence of data for DV5 and DV7 in the Parts 121 and 135 dataset, viable distributions across all IVs were not available. These gave cause for concern for further consideration. In summary, Shapiro-Wilk tests indicated that DVs were not normally distributed in all the groups.

The multivariate normality was tested by Mahalanobis D^2 using an identical critical value as the aforementioned outlier test. The majority of the Mahalanobis D^2 values from the Parts 121 and 135 and Part 91 datasets were below the critical value of 26.13 but had some values above. This result suggested a reduced level of multivariate normality for the datasets, also suggesting the need for data transformation.

Table 13

Normality Tests Based on Kolmogorov-Smirnova and Shapiro-Wilk for Parts 121 and 135, and Part 91 Datasets

		Parts 121 and 135				Part 91			
		Kolmogorov-Smirnov		Shapiro-Wilk		Kolmogorov-Smirnov		Shapiro-Wilk	
DV	IV Group	Statistic	Sig.	Statistic	Sig.	Statistic	Sig.	Statistic	Sig.
1	1	0.147	.200*	0.947	0.418	0.464	0.000	0.553	0.000^
	2	0.274	0.001	0.794	0.002^	0.194	0.090	0.853	0.012^
	3	0.262	0.003	0.782	0.001^	0.155	.200*	0.904	0.079
	4	0.116	.200*	0.938	0.291	0.094	.200*	0.984	0.985
2	1	0.261	0.003	0.790	0.002^	0.236	0.013	0.897	0.062
	2	0.199	0.073	0.868	0.020^	0.238	0.011	0.809	0.003^
	3	0.228	0.019	0.839	0.007^	0.173	0.190	0.862	0.017^
	4	0.106	.200*	0.988	0.996	0.130	.200*	0.959	0.617
3	1	0.294	0.000	0.759	0.001^	0.327	0.000	0.810	0.003^
	2	0.156	.200*	0.884	0.037^	0.184	0.131	0.864	0.018^
	3	0.118	.200*	0.959	0.616	0.225	0.023	0.886	0.040^
	4	0.173	0.190	0.952	0.483	0.223	0.024	0.874	0.025^
4	1	0.272	0.002	0.826	0.005^	0.432	0.000	0.627	0.000^
	2	0.214	0.037	0.845	0.009^	0.194	0.089	0.826	0.005^
	3	0.287	0.001	0.823	0.004^	0.108	.200*	0.979	0.946
	4	0.140	.200*	0.963	0.679	0.179	0.151	0.917	0.134
5	1	0.451	0.000	0.563	0.000^	0.464	0.000	0.580	0.000^
	2	0	0	0	0	0.537	0.000	0.262	0.000^
	3	0	0	0	0	0	0	0	0
	4	0	0	0	0	0	0	0	0
6	1	0.155	.200*	0.929	0.209	0.462	0.000	0.561	0.000^
	2	0.174	0.180	0.914	0.115	0.140	.200*	0.910	0.100
	3	0.189	0.109	0.848	0.010^	0.134	.200*	0.966	0.743
	4	0.198	0.076	0.806	0.002^	0.148	.200*	0.959	0.607
7	1	0.513	0.000	0.391	0.000^	0.467	0.000	0.574	0.000^
	2	0	0	0	0	0.537	0.000	0.262	0.000^
	3	0.137	.200*	0.955	0.539	0.249	0.006	0.856	0.013^
	4	0	0	0	0	0.537	0.000	0.262	0.000^
8	1	0.537	0.000	0.262	0.000^	0.469	0.000	0.559	0.000^
	2	0.537	0.000	0.262	0.000^	0	0	0	0
	3	0.241	0.010	0.802	0.002^	0.125	.200*	0.965	0.720
	4	0.129	.200*	0.953	0.505	0.323	0.000	0.765	0.001^

N = 17

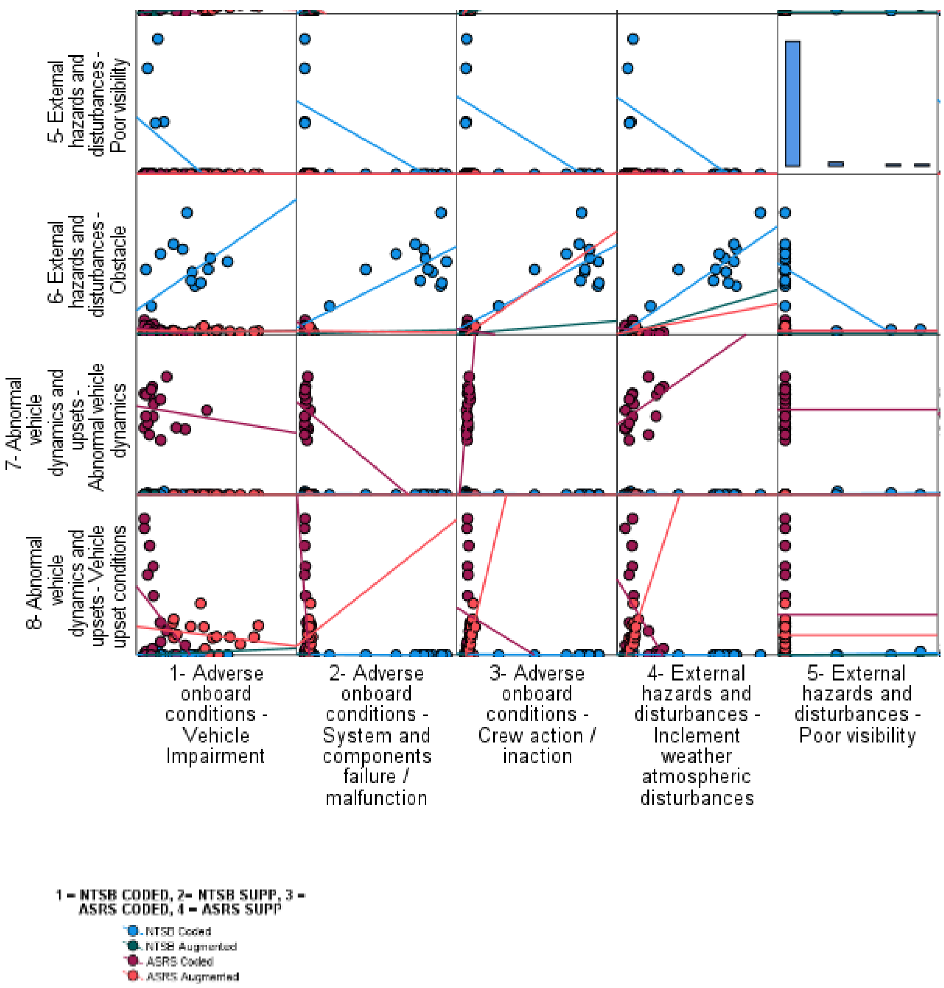
^*p* value < .05 for Shapiro-Wilk tests, *lower bound for true significance per IBM SPSS®

IV: 1 = NTSB Classified, 2 = NTSB Augmented, 3 = ASRS Classified, 4 = ASRS Augmented

Correlation, Missing Data, Linearity and Multicollinearity. Linearity was tested visually by scattered plots. The scattered plots matrix covering the eight DVs and four IV groups for both datasets was recorded in Appendices K3 and K6. A unique color of the plot was used to identify each unique IV group. Hence, the scattered plots matrix showed four colors of plots. A segment of the Parts 121 and 135 scattered plots matrix was captured in Figure 21 for illustration. A linear best-fit line was applied using IBM SPSS®. On visual examination of the Parts 121 and 135 dataset, among the eight DVs, the best-fit lines were generally representative of the data for most DVs apart from DV5 and DV7 due to the zero count of some types of IV in these DVs, suggesting the DVs were approximately linearly related within its group apart from DV5 and DV7, as shown in Figure 21. For the Part 91 dataset, the general level of linearity was similar to Parts 121 and 135 dataset, with the possibility of best-fit lines not representative of the data recorded on DV5 and DV8 due to the zero count for some IV groups. The rest of the relationships for the Part 91 dataset represented approximate linear relationships. Some plots showed a more random pattern, and best-fit lines were less representative of the data. Hair et al. (2019) stated that the linearity assumption was not necessarily broken if no non-linear patterns, such as exponential or parabolic curves, were detected. Therefore, the linearity assumption was generally met for all DVs.

Figure 21

A Segment of the Parts 121 and 135 Dataset Scattered Plots Matrix demonstrating lack of linearity for DV5 and DV7



Multicollinearity is the measure of shared variance with other variates. Hair et al. (2019) stated that the DVs were best moderately correlated with multicollinearity but should not be too high. A high level of correlation was generally defined as Pearson Correlation Coefficient, $r > 0.8$, as this indicated redundant dependent measures and decreased statistical efficiency. Multicollinearity for the datasets was tested by Pearson correlation analysis, documented in Appendices K2 and K5. Four significant relationships were identified for the Parts 121 and 135 dataset, per Appendix K2. Six relationships, DV2/DV3, DV4/DV2, DV4/DV3, DV6/DV2, DV6/DV4, DV6/DV3, were above the moderate level of correlation (i.e., $R > 0.8$) (Hair et al., 2019 pp. 386; Stevens, 2009). For the Part 91 dataset, eighteen significant relationships were identified within the Part 91 dataset, ten relationships DV4/DV1, DV5/DV1, DV6/DV1, DV8/DV1, DV5/DV4, DV6/DV4, DV8/DV4, DV6/DV5, DV8/DV5, DV8/DV6 had Pearson Correlation Index, *R*, to be higher than 0.8. The rest of the relationships were moderate, with $R < 0.8$, $n = 68$, and $p < .05$. Such results gave cause for concern that the assumption for multicollinearity was not met, needing data transformation (Hair et al., 2019), which is discussed later in this chapter.

Regarding missing data, the LOC-I yearly hazard rates were based on the directly coded data extracted from the ASRS and NTSB databases, which comprehensively provided data for the years of interest (i.e., 2004 to 2020). Provided the coding was performed adequately by NTSB and ASRS administrators, which was an assumption to this research, no data acquired by the MANOVA analysis was missing.

Homoscedasticity / Homogeneity of Variance. Homoscedasticity, or homogeneity of variance, is the "assumption that dependent variables exhibit equal levels of variance across the range of predictor variables" (Hair et al., 2019, p. 97). Multivariate homoscedasticity means the variability in the values of the continuous IV is roughly the same across all continuous DVs. Its importance in a multivariate analysis, as explained by Hair et al. (2019), is that the dependent variable explained in a dependence relationship should not be concentrated in only a limited range of the independent values.

Univariate homoscedasticity means variability in the DV is expected to be the same at all levels of the grouping variable (Tabachnick et al., 2007).

Hair et al. (2019) stated that a multivariate homoscedasticity test could be performed by Box's M's test, a sensitivity-adjusted non-significant value of $p > 0.01$ (Hair et al., 2019, p. 372) indicated no presence of heteroscedasticity (i.e., meeting the homogeneity assumption). For the MANOVA analysis, as the IV was nonmetric, the concept of multivariate homoscedasticity referred to the equality of variance matrices (multiple dependent variables) across the groups formed by nonmetric independent variables. Hence, the Box's M test analyzed the variance and covariance matrices. The results of the test are documented in Table 14. For the Parts 121 and 135, and Part 91 datasets, DV5 had to be removed to avoid IBM SPSS® generating error messages on *fewer than two nonsingular cell covariance* due to the zero content of some groups (see Table 15). On examination, the Box's M test results indicated that the assumption of multivariate homogeneity of variance-covariance was not met for both Parts 121 and 135, and 91 datasets, $p < .01$.

Table 14

Multivariate Homogeneity Test Results for Parts 121 and 135, and Part 91 Datasets

	Parts 121 and 135	Part 91
Box's M	479.712	810.063
F	13.050	11.379
df1	28	56
df2	3568.203	6581.056
Sig.	<0.001	<0.001

The results of Levene's test for univariate homogeneity are listed in Table 15, based on the median and with adjusted df results by IBM SPSS®. A significance level of $p < .05$ indicated that the univariate homogeneity requirement had not been met. All DVs for Parts 121 and 135 and Part 91 datasets failed the homoscedasticity assumption. By comparing the size of the box in the box plots in Appendix K1, the failed assumption was illustrated. This further provided the impetus for data transformation.

Table 15

Levene's Test Results for Parts 121 and 135, and Part 91 Datasets

DV	Parts 121 and 135 dataset		Part 91 dataset	
	Levene's Statistics	Significance	Levene's Statistics	Significance
1	7.352	<0.001	4.419	.019
2	14.502	<0.001	16.545	<.001
3	11.689	<0.001	8.625	<.001
4	12.751	<0.001	5.084	.011
5	DELETED	DELETED	DELETED	DELETED
6	20.633	<0.001	4.892	.013
7	56.986	<0.001	11.328	<.001
8	11.832	<0.001	4.664	.016

Independence of Datasets. As the classified and augmented groups originated from the same database, to ensure the independence of cases being analyzed, all unique case numbers were compared among groups using Microsoft Excel®. Duplications were eliminated with the respective dataset. For example, if a unique case number were found in both the Classified and Augmented groups, the case would be retained in the Classified group and eliminated from the Augmented group.

Data Transformation. Hair et al. (2019) and Field (2020) highlighted three possible ways to address the data's lack of normality and homoscedasticity. The first applies trimmed means and bootstrapping, the second uses a robust non-parametric test, and the third is data transformation. In addition, if the sample size for this study was over 30, based on the Central Limit Theorem (Field, 2020), the sampling distribution is expected to be normal. For completeness, all three methods suggested were adopted, as follows:

The descriptive statistics analysis was re-run using bootstrapping with 1,000 samples on IBM SPSS® using Bias Corrected Accelerated (BCA). Bootstrapping estimated the properties of the sample distribution from the sample data (Field, 2020). The distribution within each group was further examined. The analysis indicated that the skewness and kurtosis values were within the bounds of the bootstrapped lower and upper 95% level, making the datasets suitable for further analysis using the bootstrapping technique despite its violation of the normality assumption (Field, 2020). However, there was no provision in IBM SPSS® to implement MANOVA with bootstrapping directly without using additional software such as R (Field, 2020), so the bootstrapping method was not further pursued in this study. The non-parametric tests were also carried out. However, as the non-parametric test had less statistical power than MANOVA, it was not adopted as a preferred method going forward. Therefore, the only viable method remaining was a transformation of data.

Five transformation models, square root, cube root, quartic root, log10(DV + 1), and inverse (DV+1), were trialed in transforming the original datasets, which were highly positively skewed, into distributions with higher normality. The rationale behind the DV+1 was to adjust for the zero data and avoid the one divided by zero error. Transformed variables, prefixed by TX, were created and labeled according to the transformation applied. The transformed DVs were tested against the assumptions earlier in the chapter.

Regarding normality, Kolmogorov-Smirnova and Shapiro-Wilk tests were conducted with complete results and compared among the transformation models, documented in Appendix L1. For the Parts 121 and 135 dataset, cube root transformation models produced the best normality performance with 11 DV and IV type combinations with $p < .05$ instead of 14 with the original model. This indicated an increased level of univariate normality for the transformed model. DV5 and DV7 showed some blank results due to zero data points and, therefore, could not answer some of the hypotheses related to the NTSB dataset. DV5 was removed for this analysis as three Group Types had zero Belcastro LOC-I Hazards rates entries. For the Part 91 dataset, the improvement by transformation was not noticeable. While the square root transformation reduced the total number of extreme outliers from 31 to 29, this was at the expense of 20 significant Shapiro test results instead of 19 from the original dataset. Therefore, in terms of normality, the transformation did not notably improve the original Part 91 dataset. Table 16 illustrates the difference in Shapiro-Wilk results between the original and transformed datasets.

Table 16

The Differences in Shapiro-Wilk Results Between Original and Transformed Parts 121 and 135 and Part 91 Datasets

		Parts 121 and 135 – cube root				Part 91 – square root transform			
		Original Model		Transformed Model (Tx)		Original Model		Transformed Model (TX)	
DV	IV	Statistic	Sig.	Statistic	Sig.	Statistic	Sig.	Statistic	Sig.
1	1	0.947	0.418	.940	.322	0.553	0.000^	0.600	0.000^
	2	0.794	0.002^	.762	<.001^	0.853	0.012^	0.889	0.044^
	3	0.782	0.001^	.966	.747	0.904	0.079	0.916	0.126
	4	0.938	0.291	.937	.289	0.984	0.985	0.965	0.729
2	1	0.790	0.002^	.701	<.001^	0.897	0.062	0.855	0.013^
	2	0.868	0.020^	.878	.030^	0.809	0.003^	0.744	0.000^
	3	0.839	0.007^	.943	.354	0.862	0.017^	0.868	0.020^
	4	0.988	0.996	.989	.998	0.959	0.617	0.976	0.910
3	1	0.759	0.001^	.686	<.001^	0.810	0.003^	0.882	0.034^
	2	0.884	0.037^	.974	.890	0.864	0.018^	0.778	0.001^
	3	0.959	0.616	.975	.905	0.886	0.040^	0.890	0.046^
	4	0.952	0.483	.973	.875	0.874	0.025^	0.904	0.079
4	1	0.826	0.005^	.764	<.001^	0.627	0.000^	0.717	0.000^
	2	0.845	0.009^	.922	.158	0.826	0.005^	0.748	0.000^
	3	0.823	0.004^	.924	.170	0.979	0.946	0.961	0.649
	4	0.963	0.679	.940	.315	0.917	0.134	0.963	0.681
5	1	0.563	0.000^	.569	<.001^	0.580	0.000^	0.563	0.000^
	2	0	0	0	0	0.262	0.000^	0.262	0.000^
	3	0	0	0	0	0	0		
	4	0	0	0	0	0	0		
6	1	0.929	0.209	.828	.005^	0.561	0.000^	0.608	0.000^
	2	0.914	0.115	.848	.010^	0.910	0.100	0.881	0.033^
	3	0.848	0.010^	.959	.615	0.966	0.743	0.962	0.665
	4	0.806	0.002^	.903	.076	0.959	0.607	0.931	0.227
7	1	0.391	0.000^	.398	<.001^	0.574	0.000^	0.557	0.000^
	2	0	0	0	0	0.262	0.000^	0.262	0.000^
	3	0.955	0.539	.954	.516	0.856	0.013^	0.870	0.022^
	4	0	0	0	0	0.262	0.000^	0.262	0.000^
8	1	0.262	0.000^	.262	<.001^	0.559	0.000^	0.547	0.000^
	2	0.262	0.000^	.262	<.001^	0	0		
	3	0.802	0.002^	.885	.038^	0.965	0.720	0.952	0.485
	4	0.953	0.505	.823	.004^	0.765	0.001^	0.742	0.000^

N = 17

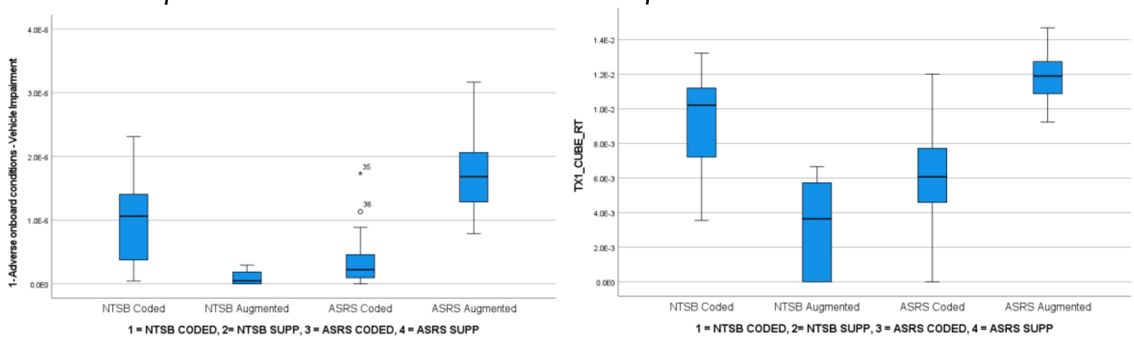
^*p* value < .05 for Shapiro-Wilk tests

IV: 1 = NTSB Classified, 2 = NTSB Augmented, 3 = ASRS Classified, 4 = ASRS Augmented

The outliers' comparison with the six transformation models is documented in Appendix L5 and L6. The cube root transformation for the Parts 121 and 135 dataset provided the optimum performance in reducing mild and extreme outliers to three and six from six to nine. Eight box plots of the cube root transformed DVs are captured in Appendix L7 for Parts 121 and 135. As an illustration for the Parts 121 and 135 dataset, Figure 22 shows that the transformation led to a more normally distributed dataset with reduced extreme outliers for DV1. Although not all the DVs were fully improved, this is acceptable based on the Central Limit Theorem (Field, 2020). The sampling distribution could be expected to be normal because there was a sufficient sample size for this study per the GPower® analysis. According to Hair et al. (2019), the transformed variables are to be retained if the distribution has a higher level of normality than the pre-transformed (p. 115). Therefore, although the Kolmogorov-Smirnov and Shapiro-Wilk results improvement were limited, the cube root transformed variables were deemed acceptable from the normality perspective. For the Part 91 dataset, the improvement made by the transformed model was marginal. After applying square root transformation, extreme outliers were reduced from 31 to 29, and moderate outliers remained at 10. Although the Kolmogorov-Smirnov and Shapiro-Wilk results did not improve, adopting the Square Root transformation improved the distribution of outliers for the Part 91 dataset. An attempt was made to re-run the normality test with four extreme outliers related to the 2004 to 2007 NTSB Classified. The result was improved with the outliers and Kolmogorov-Smirnov and Shapiro-Wilk results, per Appendix L5 and L6. The applicability of this will be discussed later.

Figure 22

Illustration of Parts 121 and 135 Cube Root Transformation Results on DV1



Multivariate normality and outliers were checked by calculating the Mahalanobis D^2 distance. For the Parts 121 and 135 dataset, with the cube root transformation, a maximum distance of 24.31 was recorded. This result demonstrated that the multivariate normality was met as this was below the critical value of 26.13. For the Part 91 dataset, the maximum Mahalanobis D^2 distance was 41.245, lower than the 63.95 with the original dataset. When the filter function was used to filter out Mahalanobis D^2 distance above the critical value of 26.13, only one datapoint was affected, 2004 NTSB Classified, with Mahalanobis D^2 distance of 41.24. Therefore, multivariate normality assumptions were met, with one data point deleted for the Part 91 dataset.

Linearity was checked by scatter plots for the transformed datasets created for Parts 121 and 135, and Part 91 datasets, as shown in Appendices L8 and L10. Apart from Parts 121 and 135 TX5, which were already removed for MANOVA, the best-fit lines represented a higher level of linearity for cube root transformed models for Parts 121 and 135, and square root transformed model for Part 91. Parts 121 and 135 and Part 91 TX5 were therefore removed from the research.

Multicollinearity was tested by Pearson correlation, with results documented in Appendix L9. On examination of the result, for Parts 121 and 135 dataset, with the cube root model, 14 significant relationships with $p < .05$ were found with overall r values decreased. All but five relationships were mildly to moderately correlated with $R < 0.8$ (Hair et al., 2019). TX3/TX2, TX6/TX2, TX4/TX3, TX6/TX3, TX6/TX4 relationships had $p < .05$ and $r > 0.8$, which improved from the original dataset by one set. For the Part 91 dataset, captured in Appendix K.7, 24 significant relationships were found with $p < .05$. Eleven relationships TX4/TX1, TX5/TX1, TX6/TX1, TX8/TX1, TX3/TX2, TX4/TX3, TX5/TX4, TX6/TX4, TX6/TX5, TX8/TX5, TX8/TX6, had $r > 0.80$. The rest of the 13 relationships were moderate, with $R < 0.8$, $N = 67$, and $p < .05$. There was also one relationship more than the Part 91 original dataset.

The last assumption to be tested on the transformed models was homogeneity at univariate and multivariate levels. Box's Test of Covariance Matrices was conducted for the multivariate homogeneity, with results documented in Table 17. The significance value of $p <.001$ indicated some level of multivariant heteroscedasticity for both transformation models.

Table 17

Multivariate Homogeneity Test Results for the Transformed Parts 121 and 135, and Part 91 Datasets

	Cube Root Transformed Parts 121 and 135	Square Root Transformed Part 91
Box's M	122.485	421.903
F	3.332	5.889
df1	28	56
df2	3568.203	6266.924
Sig.	<0.001	<0.001

Although the multivariate Box's M tests did not meet the $p > .001$ significance requirement, the literature review showed that Box's M is sensitive to large data files or uneven group sizes (Tabachnick, Fidell, & Ullman, 2007). If the group sizes were large and even, then the MANOVA would be robust against violations of the homogeneity of variance-covariance matrices assumption (Allen & Bennett, 2008). Although the sample size cannot be described as large, the number of samples in each IV group in this research was even and verified by GPower® to be adequate. Tabachnick et al. (2007) further explained;

> "It should be noted that heteroscedasticity is not fatal to an analysis of ungrouped data. The linear relationship between variables is captured by the analysis, but there is even more predictability if the heteroscedasticity is accounted for. If it is not, the analysis is weakened, but not invalidated" (p. 85).

With the use of Pillai's Trace for the MANOVA (Tabachnick et al., 2007), which is more robust to violation of assumptions and the validation of the quantitative results by the qualitative analysis, the violation of the multivariate homogeneity test was argued to have been mitigated.

Levene's test assessed the assumption of univariate homogeneity of the variance of the transformed DVs. The test used the *median with adjusted dF* criteria by IBM SPSS®. A non-significant result of $p > .05$ indicated that the homogeneity assumption had been fully met. Levene’s test results are documented in Table 18. For the Parts 121 and 135 dataset, the cube root transformed DV1 from significant to insignificant Levene’s test result with $p > .05$. The cube root transformed model further increased the significance level on other DVs. The same applied to the Part 91 dataset whereby the square root transformation increased the number of non-significant Levene tested DV, from zero to one.

Tabachnick et al. (2007) stated that Levene’s test is not typically sensitive to departures from normality. This fact is advantageous to the datasets in this research, as normality was marginal in some cases. Hair et al. (2019) stated that Levene’s homogeneity test results were acceptable even with the presence of univariate heteroscedasticity (i.e., with a significance level of $p < .05$, as experienced in this study). He argued that due to the large sample size in each group and relatively equal sizes across the groups, the presence of homoscedasticity for other groups, further corrective remedies were not needed. As discussed above in Levene’s test results, while the sample size was not large, an equal sample size was achieved, and the sample size was deemed adequate by the GPower® analysis. Also, per Allen and Bennett (2008), if homogeneity of variance cannot be assumed for one (or more) dependent variables, then an alpha level stricter than 0.05 is to be used for performing the Tests of Between-Subjects Effects (univariate ANOVAs). Therefore, an alpha of 0.001 was used to evaluate the univariate (between-subjects effects) result, discussed in the next section.

It was noted that the transformation for Part 91 only had a marginal effect on the normality improvement. The transformation only significantly improved when the outliers on NTSB Classified reports from 2004 to 2007 were removed. Removing four years of data points from the NTSB Classified data would reduce the critical information related to the period prior to the upgrade of the e-ADMS system. Therefore, those *interesting outliers* (Hair et al., 2019) were retained, compensated by a more stringent univariate test threshold of $p < .001$.

Based on these considerations, and that transformation models had been optimized, the cube root transformed DVs were accepted for the transformation of Parts 121 and 135 dataset, and the square root transformation for Part 91 dataset for the MANOVA analysis.

Table 18

Levene's Test Results for Transformed Parts 121 and 135, and Part 91 Datasets

DV	Parts 121 and 135 dataset – Cube Root Transform				Part 91 dataset – Square Root Transform			
	Levene's Statistics	Df1	Df2	Sig	Levene's Statistics	Df1	Df2	Sig
1	2.229	3	50.316	0.096	2.942	3	15.886	.065
2	4.878	3	19.929	0.011*	8.533	3	17.891	<.001*
3	5.411	3	64	0.002*	6.123	3	25.549	.003*
4	4.040	3	64	0.011*	3.586	3	27.863	.026*
5	DELETED							
6	5.092	3	64	0.003*	3.872	3	16.541	.029*
7	5.748	3	64	0.002*	4.958	3	21.922	.009*
8	20.226	3	41.608	<0.001*	2.818	3	15.625	.073*

* $p < .05$

Multivariate Test

The multivariate test assesses whether the means of Belcastro LOC-I Hazard rates differ significantly (i.e., significant main effect) across the different groups of LOC-I safety reports. Pillai's Trace was identified as the most appropriate test for multivariate analysis of variance for a smaller sample size and with some assumptions marginally violated (Allen & Bennett, 2008; Tabachnick et al., 2007); hence had been adopted for this study. The results of the analysis are documented in Table 19.

Results of the MANOVA showed that there was a significant difference among the four groups, NTSB Classified, NTSB Augmented, ASRS Classified, and ASRS Augmented, based on the combined dependent variables. For the Parts 121 and 135 dataset, Pillai's Trace = 2.584, F (21, 180) = 48.345, p <.001, partial η^2 = 0.849, observed power = 1. For the Part 91 dataset, Pillai's Trace = 2.359, F (21, 177) = 31.0, p <.001, partial $\eta^2 = 0.786$, observed power = 1. Based on these results for both datasets, evidence was sufficient to reject the null hypothesis, $H_0 1$, and conclude that the Belcastro LOC-I Hazard rates, when considered together, significantly differed based on the type of safety reports for both Parts 121 and 135 and Part 91 datasets. The effect size was large.

Table 19

Pillai's Trace Test Result for Multivariate Analysis of Variance Based on IV *CODE_TYPE.*

Statistical Test	Dataset	Value	F	Hypothesis df	Error df	Sig.	η2	Observed Power
Pillai's Trace	Parts 121 and 135	2.548	48.345	21.0	180	<0.001	0.849	1.0
	Part 91	2.359	31.0	21.0	177	<0.001	0.786	1.0

Univariate Test

In addition to the multivariate tests, univariate tests for each dependent measure were also performed using IBM SPSS®. This test aimed to examine the differences of the means of Belcastro LOC-I Hazard rates (DVs) across the four types of LOC-I reports separately. The results are documented in Table 20. If all assumptions had been met, a Bonferroni-adjusted alpha of 0.05 / 7, i.e., 0.036, should have been adopted. Due to the violation of the homogeneity assumption, as mentioned above, a stricter alpha of $p < .001$ was applied, per Allen and Bennett (2008). For the Parts 121 and 135 dataset, all DVs, apart from TX5, which was deleted earlier, showed significant results, indicating a significant difference in DVs across CODE_TYPE (type of safety reports). The η_p^2 also indicated a large effect size with a value higher than 0.14 (Field, 2013). The significance level of the results was assessed at < .001, which compensated for the partial violation of Levene's test, as mentioned in the assumptions testing section. Full results are captured in Appendix M.

Results demonstrated sufficient evidence to reject the null hypotheses for all DVs, apart from TX5, which was deleted earlier, for the Parts 121 and 135 dataset. Using TX3, crew action / inaction, as an example, there was a significant difference in TX3 based on the type of safety report, F (3, 64) = 34.427, $p < .001$, $\eta_p^2 = 0.617$, with the hazard rate highest in the NTSB Classified Group (M = 0.047, SD = 0.003) compared to the lowest, NTSB Augmented group, (M=0.013, SD=0.003). As shown in Table 18, the effect size was large for all the ANOVAs with a partial eta square larger than 0.14. The full mean and standard deviation results are captured in Appendix M1.

The situation was different for Part 91. Only TX2, TX3, TX4, and TX7 demonstrated sufficient evidence to reject the null hypotheses, per Table 20. Using TX2, System & Components Failure/Malfunction as an example, there was a significant effect of type of safety report on TX2, F (3, 63) = 20.518, p < .001, η_p2 = 0.494, with the hazard rate highest in the NTSB Augmented Group (M = 0.006, SD = 0.0005) compared to the lowest, ASRS Augmented group, (M=0.001, SD=0.0005). For the four DVs with p <.001, the effect size was large for all the ANOVAs with a partial eta square larger than 0.14. The full mean and standard deviation results are captured in Appendix M4.

Table 20

Univariate Tests Results for Parts 121&135 and Part 91 Datasets

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	$\eta_p 2$	Observed Power[h]
CODE_TYPE Parts 121 and 135	TX1_cubert	.001	3	.000	33.997	<.001	.614	1.000
	TX2_cubert	.008	3	.003	25.214	<.001	.542	1.000
	TX3_cubert	.012	3	.004	34.427	<.001	.617	1.000
	TX4_cubert	.003	3	.001	32.980	<.001	.607	1.000
				DELETED				
	TX6_cubert	.003	3	.001	43.492	<.001	.671	1.000
	TX7_cubert	.002	3	.001	1028.316	<.001	.980	1.000
	TX8_cubert	.001	3	.000	23.869	<.001	.528	1.000
CODE_TYPE Part 91	TX1_cubert	3.317E-6	3	1.1E-6	.911	.441	.042	.239
	TX2_cubert	.000	3	8.7E-5	20.518	<.001	.494	1.000
	TX3_cubert	.000	3	.000	18.946	<.001	.474	1.000
	TX4_cubert	3.888E-5	3	1.30E-5	10.105	<.001	.325	.997
	TX5_cubert			DELETED				
	TX6_cubert	6.891E-5	3	2.30E-5	5.076	.003	.195	.903
	TX7_cubert	3.773E-5	3	1.26E-5	88.571	<.001	.808	1.000
	TX8_cubert	3.354E-6	3	1.12E-6	3.470	.021	.142	.751

h. Computed using alpha = .05

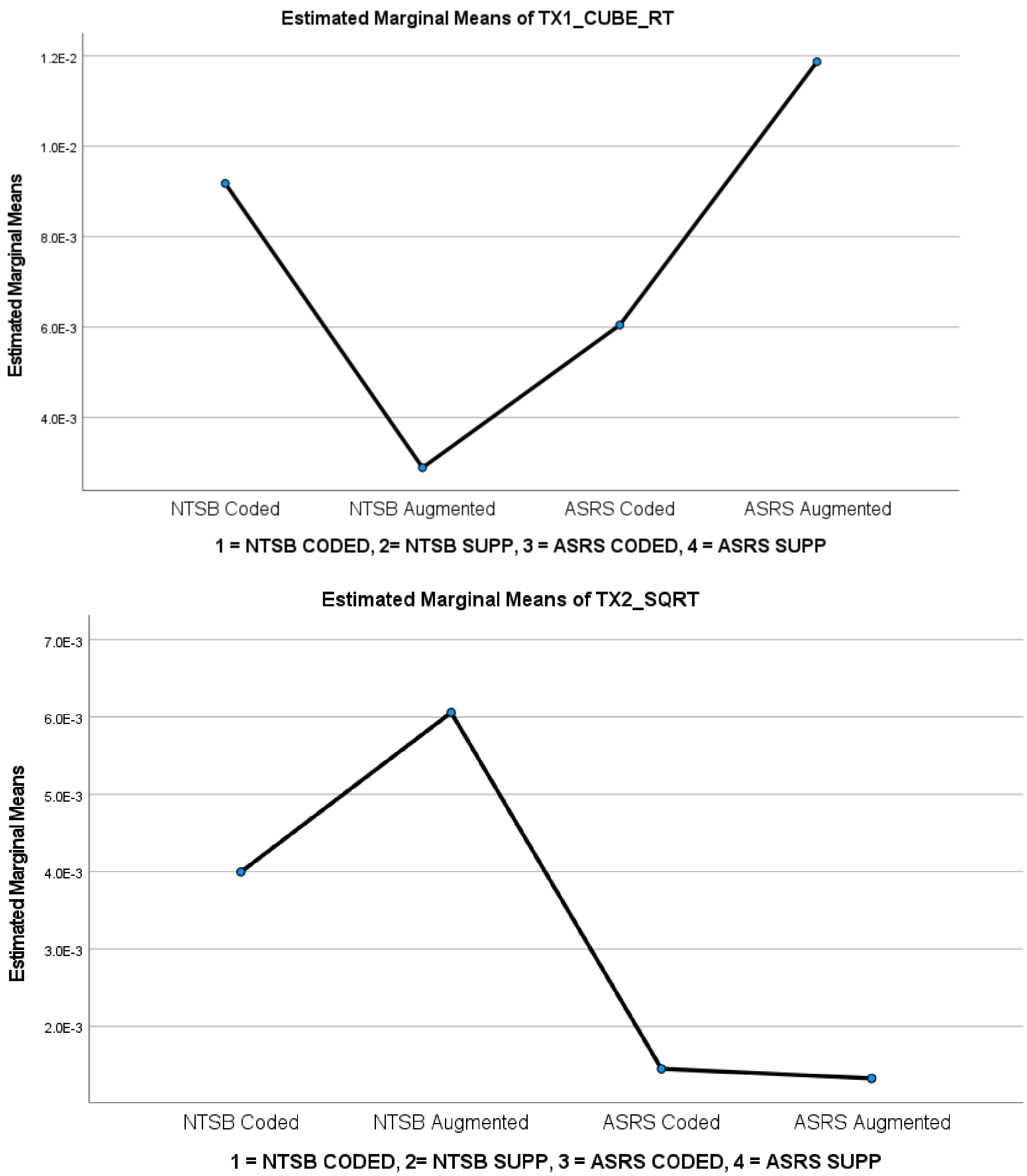
Prefix TX denotes a transformed DV. For example, TX1 denotes a transformed DV1

Post Hoc Test

On examination of the results, although the overall multivariate and univariate main effects of the IV were significant, per Tables 19 and 20, the differences between adjacent groups were not constant. Also, the differences were not all statistically significant. A significant effect by IV CODE_TYPE indicated that the total set of group differences (e.g., ASRS Classified versus ASRS Augmented, NTSB Classified versus NTSB Augmented) was large enough to be considered statistically significant. However, a significant effect did not guarantee that every group difference was significant (Hair et al., 2019). The outstanding question remained regarding individual group differences assessed while maintaining an acceptable level of overall Type I error rate. This was addressed by deploying the post hoc comparison methods based on Tukey HSD being applied to all the seven (as DV5 had been removed from the analysis) DVs across the four groups of IV, the report type, labeled as CODE_TYPE, as Steven et al. (2020) stated that Tukey HSD was most appropriate for pairwise comparison. The full results of this analysis are documented in Appendices M2 and M5 and summarized in Table 21. On examination, the results among Tukey HSD, Scheffe, and LSD were near identical. In addition to examining the statistical data, Hair et al. (2019) recommended the use of the estimate marginal means profile plots in gaining an understanding of the differences between group means; these are documented in Appendix M3 for the Parts 121 and 135 dataset and Appendix M6 for the Part 91 dataset with an illustration captured in Figure 23.

Figure 23

Examples of Estimated Marginal Means of TX1 in Parts 121 and 135 and TX2 in Part 91 Dataset



For example, per Appendix M2, for the Parts 121 and 135 dataset, regarding TX1, Vehicle Impairment, the difference between the means of NTSB Classified and NTSB augmented search was 0.0063. In contrast, the difference between ASRS Classified and ASRS Augmented was 0.0058. Upon inspection of the means scores and the EM plots, for TX1, the ASRS Augmented type has a higher mean vehicle impairment rate than NTSB Classified and ASRS Classified safety reports.

It was thus essential to determine if the differences were significant for all groups or a selection of them. Per the summary in Table 21, using TX1 as an example, all types of safety reports demonstrated pairwise significance in their differences at $p < .05$ for TX1. This was not the case for the rest of the DVs. Therefore, only a portion of the group combinations for the commercial (Parts 121 and 135) and general aviation (Part 91) datasets demonstrated significant differences in the means between groups. For the Part 91 dataset, the majority of the mean differences with TX1, TX5, TX6, and TX8 were not significant, and this was supported by the univariate test results.

Table 21

Post Hoc Comparisons for Individual Group Differences in DVs (TX5 Excluded for Both Datasets) and IV (CODE_TYPE)

Dependent Variable	Independent Variable Groups	NTSB Classified	NTSB Augmented	ASRS Classified	ASRS Augmented
TX1_cubert / sqrt	NTSB Classified		C	C	C
	NTSB Augmented	C		C	C
	ASRS Classified	C	C		C
	ASRS Augmented	C	C	C	
TX2_cubert/ sqrt	NTSB Classified		C / G	C / G	C / G
	NTSB Augmented	C / G		G	G
	ASRS Classified	C / G	G		G
	ASRS Augmented	C / G	G		
TX3_cubert/ sqrt	NTSB Classified		C	C / G	C / G
	NTSB Augmented	C		G	G
	ASRS Classified	C / G	G		
	ASRS Augmented	C / G	G		
TX4_cubert / sqrt	NTSB Classified		C / G	C / G	C / G
	NTSB Augmented	C / G		C	C
	ASRS Classified	C / G	C		
	ASRS Augmented	C / G	C		
TX5_cubert / sqrt	No significant relationship detected				
TX6_cubert / sqrt	NTSB Classified		C	C / G	C / G
	NTSB Augmented	C			C
	ASRS Classified	C / G			
	ASRS Augmented	C / G	C		
TX7_cubert / sqrt	NTSB Classified			C / G	
	NTSB Augmented			C / G	
	ASRS Classified	C / G	C / G		C / G
	ASRS Augmented			C / G	

Dependent Variable	Independent Variable Groups	NTSB Classified	NTSB Augmented	ASRS Classified	ASRS Augmented
TX8_cubert / sqrt	NTSB Classified		G	C	C
	NTSB Augmented	G		C	C
	ASRS Classified	C	C		
	ASRS Augmented	C	C		

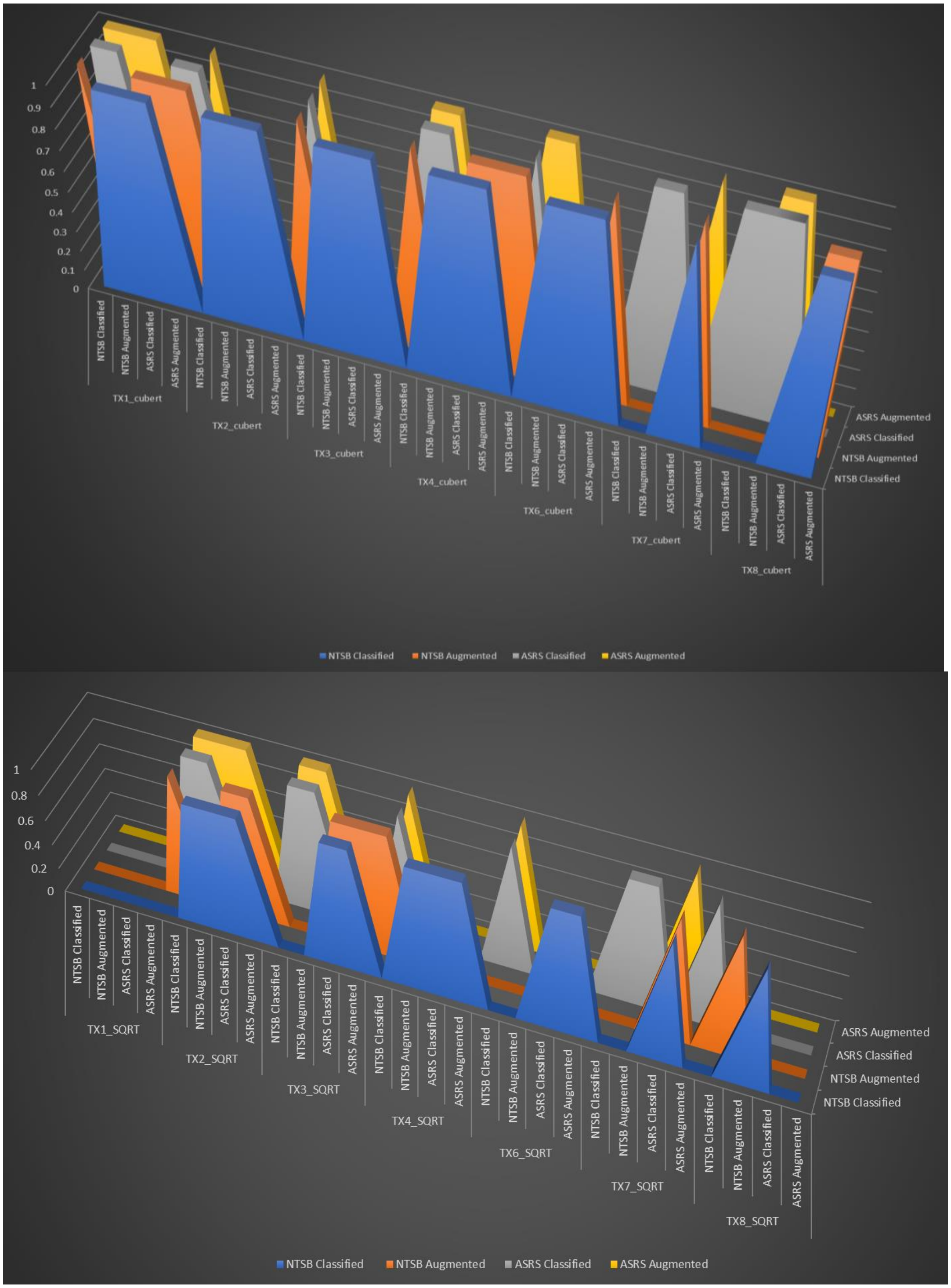
Note. C denotes the Parts 121 and 135 commercial aviation dataset indicated a significant difference at $p < .05$ between groups, G denotes Part 91 general aviation dataset indicated a significant difference at $p < .05$ between groups

Table 21 is further illustrated in an area map using Microsoft Excel® in Figure 24. From the areas map, one on the y-axis referred to a significant difference of $P < .05$ and zero to no significant difference. The areas were grouped in each DV. As seen in Figure 24, it was noted that TX1's high level of significant differences for Parts 121 and 135. The level of significant differences was reduced with other IV groups and DVs. It was also noted that the number of significant differences was less in the Part 91 dataset.

In summary, for the Parts 121 and 135 dataset, 56 out of 96 combinations for commercial aviation groups and 37 out of 96 for general aviation groups demonstrated significant differences in their means. These results supplemented the answer to research question RQ1.

Figure 24

Visualization of the Post Hoc Comparisons for Individual Group Differences in DVs (TX5 Excluded) and IV (CODE_TYPE), Parts 121 and 135 (top) Part 91 (bottom)



Hypothesis

The multivariate analysis in MANOVA identified significant differences in the means rates of reported hazards between NTSB and ASRS reports, and between classified and augmented reports, at multivariate and univariate levels. Response to each hypothesis at the multivariate and univariate level is below:

$H_A 1$

The group mean vectors in Belcastro LOC-I Hazard rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

Based on the multivariate MANOVA results, this null hypothesis was rejected with Pillai's Trace equals 2.584, F (21, 180) = 48.345, p <.001, partial $\eta^2 = 0.849$, observed power = 1 for the Parts 121 and 135 dataset and Pillai's Trace = 2.359, F (21, 177) = 31, p <.001, partial $\eta^2 = 0.786$, for the Part 91 dataset. Hence, the null hypothesis was rejected for Parts 121 and 135, and Part 91 datasets. Significant differences in the group mean vectors in Belcastro LOC-I Hazard rates across the four types of safety reports in commercial and general aviation between 2004 and 2020.

$H_A 2$

The means of adverse onboard conditions - vehicle impairment rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

Using a stricter alpha level of p < .001, results demonstrated sufficient evidence to reject the Parts 121 and 135 null hypothesis, F (3, 64) = 33.997, $p < 0.001$, $\eta_p^2 = 0.614$. However, the Part 91 null hypothesis was retained as the significance level was 0.441. For the Parts 121 and 135 ANOVA, the effect size was large. Further examination of the descriptive statistics in Appendix M2 showed ASRS Augmented reports (*M* = 0.012 and *SD* = 0.001) had the highest adverse onboard conditions - vehicle impairment rate. In contrast, NTSB Augmented reports (*M* =0.003 and *SD* = 0.001) had the lowest for the Parts 121 and 135 dataset. Therefore, the means of adverse onboard conditions - vehicle impairment rates were significantly different across the four types of safety reports in commercial aviation but not general aviation.

H_A3

The means of adverse onboard conditions - system and components failure / malfunction rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

Using a stricter alpha level of p <.001, results demonstrated sufficient evidence to reject the Parts 121 and 135 null hypothesis, F (3, 64) = 25.214, p <.001, η_p^2 = 0.542. The null hypothesis for the Part 91 dataset could also be rejected, F (3, 63) = 20.518, p <0.001, η_p2 = 0.494. Both datasets displayed a large effect size. Further examination of the descriptive statistics in Appendix M2, which showed the Parts 121 and 135 dataset, NTSB Classified reports (*M* = 0.036 and *SD* = 0.002) had the highest adverse onboard conditions - system and components failure / malfunction rate. In contrast, ASRS Classified reports (*M* =0.008 and *SD* = 0.002) had the lowest for the Parts 121 and 135 dataset. For the Part 91 dataset, as indicated in Appendix M4, NTSB Augmented was the highest (*M* =6E-03 and *SD* = 4.99E-04), with both ASRS Classified and ASRS Augmented the lowest (*M* =0.001 and *SD* = 0.0005). Therefore, the means of adverse onboard conditions - system and components failure / malfunction rates were different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A4

The means of adverse onboard conditions - crew action / inaction rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

Using a stricter alpha level of p <.001, results demonstrated sufficient evidence to reject the Parts 121 and 135 null hypothesis, F (3, 64) = 34.427, p <.001, η_p^2 = 0.617. The null hypothesis for the Part 91 dataset could also be rejected, F (3, 63) = 18.946, p <0.001, $\eta_p 2$ = 0.474. For the Parts 121 and 135 ANOVA, the effect size was large. Further examination of the descriptive statistics in Appendix M2 showed NTSB Classified reports (*M* = 0.046 and *SD* = 0.003) had the highest adverse onboard conditions - crew action / inaction rate. At the same time, NTSB Augmented reports (*M* =0.013 and *SD* = 0.003) had the lowest for the Parts 121 and 135 dataset. For the Part 91 dataset, as indicated in Appendix M4, NTSB Classified was the highest (*M* =0.008 and *SD* = 0.00067), with both ASRS Classified and ASRS Augmented the lowest (*M* =0.003 and *SD* = 0.00065). In sum, the means of adverse onboard conditions - crew action / inaction rates- differed across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A5

The means of external hazards and disturbances - inclement weather atmospheric disturbances rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

Using a stricter alpha level of p <.001, results demonstrated sufficient evidence to reject the Parts 121 and 135 null hypothesis, F (3, 64) = 32.980, p <0.001, η_p^2 = 0.607. The null hypothesis for the Part 91 dataset could also be rejected, F (3, 63) = 10.105, p <0.001, $\eta_p 2$ = 0.325. Both datasets displayed a large effect size. Further examination of the descriptive statistics in Appendix M2 showed that for the Parts 121 and 135 dataset, NTSB Classified reports (*M* = 0.023 and *SD* = 0.001) had the highest external hazards and disturbances - inclement weather atmospheric disturbances rate. In comparison, NTSB Augmented (*M* =0.006 and *SD* = 0.001) had the lowest for the Parts 121 and 135 dataset. For the Part 91 dataset, per Appendix M4, NTSB Classified was the highest (*M* =0.003 and *SD* = 0.00028), and ASRS Augmented was the lowest (*M* =0.001 and *SD* =0.00027). In sum, the means of external hazards and disturbances - inclement weather atmospheric disturbances rates were different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

$H_A 6$

The means of external hazards and disturbances - poor visibility rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

Due to the lack of statistical significance distribution in both datasets for this DV with hazard rates equal to zero, the null hypothesis could not be rejected for both Parts 121 and 135 and Part 91 datasets. In sum, the means of external hazards and disturbances - poor visibility rates were not different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A7

The means of external hazards and disturbances - obstacle rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

Using a stricter alpha level of p<.001, results demonstrated sufficient evidence to reject the Parts 121 and 135 null hypothesis, F (3, 64) = 43.492, p <0.001, η_p^2 = 0.671. However, the Part 91 null hypothesis could not be rejected as the p level was 0.003, higher than 0.001, per Table 20. For the Parts 121 and 135 ANOVA, the effect size was large. Further examination of the descriptive statistics in Appendix M2 showed NTSB Classified reports (M = 0.023 and SD = 0.001) had the highest external hazards and disturbances - obstacle rate. At the same time, ASRS Augmented reports (M =0.010 and SD = 0.001) had the lowest for the Parts 121 and 135 dataset. In summary, the means of external hazards and disturbances - obstacle rates were different across the four types of safety reports in commercial and not different in general aviation between 2004 and 2020.

H_A8

The means of abnormal vehicle dynamics and upsets - abnormal vehicle dynamics rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

Using a stricter alpha level of p<.001, results demonstrated sufficient evidence to reject the Parts 121 and 135 null hypothesis, F (3, 64) = 1028.316, p <.001, η_p^2 = 0.980. The null hypothesis for the Part 91 dataset could also be rejected, F (3, 63) = 88.571, p <0.001, η_p2 = 0.808. Both datasets displayed a large effect size. Further examination of the descriptive statistics in Appendix M2 showed that for the Parts 121 and 135 dataset, ASRS Classified reports (*M* = 0.013 and *SD* = 0.0002) had the highest abnormal vehicle dynamics and upsets - abnormal vehicle dynamics rate. In comparison, NTSB Classified and Augmented reports (*M* =0 and *SD* = 0) had the lowest for the Parts 121 and 135 dataset. For the Part 91 dataset, per Appendix M2, ASRS Classified was the highest (*M* =0.002 and *SD* = 0.00009), and ASRS Augmented was the lowest (*M* =0.00001 and *SD* =0.00009). In sum, the means of abnormal vehicle dynamics and upsets - abnormal vehicle dynamics rates differed across the four types of safety reports in commercial and general aviation between 2004 and 2020.

H_A9

The means of abnormal vehicle dynamics and upsets - vehicle upset conditions rates are different across the four types of safety reports in commercial and general aviation between 2004 and 2020.

Using a stricter alpha level of p <.001, results demonstrated sufficient evidence to reject the Parts 121 and 135 null hypothesis, F (3, 64) = 23.869, p <0.001, η_p^2 = 0.528. However, the Part 91 null hypothesis could not be rejected as the significance level was 0.021, higher than 0.001. For the Parts 121 and 135 ANOVA, the effect size was large. Further examination of the descriptive statistics in Appendix M2 showed that ASRS Classified (*M* = 0.0059 and *SD* = 0.0007) had the highest abnormal vehicle dynamics and upsets - vehicle upset conditions rate. In contrast, NTSB Classified and Augmented reports (*M* =0.0002 and *SD* = 0.0007) had the lowest for the Parts 121 and 135 dataset. In sum, the means of abnormal vehicle dynamics and upsets - vehicle upset conditions rates were different across the four types of safety reports in commercial and not different in general aviation between 2004 and 2020.

Discriminant Analysis to Verify the MANOVA Univariate Result

Discriminant analysis was used to verify the univariate analysis of the MANOVA. The analysis assessed individual outcome variables (DVs) regarding their differences across the treatment variables (IV). The objective was to profile the outcome variables in terms of their differences between groups of treatment variables. This analysis was useful when the treatment variable has three or more levels, as in this study (Field, 2013; Hair et al., 2019). The IV and DVs were reversed between MANOVA and discriminant analysis.

Assumptions Testing

Before starting the discriminant analysis, normality, linearity, and multicollinearity assumptions were explored, as specified in Hair et al. (2019). The three assumptions were already considered in the MANOVA analysis. Although the assumptions were not completely met, the cube root transformed dataset for Parts 121 and 135 and the square root transformed dataset for Part 91 were used as they produced the optimized level of adherence. Regarding homogeneity, the Box's M test results for both Parts 121 and 135, and Part 91 datasets were identical to the Box's M performed during MANOVA, as shown in Table 17 with $p > .001$. Hair et al. (2019) indicated that for discriminant analysis, the sensitivity of the test to factors other than just covariance differences (e.g., normality and sample sizes) made this an acceptable level. Therefore, it was argued that the datasets used in MANOVA were also applicable to the discriminant analysis in terms of the assumptions.

With the assumptions optimized, discriminant analysis using Wilk's Lambda, pooled within-groups matrices, tests of equality of group means, eigenvalues, standardized canonical discriminant function coefficients, structure matrix, and classification results on IBM SPSS® were carried out. The key results for the Parts 121 and 135 and Part 91 datasets are documented in Tables 22 to 25, with supplementary results in Appendix N for Parts 121 and 135 and Appendix O for Part 91 datasets. In terms of both the Parts 121 and 135 and Part 91 datasets, TX5 was the variable that induced the error message on two nonsingular group covariance matrixes, requiring removal. With TX5 removed, the analysis was a rerun.

Wilks's Lambda Tests

For the Parts 121 and 135 dataset, per Table 22, three discriminant functions were found to be statistically significant: Wilks's Λ = .012, (21) = 444, $p < .001$ for discriminant function 1 through 3; Wilks's Λ = 0.047, (12) = 9.26, $p < .001$ for discriminant function 2 through 3, Wilks's Λ = 0.230, (5) = 90.33, $p < .001$ for discriminant function 3. For the Part 91 dataset, per Table 23, three discriminant functions were found to be statistically significant: Wilks's Λ = .004, (21) = 343, $p < .001$ for discriminant function 1 through 3; Wilks's Λ = 0.084, (12) = 152, $p < .001$ for discriminant function 2 through 3, Wilks's Λ = 0.344, (5) = 66, $p < .001$ for discriminant function 3. These meant function 3, combined 2 and 3, and combined 1 and 3 were effective in discriminating among the four types of safety reports.

Table 22

Discriminant Analysis Wilk's Lambda Results for Parts 121 and 135, and Part 91 Datasets.

Dataset	Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
Parts 121 and 135	1 through 3	0.001	441.012	21	< .001
	2 through 3	0.047	187.696	12	< .001
	3	0.230	90.330	5	< .001
Part 91	1 through 3	.004	342.911	21	< .001
	2 through 3	.084	151.996	12	< .001
	3	.344	65.668	5	< .001

Equality of Group Means and Eigenvalue Tests

The tests of equality of group means in Appendix N1 examined whether mean differences exist across groups for any variables. This showed that all three functions discriminated the four groups of LOC-I safety report types. Having applied Bonferroni Adjustment ($p < .05 / 7 = 0.007$), significant differences across the groups with TX1, 2, 3, 4, 6, 7, and 8 were obtained for the Parts 121 and 135 dataset, supporting the univariate results in the MANOVA. For the Part 91 dataset, only TX2, 3, 4, 6, and 7 indicated significant differences. These were similar to the MANOVA result with a difference of TX6, which did not previously pass the univariate test in MANOVA.

For the Parts 121 and 135 dataset, by examining the eigenvalues indicated in Table 21, the first discriminant function explains 89.3% of the variance, the second discriminant function explains 5.7% of the variance, and the third discriminant function explains the rest of the variance. From Table 23, Canonical correlations are 0.992, 0.891, and 0.877 for the three discriminant functions, indicating that 99%, 89%, and 88% of variances were explained by the relationship between predictors and group membership by discriminant functions 1, 2, and 3 respectively. The canonical correlation value was also the square root of the effect size, η_p^2 (Hair et al., 2019). Therefore, the effect size was over 0.75 for all three functions.

For the Part 91 dataset, by examining the eigenvalues indicated in Table 23, the first discriminant function explains 81% of the variance, the second discriminant function explains 11.7% of the variance, and the third discriminant function explains the rest of the variance. From Table 21, Canonical correlations are 0.977, 0.869, and 0.810 for the three discriminant functions, indicating that 97.7%, 86.9%, and 81.0% of variances were explained by the relationship between predictors and group membership by discriminant functions 1, 2 and 3 respectively. Per the above, the effect size was over 0.65 for all three functions.

Table 23

Discriminant Analysis Eigenvalues for Parts 121&135 and Part 91 Datasets

	Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
Parts 121 and 135	1	60.495[a]	89.3	89.3	.992
	2	3.870[a]	5.7	95.1	.891
	3	3.344[a]	4.9	100.0	.877
Part 91	1	21.294[a]	81.0	81.0	.977
	2	3.070[a]	11.7	92.7	.869
	3	1.909[a]	7.3	100.0	.810

[a] The first three canonical discriminant functions were used in the analysis.

Standardized Canonical Discriminant Function Coefficient Tests

The Standardized Canonical Discriminant Function Coefficients in Appendix N3 showed that in terms of the Parts 121 and 135 dataset, for function 1, TX7 was substantially contributing with a value greater than 0.5; for function two, TX1 was substantially contributing; for function three, TX6, 3, 4, 2, and 8 were contributing. The structure matrix in Appendix N4 examined the extent to which each variable was correlated to the overall function. For function one, TX7 had the strongest correlation to the function. For function two, TX1 had the strongest correlation, while TX6, TX3, and TX4 had the strongest correlation for function three, per Appendix N4. The Classification Results from Appendix N5 indicated that 94.1% of the original grouped cases were correctly classified.

For the Part 91 dataset, for function one, all DVs seemed to be contributing with a standardized coefficient greater than 0.5; for function two, TX2, 3, 4, and 6 were top contributors; and for function three, TX 2, 4, and 6 were the top contributors, as TX1 and TX8 did not pass the equality of group means test earlier. On examination of the structure matrix in Appendix O4: for function one, TX7 was most correlated with the function TX8, and TX1 for function two, and TX2, 3, 4, and 6 for function three. However, it was observed that the levels of correlation were generally lower than the Parts 121 and 135 structure matrix. The highest correlation was 0.583 for Part 91 compared with 0.889 for Parts 121 and 135 in function one. The Classification Results from Appendix O5 indicated that 97.1% of the original grouped cases were correctly classified. A summary of the predicted membership results for both datasets is detailed in Table 24, and the discriminant analysis structure loadings on Function Results are detailed in Table 25.

Table 24

Percentage of Validated Predicted Membership Results from the Discriminant Analysis for Both Datasets.

Part(s)	Group	Predicted Group Membership NTSB Coded	NTSB Augmented	ASRS Coded	ASRS Augmented	Total
91	NTSB Coded	94.1	5.9	.0	.0	100.0
	NTSB Augmented	.0	94.1	.0	5.9	100.0
	ASRS Coded	.0	.0	100	.0	100.0
	ASRS Augmented	.0	.0	.0	100	100.0
121 and 135	NTSB Coded	76.5	23.5	.0	.0	100.0
	NTSB Augmented	5.9	94.1	.0	.0	100.0
	ASRS Coded	.0	.0	100.0	.0	100.0
	ASRS Augmented	.0	.0	.0	100.0	100.0

Table 25

Discriminant Analysis Structure Loadings on Function Results for Parts 121 and 135, and Part 91 Datasets

	Structure Loadings on Functions					
	Parts 121 and 135 Functions			Part 91 Functions		
	1	2	3	1	2	3
TX7_cubert	.889 [a]	-.164	.284	.392 [a]	-.273	.073
TX1_cubert	-.033	.625 [a]	.068	-.022	-.142 [a]	.071
TX6_cubert	-.029	.408	-.634 [a]	-.049	-.189	.261 [a]
TX3_cubert	-.035	.365	-.553 [a]	-.112	-.120	.575 [a]
TX4_cubert	.005	.405	-.521 [a]	-.036	-.266	.387 [a]
TX2_cubert	-.060	.256	-.461 [a]	-.100	.191	.583 [a]
TX8_cubert	.075	.259	.395 [a]	.025	-.234 [a]	.069

Note. Correlations between variables and standardized conical discriminant functions, variables were ordered by the absolute size of correlation within a function based on Parts 121 and 135 dataset

[a] Largest absolute correlation between each variable and any discriminant function

In summary, for the Parts 121 and 135 dataset, with Bonferroni correction, TX1, 2, 3, 4, 6, 7, and 8 were variables that demonstrated significant differences between groups. Three significant functions that described group differences were found with high effect sizes, and 94.1% of the original grouped cases were correctly classified. When examining standardized coefficients, all DVs contributed to the respective discriminant functions. This supported the univariate post hoc results in the MANOVA. For the Part 91 dataset, with Bonferroni correction, TX 2, 3, 4, 6, and 7 were variables that demonstrated significant differences between groups. Three significant functions that described group differences were found with high effect sizes, and 97.1% of the original grouped cases were correctly classified. When examining standardized coefficients, all contributed to the respective discriminant functions. The structure matrix also reflected the strongest correlation to functions as the standardized coefficients results, though with a lower level of correlation with the discriminant functions compared with the Parts 121 and 135 dataset. This broadly supported the univariate post hoc results in MANOVA with the difference of TX6, which did not pass the univariate test in the MANOVA while passing the equality of group means test in the discriminant analysis.

Qualitative Analysis – A Supplement

NVivo® was used to explore LOC-I reports from their synopsis and narratives in the ASRS, NTSB, and AIDS databases in a cursory manner to seek any insights relevant to the MANOVA results. AIDS contained events matching the definitions of incidents (ICAO, 2001). Hence, it was introduced as a source of reports with severity between ASRS and NTSB.

Word Clouds, Tree Maps, and Cluster Analyses on NTSB and ASRS Data

Word clouds, tree maps, cluster analyses, and word trees based on Belcastro et al.'s (2018) keywords for LOC-I were deployed. The word clouds are captured in Figure 25, while the rest of the results are captured in Appendix N. The analyses were conducted using the stemmed words setting on NVivo®. The source summary of the narratives extracted is listed in Appendix P1. The top ten frequent word comparison from the Parts 121 and 135 dataset treemaps is shown in Table 26.

Table 26

Top 10 Frequent Word Comparison for Parts 121 and 135 Dataset from Tree Maps

NTSB Classified	NTSB Augmented	ASRS Classified	ASRS Augmented
Aircraft*	Flights*	Aircraft*	Aircraft*
Pilot	Airplanes*	Turbulent	Flights*
Flights*	Engine	Flights*	Engines
Accident	Pilot	Controls	Lands
Control	Landing	Encountered	Crews
Runway	Gear	ATC	Stalls
Reported	Operators	Reports	Timing
Engine	Airport	Severity	Approaching
Operators	Left	Turns	Calls
Airport	Reported	Timing	First

Note. * indicates the same text appeared in all four groups

The hypotheses related to RQ1 and RQ2 were used in guiding the interpretation of the NVivo® study results in Table 27, as follows:

Table 27

Summary of Qualitative Analysis from NTSB, ASRS Narratives

Criteria being tested for commercial and general aviation	Insights from NVivo® study results	Guidance to alternative hypothesis
H_A1 - Linear combinations of Belcastro LOC-I Hazard rates	Examination of word clouds and tree maps in Appendix N demonstrated that each dataset and type of report shared some similarities of the highest frequency words, such as *aircraft* and *flights*. However, the order of higher-frequency words did differ across the groups. For example, *crew* or *pilots* were mentioned as the top items for the NTSB Parts 91 database, while *runways* and *landings* were the top items for ASRS. The Parts 121 and 135 dataset results displayed the same level of differences.	Supported for both general and commercial aviation.
H_A2 - Adverse onboard conditions - Vehicle Impairment	For the Parts 121 and 135 dataset, words that resembled impairment, such as *controls, engines, and stalls*, appeared in NTSB, ASRS, and AIDS groups as the highest frequency words, indicating this attribute was measured in the dataset. Examining the tree maps of the Part 91 dataset, they did not show explicit mentions of aircraft impairment-related words. Hence it was not conclusive if such hazards differed in distribution. This supported the MANOVA and discriminant univariate finding.	Supported for commercial aviation. Not supported for general aviation.
H_A3 - Adverse onboard conditions - System and components failure / malfunction	*Fuel, instrument, autopilots,* and *indicator* appeared in the top 100-word frequency treemaps for Parts 121 and 135 Classified. The appearance of such words in the rest of the groups was less pronounced. This aligned with Appendix L-3 TX2 Estimated Marginal Means graph. For the Parts 91 dataset, the NTSB Augmented group demonstrated system and components failure-related words such as *engines, fuel,* and *power* among the first eight highest frequency words. This was more apparent than other groups and corresponded with Appendix L.6 TX2 Estimated Marginal Means graph. This supported the MANOVA and discriminant univariate finding.	Supported for both general and commercial aviation.
H_A4 - Adverse onboard conditions - Crew action / inaction	Among other groups, the word pilot was the second highest frequency in the Classified Part 121 & 135 dataset. This aligned with Appendix L-3 Estimated Marginal Means plot. However, no mention of this word was found in the ASRS groups. For the Part 91 dataset, both NSTB groups had *pilots* as the second high frequency. ASRS Classified group had *pilot* featured as the fifth highest word with no mention in the top eight highest frequency words in the Augmented group. The distribution broadly matched with L.6 TX3 Estimated Marginal Means plot.	Supported for both general and commercial aviation.

Criteria being tested for commercial and general aviation	Insights from NVivo® study results	Guidance to alternative hypothesis
H_A5 - External hazards and disturbances - Inclement weather atmospheric disturbances	For Parts 121 and 135, *turbulent, encountered, winds,* and *wake* featured among the top 36 frequent words in the ASRS Classified group. Whereby no equivalent mentions could be found in the Augmented dataset. For NTSB, *weather* and *ice* were featured in the Classified group, whereby the word *meteorology* was only ranked 80th in the tree map. The rankings were broadly aligned with Appendix L-3 TX4 Estimated Marginal Means plot. The results were less conclusive for Part 91, whereby all four groups featured words in the top frequency counts that matched the criteria. For example, NTSB Part 91 featured *meteorology* as the 37th top word for the Classified group, *conditions* featured as 33rd ranked in the Augmented group, while for ASRS, *winds* and *turbulent* were 14th and 25th in the Classified group, and *winds* and *conditions* featured as 51st and 96th. As the estimated marginal plots were based on the normalized rates data, the ranking in the tree maps did not provide much useful information in this case. Therefore, the cluster analyses were examined and showed the differences in the clusters for each group regarding inclement weather, rejecting the hypothesis.	Supported for both general and commercial aviation.
H_A6 - External hazards and disturbances - Poor visibility	Having examined the treemaps for the Part 91 dataset, no direct word meaning poor visibility was found in the top 100 frequent words. For the NTSB dataset, the word *visual* featured in the Classified and Augmented, but there was no indication of whether this linked to poor visibility. The qualitative data was inconclusive. This aligned with the MANOVA findings leading to the removal of the variable.	Not supported for both general and commercial aviation.
H_A7 - External hazards and disturbances - Obstacle rate	For Parts 121 and 135, tree maps in Appendix N showed that the Classified dataset contained the word *impact* as its top 21st highest frequency word, with no related word found in the Augmented group. The ASRS groups showed no related words in the top 100. This result was in broad alignment with Appendix L-3 for TX6, whereby NTSB Classified had the highest estimated marginal means value. For the Part 91 dataset, both NTSB Classified and Augmented sets had the word *impacted* in the 18th and 19th ranks, with the word *damage* found within the top 100 rankings. By comparing the ranking order, there did not seem to be a significant difference among the groups.	Supported for commercial aviation, not rejected for general aviation.
H_A8 - Abnormal vehicle dynamics and upsets - Abnormal vehicle dynamics	Most groups in the Parts 121 and 135 dataset displayed some top 100 frequency words related to flight dynamics, such as *turn, airspeed,* and *rolls*, but the NTSB Augmented group contained none of these words in the top 100. This supported the Appendix L-3 TX7 Estimated Marginal Means plot indicating the lowest	Supported for both general and commercial aviation.

Criteria being tested for commercial and general aviation	Insights from NVivo® study results	Guidance to alternative hypothesis
	marginal means for NTSN Augmented. For the Part 91 dataset, all groups displayed some related keywords such as *rolls*, *turns*, *airspeed*, *pitching*, and *speeds.* Three keywords had been detected in the ASRS Classified group instead of the one to two for the rest of the groups. This aligned with Appendix L.6 Estimated Marginal Means plot for TX7.	
H_A9- Abnormal vehicle dynamics and upsets - Vehicle upset conditions	For the Parts 121 and 135 dataset, NTSB Classified and ASRS Augmented each had *upset* as the top 100. This word was ranked 7th in the ASRS Augmented group and 64th in NTSB Classified. This supported the ASRS Augmented as the peak in the Appendix L-3 TX8 Marginal Means Plot. However, the ASRS Classified dataset did not feature words directly connected to an upset condition suggesting the incomplete nature of the narratives. For the Part 91 dataset, no keywords directly related to the upset conditions were found. This supported the finding in the MANOVA.	Not supported for both general and *partially supported* commercial aviation.

Insights from AIDS Data

Reviewing the Word Clouds in Figure 25 and the tree maps in Appendix N suggested that, for the Part 91 dataset, AIDS' narratives provided similar coverage of the keywords compared with the NTSB and ASRS datasets. However, the volume of the data from AIDS was lower than in NTSB and ASRS groups, as indicated in the sizes from the combined hierarchy charts in Appendices P14 to 17. On closer examination, the AIDS dataset contained more mentions of factors such as engine, omitted rather than the actual consequences. Also, the prominence of the words *reported*, *causing*, and *resulting* suggested a third-person approach in the reports rather than written in first-person in the case of ASRS. The NTSB word clouds also carried this similarity, signified by the frequent words revealed. For the Parts 121 and 135 dataset, one main difference between AIDS and the rest of the groups was that the word *nose* featured centrally in the top frequent words in the Classified group, with the word *stall* in the Augmented group. This provided more information on the flight dynamics and upset conditions, DV7 and DV8. There was also less mention of *pilots* in the AIDS reports and high-frequency words suggesting that the crew was more of a focus for AIDS reports.

Figure 25

Word Clouds from the Classified and Augmented Searched LOC-I Reports Synopsis and Narratives from AIDS, ASRS and NTSB Databases

	Coded	Augmented
AIDS Part 91		
Parts 121&135		
ASRS Part 91		
Parts 121&135		

NTSB Part 91		
Parts 121&135		

Summary

Analyses have been performed to examine the coded and augmented quantitively searched LOC-I events. Cursory analysis using the qualitative technique was also conducted to provide insights into the quantitative results. By deploying four subject matter experts, the quantitative analyses have been made possible by establishing a common taxonomy between ASRS and NTSB databases. The SMEs mapped the coding taxonomies between ASRS and NTSB using Belcastro's factors of LOC-I (Belcastro et al., 2018). The results have highlighted the rejection of the multivariate null hypotheses related to RQ1, and $H_0 1$, meaning the mean hazard rates are collectively different across safety report types. Some combinations, but not all univariate null hypotheses (H_0 2 - 9), were also rejected. The results refer to the similarities of means of hazard rates between NTSB and ASRS databases for LOC-I for one Parts 121 and 135 DV and for three Part 91 DVs. Discriminant analysis was carried out and validated the univariate MANOVA results. The differences highlighted in the quantitative analyses were further substantiated in the cursory qualitative analysis using the safety reports narratives.

Chapter V: Discussion, Conclusions, and Recommendations

This study evaluated the levels of differences in eight Belcastro LOC-I Hazards (DVs) across four severity groups (IV) of LOC-I safety reports from ASRS and NTSB databases. The reports evaluated were obtained from two search methods: classified and augmented. MANOVA and discriminant analyses were deployed in the core quantitative analyses. Cursory qualitative analysis based on report narratives was used to provide additional insights. This chapter discusses the study's results, its contributions in theoretical and practical manners, and its broader implications for the effectiveness of safety reporting in aviation and safety industries where open-loop voluntary safety reporting systems (such as ASRS) are implemented.

Discussions of Results

The results of this study, as detailed in Chapter IV, have been critically examined with respect to the ground theories documented in Chapter II. From this critical review, apart from answering the research questions and their associated hypotheses specified in Chapter I, additional findings have been made to provide more insights into the relationship between safety report types and Belcastro LOC-I Hazards. These additional findings were anticipated to contribute to the knowledge base on aviation safety reporting systems.

Research Question 1

RQ1 is a multivariate research question, "Do Belcastro LOC-I Hazard rates differ across types of safety reports for commercial and general aviation?" MANOVA results on H_A1 showed that for both 121 and 135, and Part 91 datasets, when all the Belcastro LOC-I Hazards (DVs) were considered together in a multivariate manner, the means of the hazards rate vectors across the four groups of safety reports (IV) were significantly different for commercial and general aviation. In other words, Belcastro LOC-I Hazards rates collectively differ across types of safety reports for commercial and general aviation. Chapter II discusses the differences in the severity of the cases, the level of rigor and independence in the investigation, biases from the originators, and differences in the extent of follow-up for individual safety reporting systems (Mills, 2011). The likelihood is that one or a combination of such differences transpired to the differences in the content of the safety reports across types. Secondly, the differences in temporal sequence in the reporting types may lead to differences in reported hazards. NTSB reports contain accidents, typically covering the entire accident causation chain (Reason, 2016), while ASRS reports contain safety events that may exhibit only part of the causation chain of a LOC-I accident.

Research Question 2

RQ2 is a univariate research question, "Which of the Belcastro LOC-I Hazards display(s) significant difference(s) in mean hazard rate(s) across safety report types for commercial and general aviation?" Table 20 documented the univariate MANOVA results for commercial and general aviation with a strict $p < .001$ to compensate for the partial conformance with assumptions such as homogeneity. For commercial aviation, the DVs that displayed such differences in mean hazard rates across groups were:

a. adverse onboard conditions - vehicle impairment

b. adverse onboard conditions - system and components failure / malfunction

c. adverse onboard conditions - crew action / inaction

d. external hazards and disturbances - inclement weather atmospheric disturbances

e. external hazards and disturbances – obstacle

f. abnormal vehicle dynamics and upsets - abnormal vehicle dynamics, and

g. abnormal vehicle dynamics and upsets - vehicle upset

For general aviation, the DVs below displayed the differences:

i. adverse onboard conditions - system and components failure / malfunction

ii. adverse onboard conditions - crew action / inaction,

iii. external hazards and disturbances - inclement weather atmospheric disturbances, and

iv. abnormal vehicle dynamics and upsets - abnormal vehicle dynamics displayed differences.

Contrastingly, the research found a collection of Belcastro LOC-I Hazards that did not statistically differ across the four severity groups of safety reports. For both commercial and general aviation, external hazards and disturbances - poor visibility was a DV that did not demonstrate differences across the four groups of safety reports. For general aviation, the following DVs did not demonstrate significant differences across the groups:

i. adverse onboard conditions - vehicle impairment

ii. external hazards and disturbances – obstacle

iii. abnormal vehicle dynamics and upsets - vehicle upset

The lack of differences for some Belcastro LOC-I Hazards was as impactful, if not more so, than identifying differences because this highlighted a higher value of the safety report types in the lower severity groups (ASRS). A detailed discussion of this impact for each DV is documented below:

Univariate Analysis on Adverse Onboard Conditions - Vehicle Impairment. The Part 91 dataset did not pass the univariate test. This result signified that, for the Part 91 dataset, the vehicle impairment hazard rates across each group were not significantly different. Therefore, should a data analysis exercise be conducted on the four groups of safety reports in general aviation, based on this result, the vehicle impairment data rates would not be significantly different. Provided the context of the vehicle impairment data was similar across each safety report group, addressed later in this chapter, this result could provide a pathway to mitigate the causal factor of vehicle impairment for Part 91 by ASRS reports. On reflection, the Part 91 operation utilized aircraft with relatively lower complexity and automation than the Parts 121 and 135 operations. Hence, the Part 91 aircraft should have less diverse failure modes across safety report types; whether an ASRS case resulted in an NTSB case (an accident) or not might be more dependent upon the action(s) of the pilot(s).

Regarding the Parts 121 and 135 operations, the univariate MANOVA test was significant, indicating Belcastro LOC-I Hazard rates differed significantly across each group regarding vehicle impairment. The top estimated marginal means, per Appendix M3, were from the NTSB Classified and ASRS Augmented groups. For operations under Parts 121 and 135, the ASRS Augmented search revealed a higher quantity per flight hour of aircraft impairment information than NTSB accident investigations. This result demonstrated the usefulness of lower severity events from voluntary safety reporting in obtaining the volume of hazard information for vehicle impairment. An analogous scenario would be reporting B737-MAX LOC-I precursors in VSRs before the hull losses. A final observation from the analysis was that the NTSB Augmented and ASRS Classified groups yielded lower vehicle impairment rates. Therefore, it was not resource effective to deploy additional resources to perform an augmented search from the NTSB database, nor was it appropriate to rely solely on ASRS-classified data for vehicle impairment. The discrepancy between ASRS Classified and Augmented cases suggested that coding in the ASRS system for this particular Belcastro LOC-I Hazard was less effective.

Univariate Analysis on Adverse Onboard Conditions - System and Components Failure / Malfunction. Univariate tests for both Parts 121 and 135 and Part 91 datasets resulted in significant results, indicating that, for both datasets, the means of system and components failure / malfunction rates were significantly different. An examination of the relevant estimated marginal means plots on the Belcastro LOC-I Hazard rates in Appendix M3 and M5 showed different patterns between the two datasets. The NTSB Classified group gave the highest mean, followed by the ASRS Augmented group for Parts 121 and 135. The ASRS Augmented group was approximately 50% less than the NTSB Classified group for Parts 121 and 135. This result was expected given the rigor and independence of NTSB investigations, which revealed complex system and component failure and malfunction in accidents. For Part 91, the highest rate was the NTSB Augmented search. This result suggested that additional information would be available from an augmented keywords search which were precursors to a LOC-I. ASRS groups indicated around one-third of the marginal means of the NTSB groups. This suggested that even Heinrich’s common causes hypothesis was valid between lower and higher severity events, but the ratio differed from the claimed ratio in Henrich's Triangle (Heinrich, 1931). Therefore, solely using ASRS would not be sufficient to cover the hazard rate captured by NTSB on system and component failures for Part 91 operations.

Univariate Analysis on Adverse Onboard Conditions - Crew Action / Inaction. The MANOVA results showed significant differences in the means of crew action / action rates for both Parts 121 and 135 and Part 91 datasets. Both datasets indicated that the NSTB Classified group provided the highest means. However, per Appendix M3, the NTSB Augmented group shared the lowest hazard rate of the marginal mean. This suggested that the NTSB Augmented search was not useful in identifying cases with further crew action / inaction hazards.

Moreover, combining ASRS Classified and Augmented gave results approximately a third of the magnitude lower than NTSB Classified, suggesting that both Classified and Augmented Groups need to be considered when identifying crew action / action errors when only ASRS was originally to be used by the researcher. The situation was different with Part 91. Per Appendix M6, TX3 estimated marginal means plot, the NTSB Augmented group had the second highest crew action / action factor rate, compared with both ASRS groups, which recorded the lowest rates. This observation was analogous to findings from previous research on reporting biases and voluntary reports, which suggested individuals were unlikely to self-report errors voluntarily (see Chapter II). In summary, ASRS had limited utility in identifying general aviation crew action / inaction hazards.

Univariate Analysis on External Hazards and Disturbances - Inclement Weather or Atmospheric Disturbances. The MANOVA results showed divergence in the means of inclement weather / atmospheric disturbances rates for both Parts 121 and 135, and Part 91 datasets. For both datasets, per Appendices M3 and M6, the NTSB Classified group showed the highest means. In the Parts 121 and 135 dataset, the NTSB Augmented group had the lowest means of the four groups, indicating the low effectiveness of augmented search in uncovering this hazard. The ASRS Coded group and the ASRS Augmented group combined amounted to approximately half of the level of the marginal means recorded by the NTSB classified group. Therefore, when the ASRS reports were used, it would be more effective for both classified and augmented groups to be deployed by research to obtain more comprehensive information. For the Part 91 dataset, although the NTSB Classified group had the highest marginal mean, the NTSB Augmented group also indicated half of the mean level. Hence, the augmented search method could substantially provide an additional volume of inclement weather / atmospheric disturbances information in the NTSB dataset for general aviation. The rate of the ASRS Coded group was approximately half of the NTSB Classified group. Therefore, a combined NTSB, NTSB Augmented, and ASRS Coded reports dataset are expected to provide the optimum coverage of inclement weather and atmospheric disturbances. The ASRS augmented group, on its own, however, would not be sufficient in providing an adequate volume of hazard rate information for inclement weather / atmospheric disturbances.

Univariate Analysis on External Hazards and Disturbances - Poor Visibility. Both Parts 121 and 135 and Part 91 datasets did not produce significant MANOVA results that suggested variations among groups of safety reports on poor visibility. Therefore, in terms of obtaining visibility-related hazards, there was no evidence that a specific type of safety report would harvest a more superior rate. This observation inferred that any reporting type could obtain the same level of poor visibility hazard rate.

Univariate Analysis on External Hazards and Disturbances - External Hazards and Disturbances – Obstacle. The MANOVA results suggested that the differences across groups in the Parts 121 and 135 dataset were statistically significant, while no significant differences were identified in the Part 91 dataset. Hence, for the Part 91 dataset, there was no evidence to support that a specific type of safety report would harvest a more superior rate of obstacle hazards. For the Parts 121 and 135 dataset, the NSTB Classified group had the highest marginal mean while the contribution of the NTSB Augmented group was minimal. Though the combined classified and augmented groups did not meet the hazard rate for identifying obstacle hazards, for the ASRS dataset to be more effective, the classified and augmented datasets had to be used together, as separately, each only one-third of the hazard rate of the NTSB group.

Univariate Analysis on Abnormal Vehicle Dynamics and Upsets - Abnormal Vehicle Dynamics. Both Parts 121 and 135 and Part 91 datasets had significant results in the MANOVA, indicating significant differences across the safety reporting groups. Both datasets displayed patterns that were dissimilar to the other DVs. The ASRS Classified group demonstrated the peak hazard rates, whereby the hazard rates of the other three groups were minimal. This showed that the number of abnormal vehicle dynamics per flying hour was the highest in the ASRS classified case. A possible explanation for this was that abnormal vehicle dynamic was a precursor to aircraft going into upset condition, as demonstrated by the LOC-I bowtie model developed by the UK Civil Aviation Authority (Civil Aviation Authority, n.d.). Many of these events have been recovered after abnormal vehicle dynamics before an aircraft went into upset condition. Hence, these occurrences would not be classified as accidents that otherwise had to be investigated by the NTSB. These recovered cases, as well as the events that experienced abnormal vehicle dynamics and further developed into LOC-I accidents, would have been reported in voluntary safety reports by the flight crew in the ASRS system.

Univariate Analysis on Abnormal Vehicle Dynamics and Upsets - Vehicle Upset Conditions. Only the Parts 121 and 135 dataset showed significant differences across safety reporting groups. The NTSB Classified group and the NTSB Augmented Group showed the lowest hazard rates with one-twelfth of the ASRS levels per the estimated marginal means plots in Appendix M3. The ASRS Classified and Augmented groups showed similar high hazard rates in abnormal vehicle dynamics. This finding suggested that ASRS reports focused more on the later part of the causation chain (Reason, 2016), which is upset, while the NTSB tended to focus on the earlier parts of the chain (e.g., human errors or mechanical failures that caused upset conditions). For the Part 91 dataset, as there was no significant difference across groups, the researcher would not recommend using a particular type of safety report vehicle for obtaining upset condition hazard rates.

Effectiveness of Augmented Searches and Dependency on Classified Searches

Chapter II discusses the substantial human and financial resources required to implement SMS (Okwera, 2016). Hence, an effective approach to retrieve relevant hazard information using the most relevant SMS database, safety report type, and search method to obtain the highest quantity and contextual content are essential. While Belcastro et al.

(2018) uncovered LOC-I events that were not officially classified as LOC-I in the NTSB databases using the augmented *keyword search*, results from this research suggested that augmented search was ineffective in enriching the classified groups for the entire set of eight Belcastro LOC-I Hazards. This was indicated by the lack of significant differences in MANOVA univariate results and the relatively low marginal means with some augmented groups. Belcastro LOC-I Hazards which were insensitive to augmented searches were:

a. abnormal vehicle dynamics DV for both commercial and general aviation in both NTSB and ASRS groups

b. all the Belcastro LOC-I Hazards DVs in general aviation in the ASRS group

The corresponding hazard rates have not increased substantially with augmented searches, suggesting that a search of the events with the related Belcastro LOC-I Hazards using classified search was adequate. This observation could be partially explained by the rigor and independence applied in the investigations (ICAO, 2016) by NTSB. The relevant Belcastro LOC-I Hazards were identified effectively through the investigation process. This negated the need to expend resources to perform augmented search analysis. Also, as the nature of the ASRS was self-reporting, the depth of factors being reported might not be as deep as those reported by NTSB; this explains why there was a lower estimated mean from general aviation reports. Further research would be necessary to understand the reasons behind these observations conclusively:

i. NTSB augmented search was effective (i.e., higher hazard rate means) for the DVs with significant univariate MANOVA results in general aviation, apart from abnormal vehicle dynamics.

ii. The effectiveness of ASRS-augmented searches in commercial aviation was high (i.e., higher hazard rate means) but not in general aviation.

Implications for Heinrich Principles

Suppose Heinrich's common cause hypothesis (Davies, 2003) was to hold. In that case, the causes in the lower severity LOC-I events reported in ASRS should be the same as those in NTSB, and Heinrich triangle's 300:29:1 ratio (Heinrich, 1931) would be met. From the results of the MANOVA study, the multivariate analysis results showed that, in terms of hazard rates, safety report types (IV) had significant effects on the set of Belcastro LOC-I Hazards (DVs). This premise was supported by the lower severity ASRS reports which showed statistically significant differences in mean hazard rates with the higher severity NTSB reports for both Parts 121 and 135, and Part 91 datasets. The ratio implied in the Heinrich Triangle (Heinrich, 1931), 300:29:1, was also tested in this study. Based on the means values documented in Appendix M1 and M4, this ratio was not met. Therefore, the quantitative results did not support Heinrich's principles for all four types of safety reports on the complete set of Belcastro LOC-I Hazards.

However, pockets of univariate relationships were not significantly different in hazard rates across different types of safety reports. Safety report type (IV) might have no or insignificant effect on the Belcastro LOC-I Hazards, indicating that they could be the same statistically. The DVs under these conditions for general aviation were:

i. adverse onboard conditions - vehicle impairment,

ii. external hazards and disturbances - poor visibility,

iii. external hazards and disturbances – obstacle, and

iv. abnormal vehicle dynamics and upsets - vehicle upset conditions

The possibility of the same hazard across the four groups only applies to external hazards and disturbances - poor visibility for commercial aviation. For these, quantitatively, there was a potential for Heinrich Principles on common causality to be applicable regarding hazard distribution as the MANOVA results did not produce any contraindications against such application. This finding was similar to the research result from the rail industry mentioned in the literature review (Wright, 2002). Some causal factors were not significantly different across severities of rail incidents. This could be an area for further research.

From the qualitative perspective, while this study was not intended to compare the factors behind each mapped Belcastro LOC-I Hazard, the tree maps, hierarchy charts, and word clouds analyses discussed in Chapter IV suggested that, despite some similarities, not all top 10 frequent words in the narratives and synopsis were similar. This further indicated that the factors contributing to the various hazard rates differed. Moreover, the hierarchy charts shown in Appendix P indicated dissimilar patterns among each reporting type. This suggested that interrogating one database might not provide equivalent factors on a particular hazard or accident type regardless of the hazard rates. Therefore, the applicability of Heinrich's common cause hypothesis (Davies, 2003) to LOC-I cases from the contextual perspective was limited. It could therefore be inferred that, qualitatively, there was insufficient evidence to support that the causes of high-severity events were the same as those of low-severity events. Hence, viewing LOC-I through the lens of Heinrich's theories, per Figure 10, would not be appropriate.

Understanding the Strengths and Weaknesses of Reporting Systems

Despite the level of rigor and independence in investigating NTSB cases, this study found that the NTSB database was not consistently the most effective in identifying the eight Belcastro LOC-I Hazards. On the contrary, although the rigor in investigation and follow-up was less for ASRS, an *open-loop* VSR such as ASRS was not less superior in capturing some Belcastro LOC-I Hazards than NTSB. Therefore, it would not be appropriate for the NTSB database to be deployed as the default database for LOC-I research to comprehensively survey the entire set of Belcastro LOC-I Hazards in the industry. Instead, a targeted approach on the data source to be deployed based on an understanding of the limitation and effectiveness of each data source for specific Belcastro LOC-I Hazard would provide the most effective results.

Before selecting the data source, researchers and safety practitioners should consider the purpose of their research, understand the possible limitations and biases highlighted in this research, and the characteristics of each of the Belcastro LOC-I, as summarized in the recommendations section under Table 28. For example, suppose a researcher is interested in understanding how abnormal vehicle dynamics contribute to LOC-I situations for general aviation. The ASRS database may be a more appropriate option in this case due to the highest hazard rate. On the other hand, if a researcher is interested in how aircraft component system failures could lead to a LOC-I event for commercial aviation, then the NTSB database would be more appropriate.

Lastly, as this research highlighted weaknesses to specific Belcastro LOC-I Hazards in safety reporting systems, focused safety assurance activities can be arranged by regulators or the operator's assurance organization. For Belcastro LOC-I Hazards not effectively identified by an open-loop VSR such as ASRS, in the absence of other credible data, the safety assurer may decide to elevate the rigor and frequency of the safety assurance activities for these hazards to more of a *command-and-control* approach (Mills, 2011) to ascertain the hazard has been understood, assessed and mitigated.

Insights from Cursory Qualitative Analysis on the Narratives' Content and AIDS Dataset. Statistical differences in Belcastro LOC-I Hazard rates across each safety report group have been established in the formal MANOVA analysis. Implications of such differences were further explored with a cursory analysis of the narratives' content and the AIDS dataset. Statistically, a comparatively lower hazard rate inferred less information quantity per flying hour for the Belcastro LOC-I Hazard, and vice versa. For researchers and safety practitioners, the quantitative information related to each Belcastro LOC-Hazard and the context behind the identified hazards are essential for accurate diagnosis and appropriate mitigations when interrogating a safety database. Such contextual information might not reside in the coded DVs as each report was text rich.

Narratives' Content Insights. The word clouds in Figure 25 show the distribution of word counts in order of appearance, while hierarchy charts in P14 to 17 show hierarchical data as sets of nested rectangles of varying sizes, highlighting some themes of the data. The size of the rectangle represents the amount of coding at each node. Similar distribution of the word clouds or shapes of the hierarchical charts indicates similar contextual information of the safety reports. Based on Heinrich's common cause hypothesis (Heinrich, 1931), even with lower hazard rates, the context of the factors related to the hazard could be obtained, allowing appropriate mitigation measures to be applied.

As detailed in Chapter IV, examining the word clouds, hierarchy charts, and tree maps suggested that most of the top 10 keywords in the narratives were similar across different safety report groups, though some subtle differences also existed. For example, when examining the Parts 121 and 135 tree maps and hierarchy charts in Appendix N, as summarized in Table 24, the word *pilot* was missing in both the ASRS classified and augmented groups. As ASRS is a self-reporting system, reporting bias on the action of the *pilot*, in many cases, the originator of the reports, might be prevalent in the commercial aviation sector (Flynn et al., 2018; Hudson et al., 2006; Noble & Pronovost, 2010; Noort et al., 2016). This has been highlighted as an opportunity for future research. Secondly, while there were some similarities across groups in the common texts, such as *aircraft* and *flight*, based on the narratives and synopsis' qualitative analyses, there was insufficient evidence to suggest that the contextual information behind causal factors and contributory factors identified by the Belcastro LOC-I Hazards were the same across each group, meaning that some factors being retrieved in a lower hazard rate group might not be featured in a higher hazard rate group, and vice versa. This difference further supported the MANOVA multivariate results.

AIDS Data Insights. While the volume of classified and augmented searched LOC-I reports identified from the AIDS database was not as high as ASRS and NTSB, the AIDS narratives hierarchy chart and tree maps in Appendix N provided deeper insight into the technical or mechanical causal factors of the narrative of AIDS. However, they did not provide information on human factors such as crew actions / inactions. This aligns with the notion that AIDS is a safety reporting system between ASRS and NTSB in terms of investigation rigor and severity of events. The technical insight was comparatively higher, but it fell short of what NTSB investigations have offered. It is suggested that further research can be performed to apply Belcastro LOC-I Hazards coding on AIDS for statistical comparisons with the NTSB and ASRS databases in a quantitative manner.

Conclusions

The purpose of the study was to (a) establish if there are differences in the hazards identified between voluntary and mandatory LOC-I safety reports in the U.S. commercial and general aviation environments and (b) to identify the particular Belcastro LOC-I Hazard that displays significant differences between voluntary and mandatory LOC-I reports. Both purposes have been achieved by establishing the differences in the hazards between voluntary and mandatory LOC-I safety reports from the multivariate and univariate levels, using the quantitative MANOVA method, supplemented by discriminant analysis and qualitative analysis using NVivo®.

The key findings from this research are that at a multivariate level, the types of safety reports significantly affected the set of Belcastro LOC-I Hazard rates for both commercial and general aviation. Also, at the univariate level, not all Belcastro LOC-I Hazards rates varied with the types of safety reports. For general aviation, the hazard rates that did not statistically differ were:

a. adverse onboard conditions - vehicle impairment

b. external hazards and disturbances - poor visibility

c. external hazards and disturbances – obstacle, and abnormal vehicle dynamics and upsets - vehicle upset conditions

For commercial aviation, only external hazards and disturbances - poor visibility rate did not differ across the four groups of safety reports. Based on these results, it could be concluded that there is no *one-size-fits-all solution for* selecting safety databases for effective research in LOC-I, as no one safety report type consistently produced the highest hazard rates through the whole set of Belcastro LOC-I Hazards. Instead, this research highlighted the importance of considering the information to be obtained (DV) before selecting the most effective safety reporting type for research. Also, when only limited safety report types were available, the research results highlighted that augmented search could increase the level of information for some specific hazards, but not all. This applies to the NTSB database on seven Belcastro LOC-I Hazards for general aviation and all eight Belcastro LOC-I Hazards for commercial aviation in the ASRS database only.

The qualitative analysis supplemented the quantitative results and highlighted differences in the narratives and synopsis patterns across the safety report types, suggesting that the reported factors differed between the ASRS and NTSB reports. One difference was the tendency of ASRS reports to cover the factors closer to the consequence of the causation chain. In contrast, NTSB reports covered more of the earlier parts of the causation chains, such as human factors. The results of the research did not support Heinrich's common cause hypothesis.

This study has shown the potential for further research to explore the reasons behind the differences and similarities among the distributions of Belcastro LOC-I Hazards in the various safety report types. Further investigations should also be undertaken to understand how ASRS can be enhanced so that hazards found in higher severity events could be more effectively identified and mitigated by the SMS proactively.

Finally, this study pointed to the need for a targeted approach when using a safety reporting database with a clear awareness of the strengths and weaknesses of each reporting system, as well as the characteristics of the Belcastro LOC-I Hazard being researched. The findings obtained can also inform the safety assurance strategy to be deployed. Having the intelligence to evaluate the strengths and weaknesses of the reporting system will enable safety professionals to interpret and assure data from the safety reporting system in a more calibrated manner, resulting in more effective safety mitigations.

Theoretical Contributions

This research has demonstrated that some of the Belcastro LOC-I Hazard rates in lower severity voluntary safety LOC-I reports for ASRS were different from those reported in the mandatory, higher severity reports for NTSB in a univariate manner. The variations, however, differed between types of operation (commercial or general) and the Belcastro LOC-I Hazard in question. Supplementary qualitative analysis suggested that the textual content of the narratives and the synopsis of the reports were different, such as less focus on the *pilot* for voluntary safety reports in ASRS reports for commercial aviation. Hence, this research has not validated Heinrich's triangle and common cause hypothesis.

Given the studies by Flynn et al. (2018), Noort et al. (2018), and Reader et al. (2016) on reporting biases, as well as Manuele (2011) questioning the validity of

Heinrich's theory on modern safety science, this research further adds to the body of knowledge on the applicability of open-loop voluntary safety reporting systems (such as ASRS), mandatory safety reporting systems (such as NTSB) and the applicability of Heinrich's principles to LOC-I safety reports. This research contributes to existing knowledge of voluntary and mandatory safety reporting efficacies in the following ways:

a. At the multivariate level, the type of safety reporting affected Belcastro LOC-Hazards' rates for both commercial and general aviation.

b. For some Belcastro LOC-I Hazards, at a univariate level, the effect of the safety reporting type on the rates of Belcastro LOC-Hazards was not significant (i.e., adverse onboard conditions - vehicle impairment, external hazards, and disturbances - poor visibility, external hazards and disturbances – obstacle, and abnormal vehicle dynamics and upsets - vehicle upset conditions for general aviation and external hazards and disturbances - poor visibility only for commercial aviation).

c. ASRS reports are not necessarily less effective than NTSB reports in obtaining hazard information.

d. Based on this study's results, there is an opportunity to perform a targeted search on Belcastro LOC-I Hazard using the most appropriate safety report type.

e. This study should be considered as a valid source as the significant level of $p < .001$ was reached at a univariate level with a large effect size generally, validated by discriminatory analysis and supported by qualitative analysis.

f. Findings in this study on some causal factors traversing through severities of safety events were similar to the research result in the rail industry (Wright, 2002).

Practical Contributions

The primary practical contribution is to provide intelligence to aviation safety practitioners regarding the comparative strengths and weaknesses of the ASRS and NTSB reporting systems and report identification techniques as part of this research. This intelligence is particularly important as not all operators' safety reporting systems are equipped with experienced investigators to analyze safety events and provide in-depth root cause analysis. The quantitative analysis and the qualitative insights in this research have highlighted areas where NTSB and ASRS are deficient in informing hazards behind LOC-I events. The findings in this research have been condensed into a set of recommendations for safety practitioners in Table 28. Therefore, when a safety manager processes or takes reference from a publicly available open-loop VSR such as ASRS, results from this study can provide empirical evidence for alertness on possible deficiencies in the reported information. Proactive source data verification or supplementary information can be sought before deciding on mitigation. For example, pilots' actions or inactions in ASRS commercial aviation cases should be challenged when reviewing an ASRS VSR. This will prevent a disproportionate use of resources in hazard identification and mitigation, driven by the immediately available information but not appreciating the deficiencies of information, and how to seek data augmentation. Lastly, this research also provided safety practitioners a cautionary message on the danger of relying on a single source of data when obtaining safety information, and the danger of blindly following Heinrich's principle of same causality and Heinrich's triangle. The same message also applies to regulators whereby a focused approach with strengthened rigor might be required for some Belcastro LOC-I Hazards not identified adequately by open-loop VSR, such as ASRS, if that was the only available source of data.

Limitations of the Findings

This research is limited to LOC-I events only. However, the characteristics of the data analyzed are not anticipated to be different from other critical hazards in aviation, such as runway excursions and controlled flight into terrain. This is worth validating and has been included as a recommendation for future research.

There were limitations in the methodology applied. Although the sample size was over the required 44 for MANOVA, as determined by GPower®, the sample size of 68 was relatively small. Also, not all the assumptions for the quantitative analysis were fully met, such as homogeneity Box's M and normality tests. However, this was mitigated by a strict alpha value of $p < .001$ and data transformation application.

Although this research was limited to commercial and general aviation, other mass transportation industries such as rail or marine also collect vast data in their management systems. The challenge is the lack of international standardization on taxonomy and coding for rail. Hence the effectiveness of the augmented search for trains' safety systems is also worth exploring and has been included as a recommendation for future research.

Finally, this research was based on voluntary safety reports from ASRS. The rigor of these investigations and the feedback loop are unique to ASRS administered in the

United States. Therefore, it is important not to generalize the results to other voluntary safety report systems without validating the systems concerned.

Recommendations

Two sets of recommendations have been suggested: (a) to guide safety practitioners in making use of the research results so that databases from safety management systems are interrogated in an optimized manner to avoid the potential pitfalls discovered in this research, and (b) to set the strategy for future research. Table 26 shows the recommendations made from answering RQ1 and RQ2.

Table 28

Recommendations from Answers to RQ1 and RQ2 to Safety Practitioners

Areas of Interest (Hypothesis)	Recommendations – Commercial Aviation	Recommendations – General Aviation
Multivariate analysis – application of safety reporting database	• Despite the difference in hazard rates predicted, the researcher should be cognizant of the potential biases in the content. For example, Voluntary Safety Reports have a bias of not mentioning the pilot in commercial aviation. • To address the biases, keyword searches should be considered when searching safety databases, requiring the researcher to understand the precursors to the hazards they are interested in.	
Vehicle Impairment	Apply Augmented Search in ASRS. NTSB classified data provides a high level of content.	Any safety report type can be used.
System and components failure / malfunction	NTSB augmented search is unnecessary to enrich the hazards identified from the classified search. ASRS classified and augmented searches are to be considered together.	NTSB Augmented search is recommended alongside NTSB classified search. ASRS is not to be solely depended upon for comprehensive LOC-I data provision.
Crew action / inaction	NTSB augmented search is unnecessary to enrich the hazards identified from the classified search. ASRS classified and augmented searches are to be considered together.	ASRS is not to be depended upon solely for the comprehensive provision of LOC-I data on crew action / inaction.
Inclement weather / atmospheric disturbances	NTSB augmented search is unnecessary to enrich the hazards identified from the classified search. ASRS classified and augmented searches are to be considered together.	NTSB classified provides a high level of content. ASRS augmented is unnecessary to enrich the hazards identified.
Poor visibility	Any safety report type can be used.	Any safety report type can be used.
Obstacle	NTSB augmented is unnecessary to enrich the hazards identified from the classified search. ASRS classified and augmented searches are to be considered together.	Any safety report type can be used.
Abnormal vehicle dynamics	ASRS classified can be used as the primary source.	ASRS Classified is to be used as the dominant source of abnormal vehicle dynamics.
Vehicle upset conditions	ASRS classified or augmented reports are to be used when identifying hazard rates.	Any safety report type can be used.

Further recommendations to safety practitioners are:

a) Promote the application of VSRs for the Belcastro LOC-I Hazards identified in this research that did not show statistical differences between ASRS and NTSB databases.

b) Provide proforma-based VSR reporting forms to encourage reporting on the areas of deficiencies identified in this research.

c) Strengthen the rigor of safety assurance activities to more command and control (Mills, 2011) for the Belcastro LOC-I Hazards highlighted as significantly different across report types, if only lower severity reports are available.

In terms of recommendations for future research, the limitations identified by this research could be further explored to allow for a broader generalization of the research results.

d) Extend this research to other types of aviation safety-critical events in aviation such as CFIT and runway excursions. Establish a bow-tie model into the hazard and search on the precursors identified from the *bow-tie* as the keywords. Test the validity of the findings in this study in other accident types.

e) While many of the ASRS cases have been mitigated in flight, preventing them from developing into accidents, a *what has gone right* research is to be conducted to capture the effective barriers deployed.

f) Assess the level of self-reporting biases described by scholars in Chapter II (Flynn et al., 2018; Hudson et al., 2006; Noble & Pronovost, 2010; Noort et al., 2016) in voluntary safety reports. The focus should be on reporting critical hazards such as LOC-I and CFIT. If there is research evidence that the reporting has been biased against comprehensive reporting, introducing specific education programs or amending the reporting form can help encourage relevant personnel to actively report areas that have traditionally been underreported.

g) Explore why NTSB augmented search was adequate for the DVs with significant univariate MANOVA results in general aviation, apart from abnormal vehicle dynamics, and why ASRS augmented search's effectiveness in commercial aviation was high, but not in general aviation.

h) Repeat the same quantitative research, including Belcastro LOC-I Hazards coding on AIDS for statistical comparison with the NTSB and ASRS databases.

i) Extend the research to other VSRs, such as those administered by airlines within their systems, whereby an increased rigor of investigation and feedback with the originators are possible.

The research has added to the body of knowledge in ASRS as a data source for informing hazards in high-severity LOC-I events. The results provide further contributions regarding Heinrich's (1931) principles theoretically, and to modern safety management in aviation practically. The gaps for ASRS in providing the level of information have been highlighted with recommendations on how these can be filled, or how safety practitioners should interpret ASRS. Finally, this research has highlighted the potential for further research to understand the reasons behind the deficiencies in ASRS in the context of this research, which spans across human biases and safety reporting systems design.

References

Aeronautics and Space, 14 C.F.R. § 1-199 (2022).

Alamgir, H., Yu, S., Gorman, E., Ngan, K., & Guzman, J. (2009). Near miss and minor occupational injury: Does it share a common causal pathway with major injury? *American Journal of Industrial Medicine*, *52*(1), 69-75. doi:10.1002/ajim.20641

Allen, P. J., & Bennett, K. (2008*). SPSS for the health and behavioral sciences* (1st ed.).Thomson Learning.

Anderson, C. L. (2013). *The effects of aircraft certification rules on general aviation accidents*. Embry-Riddle Aeronautical University

Bałaj, B., Lewkowicz, R., Francuz, P., Augustynowicz, P., Fudali-Czyż, A., Stróżak, P., & Truszczyński, O. (2018). Spatial disorientation cue effects on gaze behaviour in pilots and non-pilots. *Cognition, Technology & Work*, 21(3), 473-486. doi:10.1007/s10111-018-0534-7

Balogh, S. D. (2017). Use of FOQA Data to Estimate the Probability of Vehicle Upset or Loss of Control In-Flight. ProQuest Dissertations Publishing.

Barstow, D. (2010). Regulators failed to address risks in oil rig fail-safe device. *New York Times.* June 20, 2010 edition.

Basford, B. (2017). The Heinrich model: Determining contemporary relevance. ProQuest Dissertations Publishing.

Bazeley, P. (2013). *Qualitative data analysis: Practical strategies* (1st ed.). Sage.

Belcastro, C., & Foster, J. (2010). Aircraft loss-of-control accident analysis. doi:10.2514/6.2010-8004. Langley Research Center.

Belcastro, C. M. (2012). Validation and verification (V&V) of safety-critical systems operating under off-nominal conditions. Langley Research Center.

Belcastro, C.M., Foster, J., Newman, R.L., Groff, L., Crider, D.A., & Klyde. D.H., (2014). Preliminary Analysis of Aircraft Loss of Control Accidents: Worst Case Precursor Combinations and Temporal Sequencing. In AIAA Guidance, Navigation, and Control Conference. https://doi.org/10.2514/6.2014-0612

Belcastro, C. M., Newman, R. L., Crider, D. A., Klyde, D. H., Foster, J. V., & Groff, L. (2016). Aircraft loss of control: Problem analysis for the development and validation of technology solutions. Langley Research Center.

Belcastro, C. M., Foster, J. V., Shah, G. H., Gregory, I. M., Cox, D. E., Crider, D. A., . . . Klyde, D. H. (2017). Aircraft loss of control problem analysis and research toward a holistic solution. *Journal of Guidance, Control, and Dynamics*, *40*(4), 733-775. doi:10.2514/1.G002815

Bellamy, L. J., Ale, B. J. M., Whiston, J. Y., Mud, M. L., Baksteen, H., Hale, A. R., Papazoglou, I. A., Bloemhoff, A., Damen, M., & Oh, J. I. H. (2008). The software tool storybuilder and the analysis of the horrible stories of occupational accidents. Safety Science, 46(2), 186–197. https://doi.org/10.1016/j.ssci.2007.06.022

Bellamy, L. J. (2015). Exploring the relationship between major hazard, fatal and non-fatal accidents through outcomes and causes. *Safety Science*, *71*, 93-103. doi:10.1016/j.ssci.2014.02.009

Bird, F. E. (1976). *Loss control management*. Institute Press.

Bureau of Transportation Statistics, (2020). Air Carrier Statistics. https://www.bts.gov/topics/airlines-and-airports/quick-links-popular-air-carrier-statistics

Bureau d'Enquêtes sur les Accidents de transport terrestre, (2020). Accident to the AIRBUS A380-861 equipped with Engine Alliance GP7270 engines registered F-HPJE operated by Air France on 30 September 2017 in cruise over Greenland (Denmark). *Investigation Report.* September 2020 Edition.

Boeing, (2022). Statistical Summary of Commercial Jet Airplane Accidents Worldwide Operations 1959-2021. 53rd Edition. Boeing.

Cacciabue, P. C., Cassani, M., Licata, V., Oddone, I., & Ottomaniello, A. (2015). A practical approach to assess risk in aviation domains for safety management systems. *Cogn Tech Work*, 17, 249-267. doi:10.1007/s10111-014-0294-y

Campbell, D. M. (2015). An assessment of predominant causal factors of pilot deviations that contribute to runway incursions. ProQuest Dissertations Publishing.

Castellanos, E., Flores-Andrade, A., & Flynn, M. A. (2018). Safety across cultures: Understanding the challenges. *Professional Safety*, *63*(1), 28-32.

Chilakalapudi, N. S. K. (2016). Comparative Qualitative Research Distinguishing Safety Features Among Aviation Safety Action Programs in the United States Airlines. ProQuest Dissertations Publishing.

Civil Aviation Authority. (n.d.). Bow-tie analysis on loss of controls. Retrieved April 22, 2019,from https://www.caa.co.uk/Safety-Initiatives-and-Resources/Working-with-industry/Bowtie/

Cooper, M. D. (2019). The efficacy of industrial safety science constructs for addressing serious injuries & fatalities (SIFs). *Safety Science*, *120*, 164-178. doi:10.1016/j.ssci.2019.06.038

Darveau, K. (2015). Roadmap to the implementation of best practices in human factors voluntary reporting for safety. ProQuest Dissertations Publishing.

Davies, J., Ross, A., & Wallace, B. (2017). Safety Management : A Qualitative Systems Approach (First edition.). CRC Press. https://doi.org/10.1201/9781315272764

De Veaux, R. D., Velleman, P. F., & Bock, D. E. (2013). *Stats: Data and models*. Pearson Education.

Dekker, S. (2019). *Foundations of safety science: A century of understanding accidents and disasters*. CRC Press.

Ethiopian Airplane Accident Investigation Bureau. (2022). Investigation Report on Accident to the B737-Max 8 Reg. ET-AVJ Operated by Ethiopian Airlines 10 March 2019. http://avherald.com/files/ethiopian_b38m_et-avj_190310_final_report_20221226.pdf

European Aviation Safety Agency. (2017). Internal Occurrence Reporting System Facts and Figures 2016. https://www.easa.europa.eu/en/downloads/23385/en

European Aviation Safety Agency. (nd). Internal Occurrence Reporting Access, Dissemination and Data Protection Policy. https://www.easa.europa.eu/en/domains/safety-management/aviation-safety-reporting/occurrence-reporting-protection-information-sources

European Commission. (2015). The European Aviation Safety Programme Document (2nd ed.): Report from the commission to the European parliament and the council (No. COM 599). https://eur-lex.europa.eu/resource.html?uri=cellar:f0a0e4cd-9ce8-11e5-8781-01aa75ed71a1.0020.02/DOC_2&format=PDF

European Aviation Safety Agency. (2021). The European Plan for Aviation Safety. https://www.easa.europa.eu/en/downloads/134920/en

European Union Aviation Safety Agency. (2023). *The difference between EPAS, SPAS and SMS - An article from EASA supported by safety management partners and the European safety promotion network.* https://www.easa.europa.eu/en/difference-between-epas-spas-and-sms

Federal Aviation Administration, (2006). Presentation to Aerospace Control and Guidance Systems Committee of ACSGC on Safety Management System. Federal Aviation Administration.

Federal Aviation Administration, (2007). Presentation to Second ICAO Global Symposium on TEM / NOSS in Air Traffic Control on Normal Operations Data : Air Traffic Facility Evaluations and NOSS. Federal Aviation Administration.

Federal Aviation Administration, (2010). *Element 3.2 Management of Change. Safety Management System Framework*, June 1, revision 3, p.23.

Federal Aviation Administration, (2015). Safety Management Systems for Domestic, Flag, and Supplemental Operations Certificate Holders. *14 C.F.R. § 5 & 119*. Docket No. FAA–2009–0671; Amendment Nos. 5–1 and 119–17.

Federal Aviation Administration, (2020). General Aviation and Part 135 Activity Surveys. Federal Aviation Administration. https://www.faa.gov/data_research/aviation_data_statistics/general_aviation\

Field, A. P., & Field, A. P. (2020). Discovering statistics using IBM SPSS statistics (Fourth edition.). SAGE Publications.

Flynn, M. A., Castellanos, E., & Flores-Andrade, A. (2018). Safety across cultures: Understanding the challenges. *Professional Safety, 63*(1), 28-32.

Friend, M. A., & Kohn, J. P. (2014). *Fundamentals of occupational safety and health* (6th ed.). Bernan Press.

Gallivan, S., Taxis, K., Franklin, B. D., & Barber, N. (2008). Is the principle of a stable Heinrich ratio a myth?: A multimethod analysis. *Drug Safety*, *31*(8), 637-642. doi:10.2165/00002018-200831080-00001

Gesinger, S. (2017). *The fearless world of professional safety in the 21st century*. Routledge. doi:10.4324/9781315178493

Gilbey, A., Tani, K., & Tsui, W. H. K. (2016). Outcome knowledge and under-reporting of safety concerns in aviation. *Applied Cognitive Psychology*, *30*(2), 141-151. doi:10.1002/acp.3179

Gill, G. K., & Shergill, G. S. (2004). Perceptions of safety management and safety culture in the aviation industry in New Zealand. *Journal of Air Transport Management*, *10*(4), 231-237. doi:10.1016/j.jairtraman.2004.02.002

Gore, A. (1997). White House Commission on Aviation Safety and Security Final Report. February 12, 1997. http://www.fas.org/irp/threat/212fin%7E1.html

Grami, A. (2019). *Probability, random variables, statistics, and random processes: Fundamentals & applications* (1st ed.). Wiley. doi:10.1002/9781119300847

Grami, A. (2020). *Probability, random variables, statistics, and random processes: Fundamentals & applications* (1st ed.). Wiley.

Gunningham, N., & Grabosky, P. (1998). Smart Regulation. Oxford University Press.

Hair J. F., Black W. C., Babin B. J. & Anderson R. E. (2019). *Multivariate data analysis* (Eighth). Cengage Learning EMEA.

Heinrich, H. W. (1931). Industrial accident prevention: A scientific approach (1sted.). McGraw-Hill. doi:571338960

Hudson, P., Parker, D., & Lawrie, M. (2006). A framework for understanding the development of organisational safety culture. Safety Science, 44(6), 551–562. https://doi.org/10.1016/j.ssci.2005.10.004

International Air Transport Association. (2015). Loss of control in-flight (LOC-I) prevention: beyond the control of pilots. 1st Edition.

International Air Transport Association. (2018). Safety report 2017. (No. 54th Edition). Retrieved from http://www.iata.org/publications/Pages/safety-report.aspx in May, 2018

International Air Transport Association. (2019). Guidance Material and Best Practicees. *Loss of Control In-Flight Accident Analysis Report Edition 2019*. Retrieved from chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.iata.org/contentassets/b6eb2adc248c484192101edd1ed36015/loc-i_2019.pdf

International Civil Aviation Organization. (2001a). Annual Report 2019. *The Word of Air Transport in 2019.* https://www.icao.int/annual-report-2019/Pages/the-world-of-air-transport-in-2019.aspx

International Civil Aviation Organization. (2001b). International standards and recommended practices: Annex 13 to the convention on international civil aviation aircraft accident and incident investigation. http://www.emsa.europa.eu/retro/Docs/marine_casualties/annex_13.pdf

International Civil Aviation Organization. (2009). Review of the classification and definitions used for civil aviation activities.

International Civil Aviation Organization. (2013). Doc 9859: Safety management manual. (3rd ed.). Retrieved from https://www.icao.int/safety/SafetyManagement/Documents/Doc.9859.3rd%20Edition.alltext.en.pdf in December 2019

International Civil Aviation Organization. (2016a). Global aviation safety plan. (Doc. No. 10004).

International Civil Aviation Organization. (2016b). International standards and recommended practices: Annex 19 to the convention on international civil aviation safety management. (2nd ed.). https://elibrary.icao.int/reader/250466/&returnUrl%3DaHR0cHM6Ly9lbGlicmFyeS5pY2FvLmludC9leHBsb3JlO3NlYXJjaFRleHQ9YW5uZXglMjAxOSUyMCVFMiU4MCU5NCUyMHNhZmV0eSUyMG1hbmFnZW1lbnQ7cGhyYXNlTWF0Y2g9MC9wcm9kdWN0LWRldGFpbHMvMjUwNDY2?productType=ebook

International Civil Aviation Organization. (2016c). In ICAO (Ed.), Annex 13 to the convention on international civil aviation aircraft accident and incident investigation (11th ed.). chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.kenyoninternational.com/Assets/Downloads/cos-fa-icao-annex-13-11-ed.pdf

International Civil Aviation Organization. (2019). State of Global Aviation Safety. Retrieved from chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.icao.int/safety/Documents/ICAO_SR_2019_final_web.pdf

International Civil Aviation Organization. (2020). Safety Report 2020. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.icao.int/safety/Documents/ICAO_SR_2020_final_web.pdf

International Civil Aviation Organization. (2022). Doc 10004, Global Aviation Safety Plan. https://www.icao.int/safety/GASP/Documents/10004_en.pdf

International Committee for Aviation Training in Extended Envelopes, (nd). Defining the Loss of Control In-Flight Threat. ICAO. Retrieved from http://icatee.org/defining-the-loss-of-control-in-flight-threat/ on December 1, 2020.

Irwin, W. J. (2017). *Airline Pilot Situation Awareness Models: Providing a Framework for Meta-Cognition, Reflection, and Education*. ProQuest Dissertations Publishing

Janic, M. (2000). An assessment of risk and safety in civil aviation. *Journal of Air Transport Management*, *6*(1), 43-50. doi:10.1016/S0969-6997(99)00021-6

Jausan, M., Silva, J., & Sabatini, R. (2017). A holistic approach to evaluating the effect of safety barriers on the performance of safety reporting systems in aviation organizations. *Journal of Air Transport Management*, *63*, 95-107. doi:10.1016/j.jairtraman.2017.06.004

Jeelani, I., Han, K., & Albert, A. (2018). Automating and scaling personalized safety training using eye-tracking data. *Automation in Construction*, *93*, 63-77. doi:10.1016/j.autcon.2018.05.006

Kenyi, L. (2019). General Aviation Accident Modeling and Causal Determination of Pilot Loss of Aircraft Control. ProQuest Dissertations Publishing.Kim, H., Kriebel, D., Quinn, M. M., & Davis, L. (2010). The snowman: A model of injuries and near-misses for the prevention of sharps injuries. *American Journal of Industrial Medicine*, *53*(11), 1119-1127. doi:10.1002/ajim.20871

Kim, H., Kriebel, D., Quinn, M. M., & Davis, L. (2010). The snowman: A model of injuries and near-misses for the prevention of sharps injuries. American Journal of Industrial Medicine, 53(11), 1119–1127. https://doi.org/10.1002/ajim.20871

Knecht, W. R. (2013). The “killing zone” revisited: Serial nonlinearities predict general aviation accident rates from pilot total flight hours. *Accident Analysis and Prevention*, *60*, 50-56. doi:10.1016/j.aap.2013.08.012

Kongsvik, T., Fenstad, J., & Wendelborg, C. (2012). Between a rock and a hard place: Accident and near-miss reporting on offshore service vessels. *Safety Science*, *50*(9), 1839-1846. doi:10.1016/j.ssci.2012.02.003

Kwatny, H. G., Dongmo, J. T., Chang, B., Bajpai, G., Yasar, M., & Belcastro, C. (2013). Nonlinear analysis of aircraft loss of control. *Journal of Guidance, Control, and Dynamics, 36(1),* 149-162. doi:10.2514/1.56948

Lambregts, A., Nesemeier, G., Newman, R., & Wilborn, J. (2012). Airplane upsets: Old problem, new issues. *AIAA Modeling and Simulation Technologies Conference and Exhibit (18 August 2008 - 21 August 2008).*

Leedy, P. D., & Ormrod, J. E. (2013). *Practical research: Planning and design* (10th ed.). Pearson.

Li, W.-C., Nichanian, A., Lin, J., & Braithwaite, G. (2022). Investigating the impacts of COVID-19 on aviation safety based on occurrences captured through flight data monitoring. *Ergonomics, ahead-of-print(ahead-of-print)*, 1–15. https://doi.org/10.1080/00140139.2022.2155317

Liao, M. (2015). Safety culture in commercial aviation: Differences in perspective between Chinese and western pilots. *Safety Science*, *79*, 193-205. doi:10.1016/j.ssci.2015.05.011

Lower, M., Magott, J., & Skorupski, J. (2016). Analysis of air traffic incidents using event trees with fuzzy probabilities. *Fuzzy Sets and Systems*, *293*, 50-79. doi:10.1016/j.fss.2015.11.004

Manuele, F. A. (2003). *On the practice of safety* (3rd ed.). J. Wiley.

Manuele, F. A. (2011). Reviewing Heinrich: Dislodging two myths from the practice of safety. *Professional Safety*, *56*(10), 52-61.

Maris, J. M.. (2017). An archival analysis of stall warning system effectiveness during airborne icing encounters [Doctoral dissertation, Embry-Riddle Aeronautical University]. Scholarly Commons. https://commons.erau.edu/cgi/viewcontent.cgi?article=1291&context=edt

Marsh, T. (2017). *A definitive guide to behavioral safety*. Routledge.

Marshall, P., Hirmas, A., & Singer, M. (2018). Heinrich's pyramid and occupational safety: A statistical validation methodology. *Safety Science, 101*, 180-189. doi:10.1016/j.ssci.2017.09.005

Maurino, D. (2017). Why SMS: An introduction and overview of safety management systems. *International Transport Forum Discussion Papers (2017-16)*.

May, P. J. (1992, October). Policy learning and failure. *Journal of Public Policy*, *12*(4), 331-354. doi:https://doi.org/10.1017/S0143814X00005602

McKinnon, R. C. (2017). *Risk-based, management-led, audit-driven, safety management systems*. CRC Press.

McPherson, G. (Ed.). (2001). *Applying and interpreting statistics: A comprehensive guide* (2nd ed.). Springer-Verlag.

Merry, A. F., & Henderson, B. (2017). Incident reporting, aviation and anaesthesia. *Anaesthesia and Intensive Care*, *45*(3), 291-294. doi:10.1177/0310057X1704500304

Midway, S., Robertson, M., Flinn, S., & Kaller, M. (2020). Comparing multiple comparisons: Practical guidance for choosing the best multiple comparisons test. *PeerJ*, *8*.doi:10.7717/peerj.10387

Mills, R. W. (2011). *Collaborating with industry to ensure regulatory oversight: The use of voluntary safety reporting programs by the Federal Aviation Administration*. ProQuest Dissertations Publishing.

Moore, S. M., Yorio, P. L., Haas, E. J., Bell, J. L., & Greenawald, L. A. (2020). Heinrich Revisited: a New Data-Driven Examination of the Safety Pyramid. *Mining, Metallurgy & Exploration*, *37*(6), 1857–1863. https://doi.org/10.1007/s42461-020-00263-0

Moss, S., Ryan, C., & Moss, J. (2016). Airline quality, load factors and performance. *Journal of Management Information and Decision Sciences*, *19*(1), 68.

NASA (2022a). Purpose Section. Aviation Safety Reporting System. Retrieved November 26, 2022, from https://asrs.arc.nasa.gov/overview/summary.html

NASA (2022b). ASRS Program Briefing. Retrieved 3 December, 2022, from https://asrs.arc.nasa.gov/docs/ASRS_ProgramBriefing.pdf

National Transportation Safety Board. (1995). *Aviation coding manual*. U.S. Department of Transportation.

National Transportation Safety Board. (2021). Aviation Accident Database & Synopses. Retrieved from https://www.ntsb.gov/_layouts/ntsb.aviation/index.aspx on March 1, 2021.

Noble, D. J., & Pronovost, P. J. (2010). Underreporting of patient safety incidents reduces health care's ability to quantify and accurately measure harm reduction. *Journal of Patient Safety*, *6*(4), 247-250. doi:10.1097/PTS.0b013e3181fd1697

Noort, M. C., Reader, T. W., Shorrock, S., & Kirwan, B. (2016). The relationship between national culture and safety culture: Implications for international safety culture assessments. *Journal of Occupational and Organizational Psychology*, *89*(3), 515-538. doi:10.1111/joop.12139

Okwera, P. (2016). A cost-benefit analysis of safety management system implementation in the transportation industry. ProQuest Dissertations Publishing.

Oldham, K., Stanton, J., Bilderbeck, M., & Spinetto, J. (2017). Case study of a voluntary aviation safety and environmental accreditation programme. *Safety Science, 96*, 41-51. doi:10.1016/j.ssci.2017.02.015

Oster, C. V., Jr., Strong, J. S., & Zorn, C. K. (2001). Improving air safety: Long-term challenges. *Issues in Science and Technology*, *17*(2), 77-83.

Patankar, M. S., & ebrary, I. (2012). *Safety culture: Building and sustaining a cultural change in aviation and healthcare*. Ashgate Pub. Co.

Petitt, K. (2017). SMS, safety culture, and the four pillars of safety applied to airline pilot training: NextGen demands to improve safety. *International Journal of Aviation Systems, Operations and Training*, *4*(2), 45-61. doi:10.4018/IJASOT.2017070104

Rao, A. H., & Marais, K. (2020). A state-based approach to modeling general aviation accidents. *Reliability Engineering & System Safety*, *193*.doi:10.1016/j.ress.2019.106670

Reason, J. T. (2016). *Managing the risks of organizational accidents*. Taylor and Francis. doi:10.4324/9781315543543

Schnittker, R., Marshall, S. D., & Berecki-Gisolf, J. (2020). Patient and surgery factors associated with the incidence of failed and difficult intubation. *Anaesthesia*, *75*(6), 756-766. doi:10.1111/anae.14997

Spigelman, A. D., & Swan, J. (2005). Review of the Australian incident monitoring system. *ANZ Journal of Surgery*, *75*(8), 657-661. Doi:10.1111/j.1445-2197.2005.03482.x

Steckel, R. J. (2014). Developing and establishing the validity and reliability of the perceptions toward Aviation Safety Action Program (ASAP) and Line Operations Safety Audit (LOSA) questionnaires. ProQuest Dissertations Publishing

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). Routledge. doi:10.4324/9780203843130

Stolzer, A. J. (2017). *Safety management systems in aviation* (1st ed.). Taylor and Francis. doi:10.4324/9781351149648

Stolzer, A. J., Friend, M. A., Truong, D., Tuccio, W. A., & Aguiar, M. (2018). Measuring and evaluating safety management system effectiveness using data envelopment analysis. *Safety Science*, *104*, 55-69. doi:10.1016/j.ssci.2017.12.037

Stolzer, A. J., Halford, C. D., & Goglia, J. J. (2018). *Safety management systems in aviation*. Routledge.

Stolzer, A. J., Halford, C. D., & Goglia, J. J. (2011). *Implementing safety management systems in aviation.* Ashgate. Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson Education, Inc./Allyn & Bacon.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. SAGE.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using Multivariate Statistics (Vol. 5)*. Boston, MA: Pearson.

Tekles, N., Chongvisal, J., Xargay, E., Choe, R., Talleur, D. A., Hovakimyan, N., & Belcastro, C. M. (2017). Design of a flight envelope protection system for NASA's transport class model. *Journal of Guidance, Control, and Dynamics*, 40(4), 863-877. doi:10.2514/1.G001728

Tiller, L. N., & Bliss, J. P. (2017). Categorization of near-collision close calls reported to the aviation safety reporting system. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *61*(1), 1866-1870. doi:10.1177/1541931213601947

Truong, D. (2022). Predicting the Impact of COVID-19 on Air Transportation Volumes. *Scholarly Commons.* Embry-Riddle Aeronautical University. https://commons.erau.edu/publication/1860/

Velazquez, J. (2018). The presence of behavioral traps in U.S. airline accidents: A qualitative analysis. *Safety*, *4*(1), 2. doi:10.3390/safety4010002

Wang, Z., Lu, Y., & Fu, S. (2020). A multidimensional method for pilot workload assessment and diagnosis. *Aerospace Medicine and Human Performance*, *91*(12), 932-939. doi:10.3357/AMHP.5592.2020

Warne, R. (2014). A primer on multivariate analysis of variance (MANOVA) for behavioral scientists. *Practical Assessment, Research & Evaluation*, *19*, 17.

Wilborn, J. & Foster, J. (n.d.). Defining Commercial Transport Loss-of-Control: A Quantitative Approach. In AIAA Atmospheric Flight Mechanics Conference and Exhibit. https://doi.org/10.2514/6.2004-4811

Wright, L. B. (2002). The analysis of UK railway accidents and incidents: a comparison of their causal patterns. ProQuest Dissertations Publishing.

Wilke, S., Majumdar, A., & Ochieng, W. Y. (2014). A framework for assessing the quality of aviation safety databases. *Safety Science*, *63*, 133-145. doi:10.1016/j.ssci.2013.11.005

Yesilbas, V. (2014). The relationship among HFACS levels and analysis of human factors in unmanned and manned air vehicles. ProQuest Dissertations Publishing.Yorio, P. L., & Moore, S. M. (2018). Examining factors that influence the existence of Heinrich's safety triangle using site-specific H&S data from more than 25,000 establishments. *Risk Analysis*, *38*(4), 839-852. doi:10.1111/risa.12869

Yorio, P. L., & Moore, S. M. (2018). Examining Factors that Influence the Existence of Heinrich's Safety Triangle Using Site-Specific H&S Data from More than 25,000 Establishments. *Risk Analysis*, *38*(4), 839–852. https://doi.org/10.1111/risa.12869

Appendix A

ASRS Related Research Dissertations

Table A1

ASRS-Related Research Dissertations, 2015–2019

Author	Year	Title	Brief Summary
Maris, John Michael	2017	An archival analysis of stall warning system effectiveness during airborne icing encounters	Used 132 ASRS reports and NASA's accident databases to create a combined Bayes' theorem with signal detection theory and binary logistic regression model to provide a high-reliability stall warning system for icing conditions called Conservative Icing Response Bias (CIRB).
Irwin, William J.	2017	Airline pilot situation awareness models: Proving a framework for meta-cognition, reflection, and education	Used grounded theory methods to develop a pilot situation awareness model from an initial sample of 48 ASRS report narrative descriptions from a population of 433 reports. Latent Semantic Analysis was then used for report sampling to augment the initial sample.
Campbell, Denado M.	2015	An assessment of predominant causal factors of pilot deviations that contribute to runway incursions	A qualitative study to identify predominant causal factors of pilot deviations in runway incursions over a two-year (2013–14) period based on coding ASRS reports. Coding was done based on previous research on runway safety conducted by the Aircraft Owners and Pilots Association (AOPA).
Kenyi, Likambo	2019	General aviation accident modeling and causal determination of pilot loss of aircraft control	Response to a 2018 NTSB petition regarding pilot induced LOC-I (PLOC-I) events in general aviation. General aviation-related ASRS reports were analyzed to validate NTSB PLOC-I predictors.

Appendix B

Examples of Serious Incidents from ICAO Annex 13 (ICAO, 2001)

Serious incident. An incident involving circumstances indicating that an accident nearly occurred.

2. The incidents listed are typical examples of incidents that are likely to be serious incidents. The list is not exhaustive and only serves as guidance to the definition of serious incident.

Near collisions requiring an avoidance manoeuvre to avoid a collision or an unsafe situation or when an avoidance action would have been appropriate.

Controlled flight into terrain only marginally avoided.

Aborted take-offs on a closed or engaged runway.

Take-offs from a closed or engaged runway with marginal separation from obstacle(s).

Landings or attempted landings on a closed or engaged runway.

Gross failures to achieve predicted performance during take-off or initial climb.
Fires and smoke in the passenger compartment, in cargo compartments or engine fires, even though such fires were extinguished by the use of extinguishing agents.

Events requiring the emergency use of oxygen by the flight crew.

Aircraft structural failures or engine disintegrations not classified as an accident.

Multiple malfunctions of one or more aircraft systems seriously affecting the operation of the aircraft.

Flight crew incapacitation in flight.

Fuel quantity requiring the declaration of an emergency by the pilot.

Take-off or landing incidents. Incidents such as under-shooting, overrunning or running off the side of runways.

System failures, weather phenomena, operations outside the approved flight envelope or other occurrences which could have caused difficulties controlling the aircraft.

Failures of more than one system in a redundancy system mandatory for flight guidance and navigation.

Appendix C

Global Safety Reporting Programs Overview

Table C1

VSR Programs in the United Kingdom, Australia, Hong Kong SAR, and New Zealand

Component	United Kingdom	Australia	Hong Kong SAR	New Zealand
Type of aviation safety reporting	Internal Safety Report [a], Occurrence Report, Whistleblower report, Chirp Report*. The Voluntary Safety Report aims to report occurrence and hazards. Connectivity to the European Central Repository. Each aviation organization is required to establish a VSR. Each aviation organization and member state shall establish a VSR for occurrence not fulfilling MOR criteria or potential hazards. Confidential Human Factors Incident Reporting Programme	MORs, ASRS, and REPCONs Aviation accident or incident notification - mandatory occurrence notification system required by the *Transport Safety Investigation Act 2003* for Immediately Reportable Matters or Routine Reportable Matters. These reports of accidents and incidents must be made to the Executive Director of Transport Safety Investigation through the ATSB's mandatory open reporting scheme. Administered by the Air Transport Safety Bureau (ATSB) Aviation self reporting - Under the ASRS*, the holder of a Civil Aviation Authorization may report a reportable contravention committed by the holder.	Mandatory Occurrence Report [a] Operators voluntary safety reporting system as part of the hazard identification element of the SMS.	Mandatory reporting on incidents and accidents to CAANZ [a] Centralized aviation reporting platform for mandatory and voluntary* reports Under SMS: safety occurrence reporting · hazard reporting · confidential reporting system

Component	United Kingdom	Australia	Hong Kong SAR	New Zealand
		Mandatory aviation accident or incident notification REPCON – Aviation Confidential Reporting Scheme		
Regulatory protections offered by VSR	Regulation (EU) No 376/2014 on the reporting, analysis, and follow-up of occurrences in civil aviation covering mandatory and voluntary safety reporting. Effective 15 Nov 2015. Originator shall not be penalized for reporting legal infringements or raising a report.	ASRS Reporters submitting eligible reports can claim protection from administrative action by CASA, in accordance with section 30DO of the *Civil Aviation Act 1988*, once every five years. Originator identity will be kept confidential in accordance with Division 3C of the *Civil Aviation Amendment Act 2003* and Division 13.K.1 of Subpart 13.K of the *Civil Aviation Safety Regulations 1998*.	CAD712 mentioned non-punitive (Just Culture) policy. CAD712: Hazards may be identified from the organization's reactive, proactive, and predictive processes. This should include the company's voluntary reporting system, audits and surveys, accident/incident reports as well as industry incident/accident reports.	Advisory Circulars AC12-1 Mandatory occurrence notification and information and AC12-2 Incident investigation. Data privacy protected by Privacy Act 1993 and the Official Information Act 1982. Rule Part 100 Safety Management that contains safety reporting process in service providers. AC 100-1 mentioned non-punitive safety reporting policy (Just Culture)

[a] denotes voluntary in nature

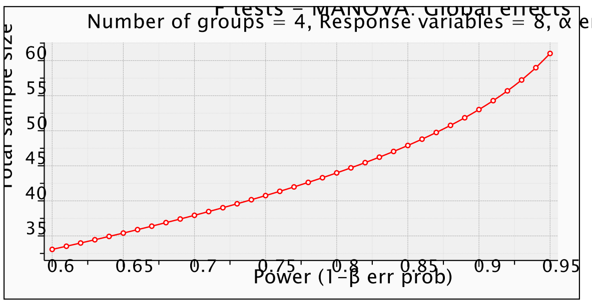
Appendix D

GPower® calculations on the MANOVA and Linear Regression Sample Size Requirement

Figure D1

GPower® Graph Showing MANOVA Sample Size Calculation

MANOVA:



F tests – MANOVA: Global effects

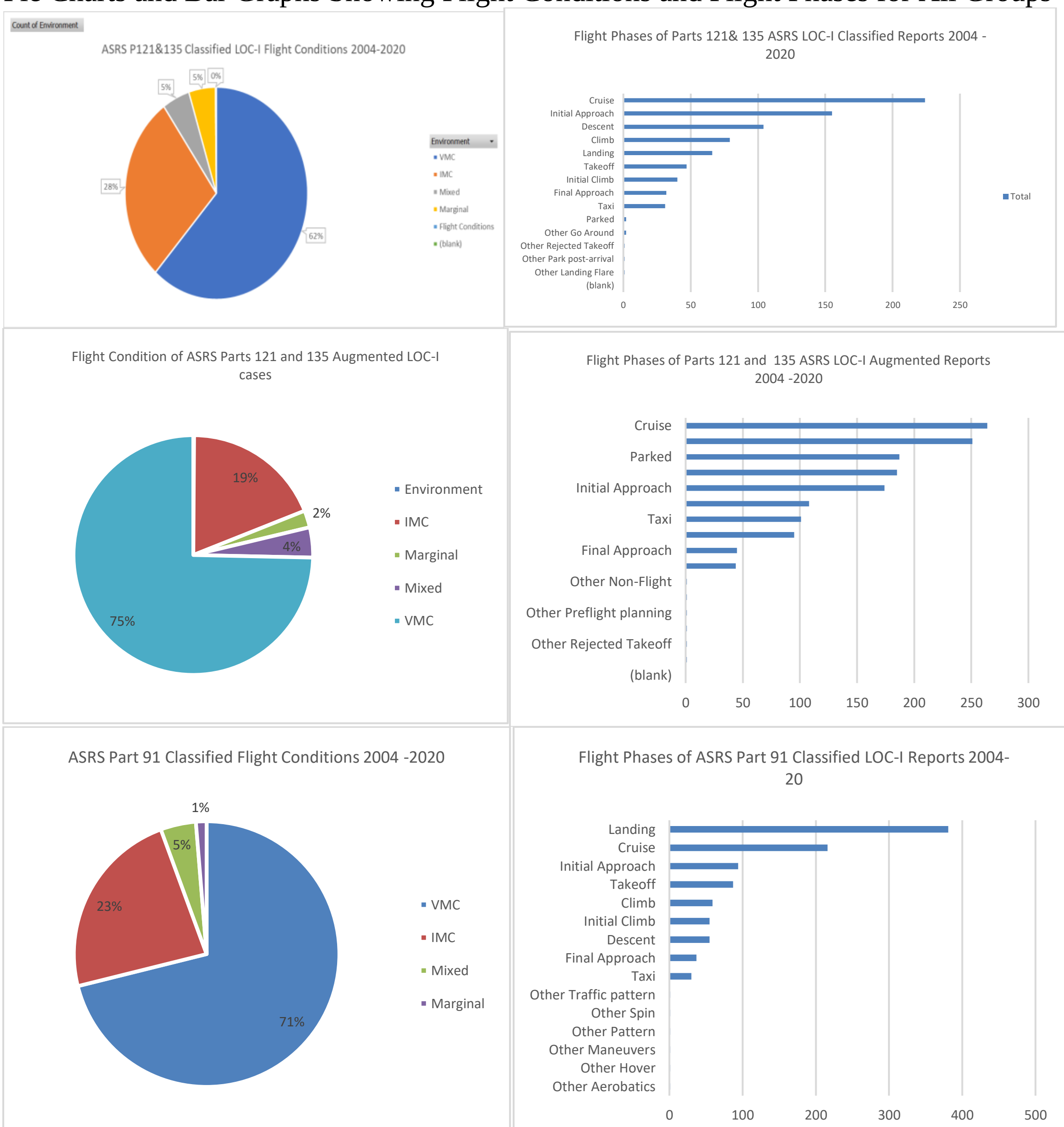
Options:	Pillai V, O'Brien-Shieh Algorithm		
Analysis:	A priori: Compute required sample size		
Input:	Effect size $f^2(V)$	=	0.2
	α err prob	=	0.05
	Power (1–β err prob)	=	0.8
	Number of groups	=	4
	Response variables	=	8
Output:	Noncentrality parameter λ	=	26.4000000
	Critical F	=	1.6214852
	Numerator df	=	24.0000000
	Denominator df	=	105
	Total sample size	=	44
	Actual power	=	0.8001268
	Pillai V	=	0.5000000

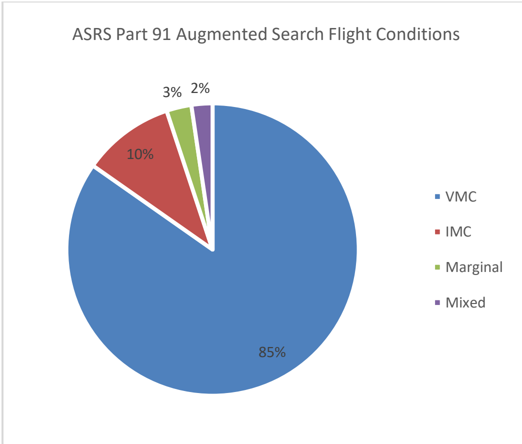
Appendix E

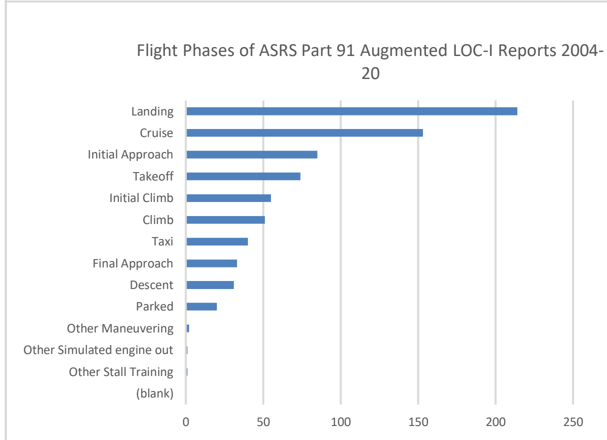
Demographics Analyses on the ASRS and NTSB Datasets

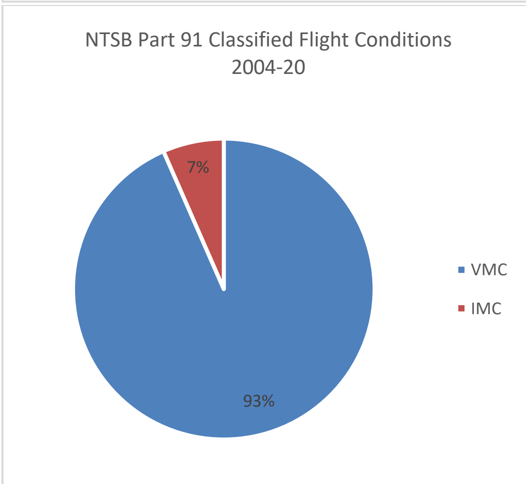
Figure E1

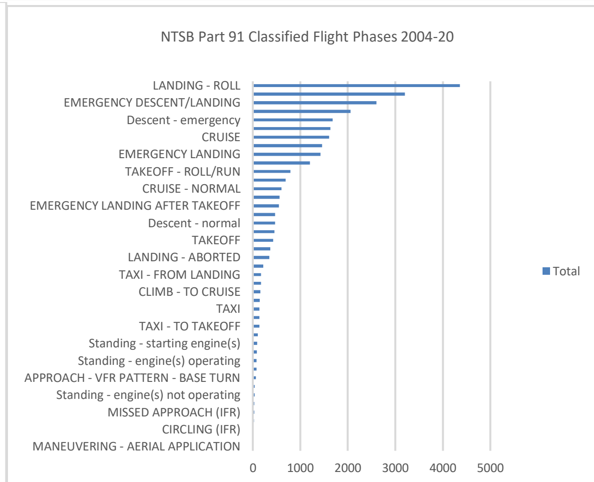
Pie Charts and Bar Graphs Showing Flight Conditions and Flight Phases for All Groups

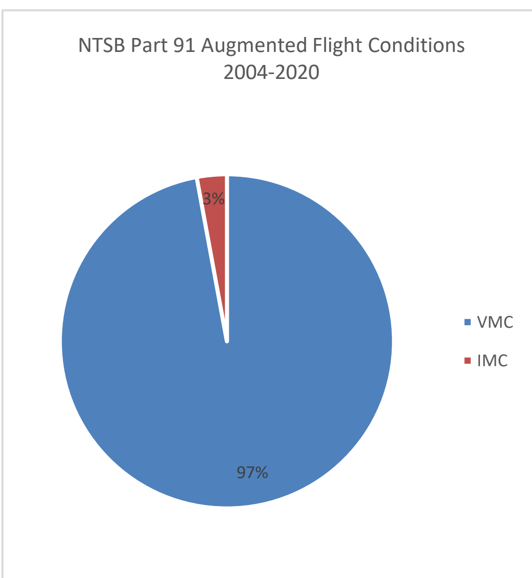


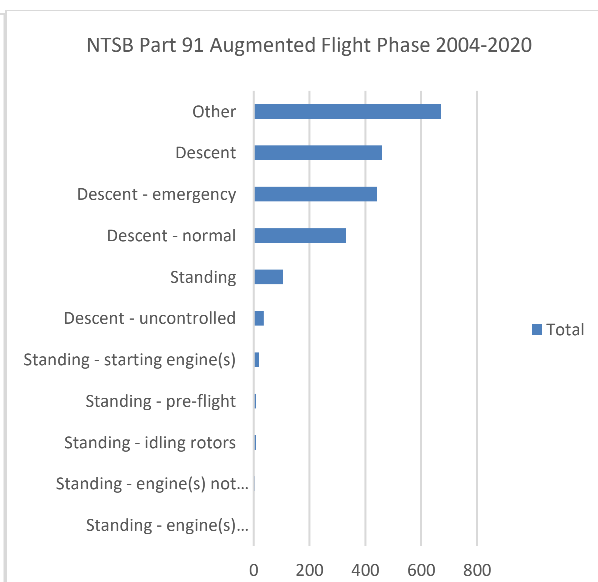

ASRS Part 91 Augmented Search Flight Conditions
3%
2%
10%
85%
VMC
IMC
Marginal
Mixed

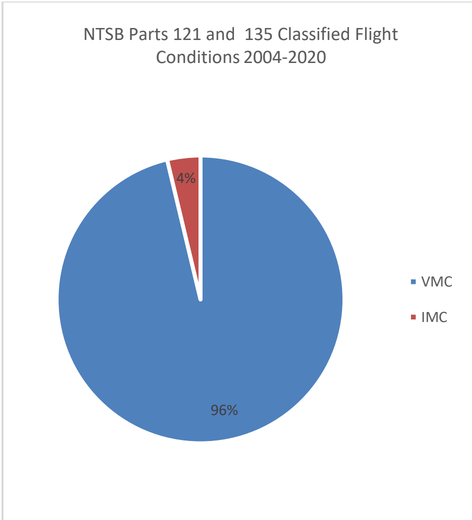

Flight Phases of ASRS Part 91 Augmented LOC-I Reports 2004-20
Landing
Cruise
Initial Approach
Takeoff
Initial Climb
Climb
Taxi
Final Approach
Descent
Parked
Other Maneuvering
Other Simulated engine out
Other Stall Training
(blank)
0
50
100
150
200
250

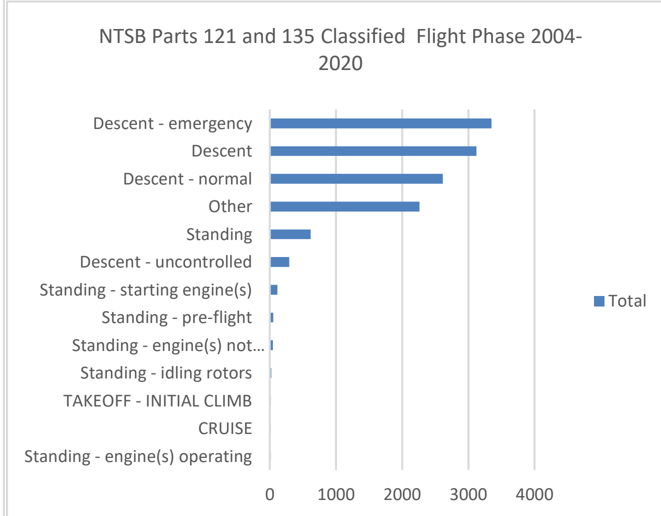

NTSB Part 91 Classified Flight Conditions 2004-20
7%
93%
VMC
IMC

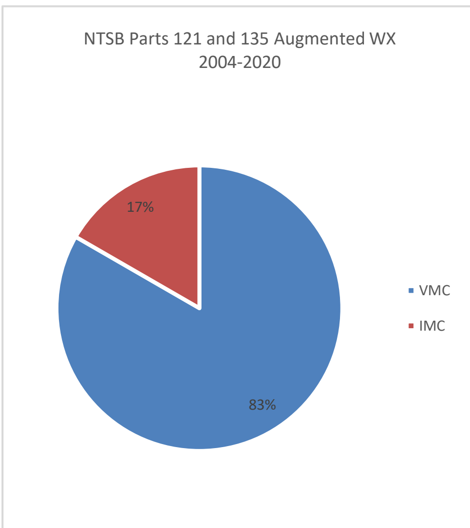

NTSB Part 91 Classified Flight Phases 2004-20
LANDING - ROLL
EMERGENCY DESCENT/LANDING
Descent - emergency
CRUISE
EMERGENCY LANDING
TAKEOFF - ROLL/RUN
CRUISE - NORMAL
EMERGENCY LANDING AFTER TAKEOFF
Descent - normal
TAKEOFF
LANDING - ABORTED
TAXI - FROM LANDING
CLIMB - TO CRUISE
TAXI
TAXI - TO TAKEOFF
Standing - starting engine(s)
Standing - engine(s) operating
APPROACH - VFR PATTERN - BASE TURN
Standing - engine(s) not operating
MISSED APPROACH (IFR)
CIRCLING (IFR)
MANEUVERING - AERIAL APPLICATION
0
1000
2000
3000
4000
5000
Total


NTSB Part 91 Augmented Flight Conditions 2004-2020
3%
97%
VMC
IMC


NTSB Part 91 Augmented Flight Phase 2004-2020
Other
Descent
Descent - emergency
Descent - normal
Standing
Descent - uncontrolled
Standing - starting engine(s)
Standing - pre-flight
Standing - idling rotors
Standing - engine(s) not...
Standing - engine(s)...
0
200
400
600
800
Total


NTSB Parts 121 and 135 Classified Flight Conditions 2004-2020
4%
96%
VMC
IMC


NTSB Parts 121 and 135 Classified Flight Phase 2004-2020
Descent - emergency
Descent
Descent - normal
Other
Standing
Descent - uncontrolled
Standing - starting engine(s)
Standing - pre-flight
Standing - engine(s) not...
Standing - idling rotors
TAKEOFF - INITIAL CLIMB
CRUISE
Standing - engine(s) operating
0
1000
2000
3000
4000
Total


NTSB Parts 121 and 135 Augmented WX 2004-2020
17%
83%
VMC
IMC

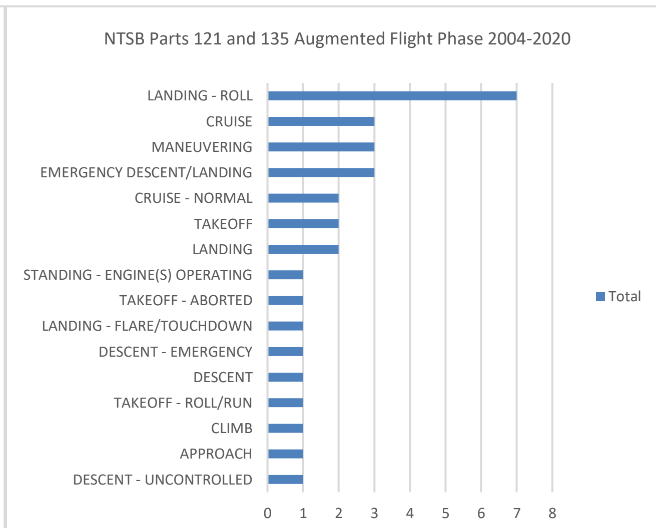

NTSB Parts 121 and 135 Augmented Flight Phase 2004-2020
LANDING - ROLL
CRUISE
MANEUVERING
EMERGENCY DESCENT/LANDING
CRUISE - NORMAL
TAKEOFF
LANDING
STANDING - ENGINE(S) OPERATING
TAKEOFF - ABORTED
LANDING - FLARE/TOUCHDOWN
DESCENT - EMERGENCY
DESCENT
TAKEOFF - ROLL/RUN
CLIMB
APPROACH
DESCENT - UNCONTROLLED
0
1
2
3
4
5
6
7
8
Total

Table E1

Top Five Flight Phases for Groups in IV

Classified				Augmented			
ASRS		NTSB		ASRS		NTSB	
Part 91	Parts 121 and 135	Part 91	Parts 121 and 135	Part 91	Parts 121 and 135	Part 91	Parts 121 and 135
Landing	Descent	Emergency Landing	Standing	Landing	Parked	Standing	Maneuvering
Cruise	Cruise	Cruise	Descent	Cruise	Cruise	Descent	Cruise
Initial Approach	Initial Approach	Descent – emergency	Descent – emergency	Initial Approach	Initial Approach	Descent – emergency	Descent – emergency
Climb	Climb	Descent / Landing	Descent - normal	Climb	Climb	Descent - normal	Landing - Roll
Takeoff	Landing	Landing-Roll	Other	Takeoff	Takeoff	Other	Takeoff

Appendix F

Biographies of the Subject Matter Experts in the Taxonomies Mapping Exercise

Mr. Thian Chow Vi (CV)

Mr. Thian Chow Vi (CV) is currently the Head of Standards and Process Improvement at Teleport by AirAsia, the cargo and logistics subsidiary of the AirAsia Group. This role looks after safety and risk management systems, as well as operational and corporate processes for the company. Prior to that, he was Senior Manager for AirAsia Group Safety, a department that oversees Safety Management for all nine AirAsia airlines in the AirAsia Group including flight, cabin, ground and corporate safety functions. He joined AirAsia X as Safety Risk Manager in 2015 before moving to the wider group function in 2017. Before that, he was Senior Associate of Technical Affairs at the Association of Asia Pacific Airlines (AAPA) for four years where he was the secretary for various committees and working groups, including the Flight Operations Safety Working Group (FOSWG). CV graduated with distinction from RMIT University, Australia in 2010 with a Bachelor of Science (Aviation). He obtained a Master's in Aviation Management in 2014 from RMIT University as well. CV holds a Private Pilot License (PPL) and in his spare time likes to hop around the islands of Malaysia in a single engine aircraft.

Capt. Peter Lawrie

Capt. Peter Lawrie joined the working aviation community in General Aviation in Remote Regions of Australia in 1994, progressing up through to the Regional Airlines.

In 2005, Capt. Lawrie made the first tentative steps in joining the International Aviation Community, as a Direct Entry Captain in startup International Airlines, in Hong

Kong, India and Australia. With a strong background also in IT, translated well into becoming involved in establishing each of the Airlines FOQA / FDM Programs.

2010 afforded the opportunity for Capt. Lawrie to join the International Corporate Aviation Community as a Captain operating multiple Gulfstream Models, and additionally tasked with Flight Data Analysis, aiding to adapt FDM programs tuned to the special needs of Corporate Aviation. Recognized by the NBAA for Contributions to Safety for 9 years continuous service. Currently operating International Long Haul flying and having served in FDM / FOQA Programs continuously for the past 18 years.

Capt. Denis Portier

Capt. Portier started his aviation career in 1990 (Jet Express dba Trans World Express). He has extensive training and managerial background with FAA, EASA & HK ATPL's (8 type ratings: CA-212 / LR-JET / G-IV / G-V / G-VI / BD-700 / CL-65 / B-737). He has been Post-Holder Training (Manager Flight Training - MFT & Chief Training Captain - CTC): TCE (FAA), TRE (EASA), AEX (HK-CAD), Post-Holder Flight Operations (Chief Pilot).

Capt. Portier graduated in Marketing (University Institute of Technology - 1985) and International Business (Ecole Supérieure de Commerce International - 1987). He obtained a certificate in safety management systems (Southern California Safety Institute - SCSI) in Aug. 2006. He has been Line Training Captain - LTC (Murray Aviation 1995) / Sim Instructor & Line Check Airman (Midway Airlines 2000-2003) / Training Centre Evaluator (TCE/SFE/TRE) CAE-DXB 2003-2005 / MFT-CTC (Metrojet 2013-2015) / MFO-CP (Gama Aviation - Asia 2017-2019) / LTC (Global Jet 2019-2020).

Capt. François Lassale, MSc FRAeS

Capt. Francois Lassale is the Chief Executive Officer for HeliSGI, an organization providing rotary wing and fixed wing services. Before joining HeliSGI, Francois was the Chief Operating Officer for HeliOffshore, a safety focused organization working with the global offshore helicopter transport industry and Managing Director for a firm in the USA bringing turnkey solutions to the fixed wing industry.

Francois has been in aviation for thirty years with a military background in the South African Army and Royal Air Force, flying both fixed wing and rotary wing. Since leaving the military he flew for an airline, freight, and was VIP and Head of State operations. Francois has been an instructor of TRI, TRE and CRMI. He served on the Flight Safety Foundation's Business Aviation Board for thirteen years. He currently serves as Vice Chairman of the European Helicopter Association, Vice Chairman of the International Pilot Training Association, and Vice Chairman of the International Association of Aeronautical Flight Auditors. He is a certified IS-BAO auditor, a Fellow with the Royal Aeronautical Society, and is an Upper Freeman with the Honorable Company of Air Pilots.

Capt. Richard Boswell

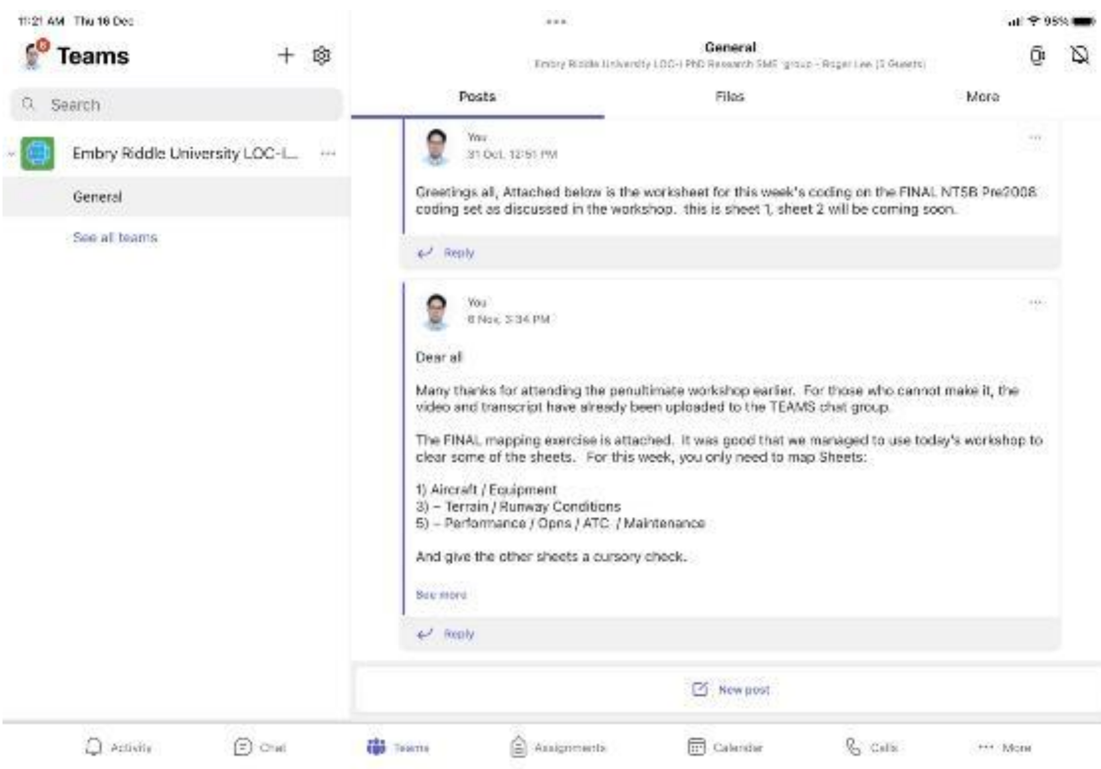
Capt. Richard Boswell joined the aviation industry in 1985 as a military pilot flying both fixed wing aircraft and helicopters. On completion of military service, he moved into commercial aviation and has extensive experience as an airline/corporate pilot and a HEMS, police, charter and utility helicopter pilot. He is an instructor and examiner and has over 15 years management experience in Europe, Africa and Asia as Accountable Manager, Safety Manager and Head of Training. He remains an active pilot and is Managing Director of Spotlight on Safety, an international consultancy firm specializing in enhancing aviation safety.

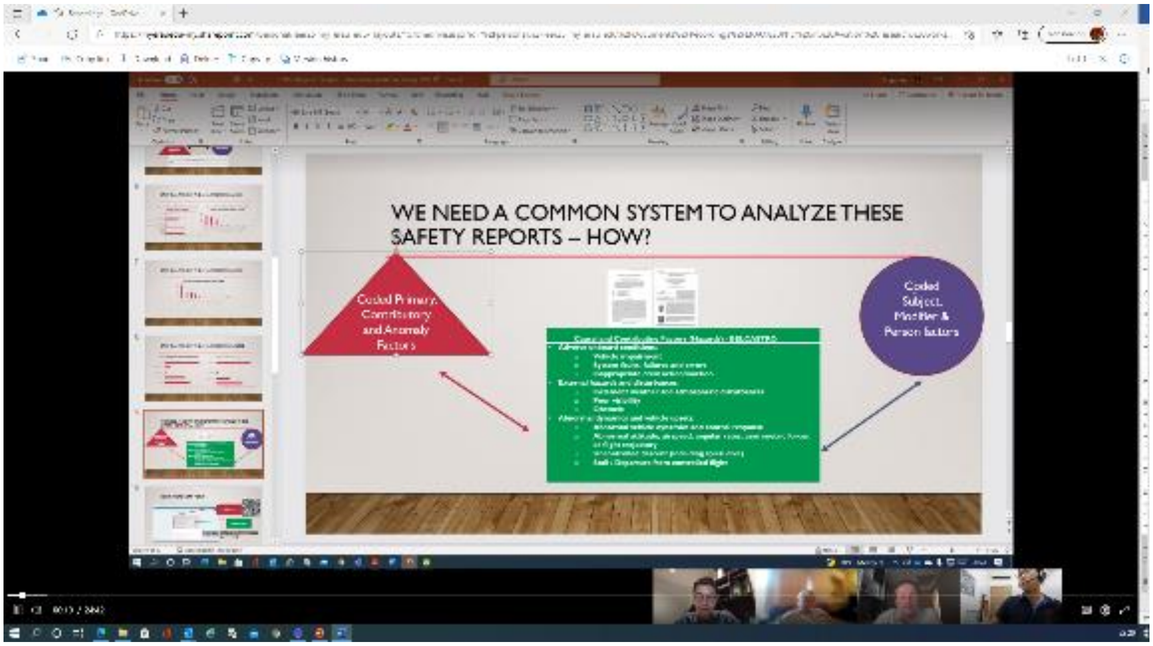
Appendix G

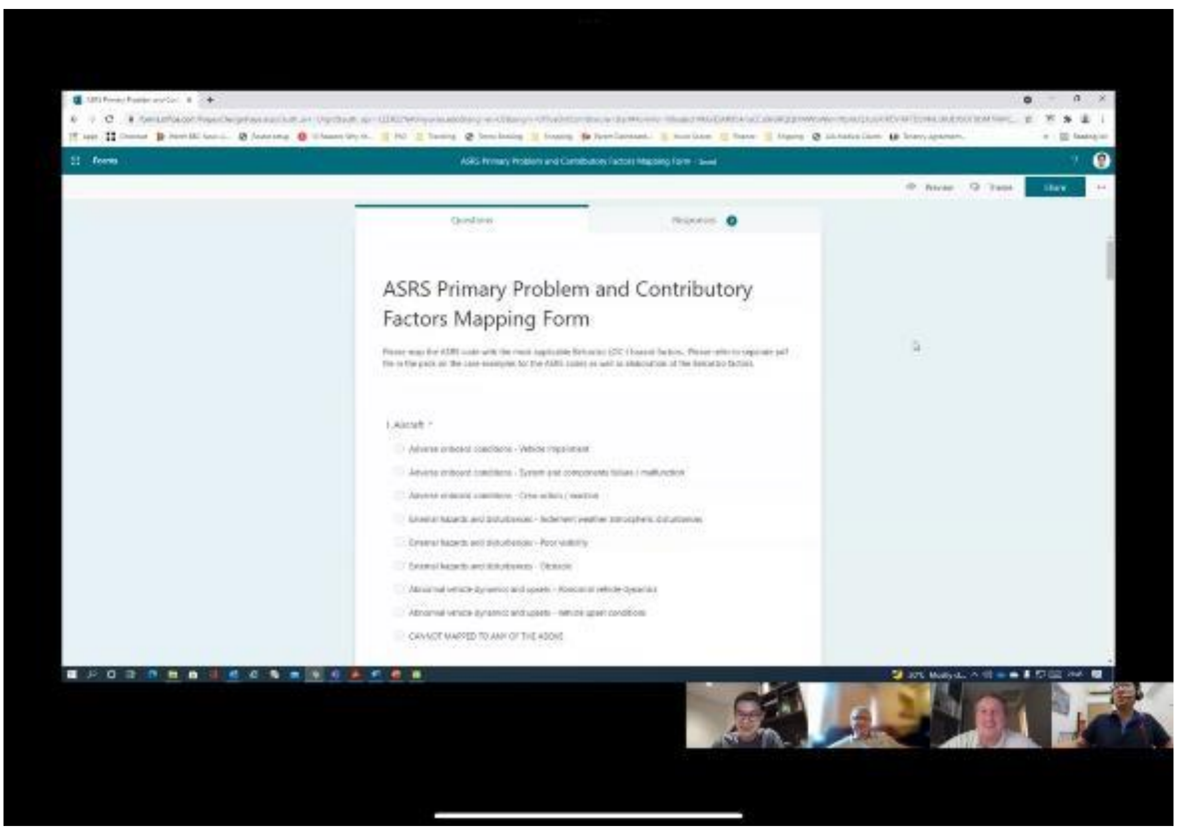
Online Workshop, Microsoft Teams® Group, and Microsoft Forms® for Taxonomies Mapping with ASRS and NTSB LOC-I Codes

Figure G1

Screenshots from Online Workshops with Subject Matter Experts




WE NEED A COMMON SYSTEM TO ANALYZE THESE SAFETY REPORTS – HOW?
Coded Primary, Contributory and Anomaly Factors
Coded Subject, Modifier & Person factors


ASRS Primary Problem and Contributory Factors Mapping Form

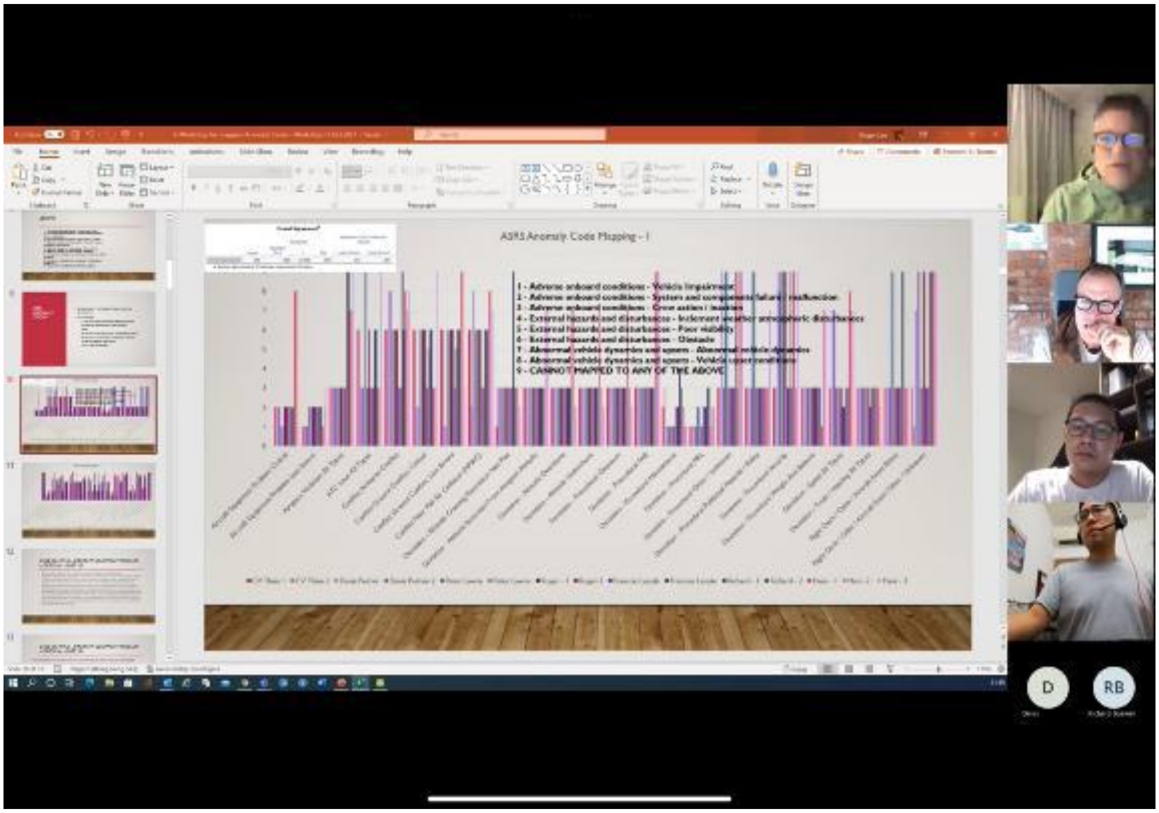


Table G2

Taxonomy Mapping Process Underwent by the SME Panel

Date	Purpose	Medium	Output
18 September 2021	Induction Video	Video File made by facilitator	Induction on the background of this research
25 September, 2021	Induction workshop	Online workshop	Equipped with the knowledge and process of the taxonomy mapping
25 September, 2021	Microsoft Forms	SME Mapping 1- ASRS	Platform to perform mapping
2 October, 2021	ASRS Primary Problems Mapping discussion - 1	Online workshop	ASRS -1 mapping reviewed and discussed
2 October, 2021	Microsoft Forms	SME Mapping 2- ASRS	Platform to perform mapping
9 October, 2021	ASRS Contributory Factors / Situation	Online workshop	ASRS -2 mapping reviewed

	Mapping discussion - 2		
16 October, 2021	ASRS Contributory Factors / Situation Mapping discussion - 3	Online Workshop	ASRS-3 Contributory Factors / Situation mapping discussed and concluded with Kappa >0.7.
23 October, 2021	ASRS Anomaly Codes mapping	Online Workshop	ASRS -3 Anomaly code mapped with Kappa >0.7
16 October, 2021	Inducting NTSB mapping	Video File made by facilitator and Microsoft Excel mapping template for submission	Induction on the NTSB codes mapping
31October, 2021	NTSB Categories mapping	Online Workshop	NTSB Subcategories mapped with Kappa >0.7
6 November, 2021	NTSB Subcategories mapping	Online Workshop	NTSB Subcategories mapped with Kappa >0.7
14 November, 2021	NTSB Modifier Coding Finalization	Online Workshop	NTSB Modifiers mapped with Kappa > 0.7

Appendix H

ASRS Result from the Taxonomy Mapping Exercises

Table H1

ASRS Result from the Taxonomy Mapping Exercises, 2015–2019

Primary and Contributory Factors Mapped	DV Code	Anomaly Mapped	DV Code
Aircraft[a]	2	Flight Deck / Cabin / Aircraft Event Passenger Misconduct	9
Airport	9	Flight Deck / Cabin / Aircraft Event Smoke / Fire / Fumes / Odor	9
Airspace Structure	9	Ground Event / Encounter Gear Up Landing	9
ATC Equipment / Nav Facility / Buildings	6	Ground Event / Encounter Ground Strike - Aircraft	3
Chart or Publication	9	Ground Event / Encounter Loss Of Aircraft Control	3
Company Policy*	3	Ground Event / Encounter Object	9
Environment - Non Weather Related*	6	Ground Event / Encounter Other / Unknown	9
Equipment / Tooling	2	Ground Event / Encounter Person / Animal / Bird	9
Human Factors*	3	Ground Event / Encounter Vehicle	9
Incorrect / Not Installed / Unavailable Part*	1	Ground Excursion Runway	3
Logbook Entry	3	Ground Excursion Taxiway	9
Manuals	9	Ground Incursion Runway	3
MEL*	1	Ground Incursion Taxiway	9
Procedure*	3	Inflight Event / Encounter Bird / Animal	2
Staffing	9	Inflight Event / Encounter CFTT / CFIT	3
Weather*	4	Inflight Event / Encounter Fuel Issue	1
		Inflight Event / Encounter Loss Of Aircraft Control	7
		Inflight Event / Encounter Object	6
		Inflight Event / Encounter Other / Unknown	4
		Inflight Event / Encounter Unstabilized Approach	3
		Inflight Event / Encounter VFR In IMC	4
		Inflight Event / Encounter Wake Vortex Encounter	8
		Inflight Event / Encounter Weather / Turbulence	4

[a] Code 9 denotes cannot be mapped with Belcastro et al. (2018) factor.

Appendix I

NTSB Result from the Taxonomy Mapping Exercises

Table I1

Taxonomy Mapping Exercise Results

Post 2008

Subject Code	DV Code	SubCat Code	DV Code
AIRCRAFT	2	Aircraft handling/service	1
PERSONNEL	3	Aircraft systems	2
ENVIRONMENTAL	9	Aircraft structures	9
ORGANIZATIONAL	9	Aircraft propeller/rotor	2
		Aircraft power plant	2
		Aircraft oper/perf/capability	3
		Fluids/misc hardware	2
		Environment: Operating environment	3
		Physical environment	6
		Conditions/weather/phenomena	4
		Task environment	3
		Organizational : Development	9
		Management	9
		Support/oversight/monitoring	9
		Personnel: Physical	3
		Psychological	3
		Experience/knowledge	3
		Action/decision	3
		Miscellaneous	9
		Task performance	3

Pre 2008

Subject Code	DV Code	Modifier Code	DV Code
Landing gear, nose gear	1	Dark night	5
Landing gear, nose gear assembly	1	Night	5
Miscellaneous, bolt/nut/fastener/clamp/spring	1	Other	9
Landing gear, main gear strut	1	Dusk	5

Wing, spar	1	Sunglare	5
Airframe	1	Crosswind	4
Landing gear, gear locking mechanism	1	Gusts	4
Wing	1	Tailwind	4
Flight control surfaces/attachments	1	Low ceiling	4
Flight control, rudder	1	Clouds	4
1engine	2	High density altitude	4
Engine assembly, cylinder	1	Carburetor icing conditions	4
Fuel system, line	2	Fog	4
Ignition system, magneto	1	Downdraft	4
Engine assembly, bearing	1	Other	4
Lubricating system, oil filler cap	1	High wind	4
Flight/navigation instruments, airspeed indicator	2	Icing conditions	4
Flight/navigation instruments, attitude gyro	2	Thunderstorm	4
Autopilot/flight director, transmitter (autopilot)	2	Turbulence, terrain induced	4
Vacuum system	2	Temperature, high	4
Reduction gear assembly, reduction gear bearing	1	Microburst/wet	4
Terrain condition	6	Mountain wave	4
Light condition	5	Rain	4
Fluid, fuel	1	Turbulence	4
Fluid, oil	2	Windshear	4
Aircraft performance, climb capability	1	Below approach/landing minimums	4
Object	6	Dust devil/whirlwind	4
Weather condition	4	Variable wind	4
Landing gear extension	1	Sudden windshift	4
Landing gear, normal brake system	1	no thermal lift	4
Carburetor heat	3	Unfavorable wind	4
Fuel supply	3	Snow	4
Fuel tank selector position	3	Obscuration	4
Raising of flaps	3	Not maintained	3
Propeller feathering	3	Improper	3
Rudder	3	Inadvertent	3
Landing gear retraction	3	Encountered	3
Elevator	1	Not performed	3
Mixture	3	Misjudged	3

Nosewheel steering	1	Not possible	3
Flight controls	3	Performed	3
Trim setting	3	Delayed	3
Emergency floats	9	Attempted	3
Throttle/power control	3	Selected	3
Flaps	3	Low	3
Aircraft control	3	Excessive	3
Airspeed	3	Not followed	3
Clearance	3	Not attained	3
Visual lookout	3	Continued	3
Ground loop/swerve	3	Intentional	3
Preflight planning/preparation	3	Initiated	3
Altitude/clearance	3	Exceeded	3
Stall/mush	3	Improper use of	3
Maintenance, installation	1	Not corrected	3
Proper touchdown point	3	Not obtained/maintained	3
Visual flight rules (VFR) flight into instrument meteorological	3	Not used	3
Go-around	3	Not obtained	3
Emergency procedure	3	Simulated	3
Precautionary landing	3	Not complied with	3
Distance/altitude	9	Not verified	3
Planning/decision	3	Incorrect	3
Refueling	3	Abrupt	3
Porpoise/pilot-induced oscillation	3	Not understood	3
Flight into adverse weather	3	Uncontrolled	3
Weather evaluation	3	Not recognized	3
Operation with known deficiencies in equipment	3	Not calculated	3
Checklist	3	Inadvertent activation	3
Aerobatics	3	Not selected	3
Proper glidepath	3	Diminished	3
Proper alignment	3	High	3
Maintenance, annual inspection	2	Activated	3
Wheels-up landing	3	Not issued	3
Instrument flight rules (IFR) procedure	3	Not successful	3
Maintenance, service bulletin/letter	2	Premature	3
Stall	3	Restricted	3
Stall/spin	3	Poor	3
Reason for occurrence undetermined	9	Not available	3

Directional control	3	Tree(s)	6
Compensation for wind conditions	3	Fence	6
Remedial action	3	Wire, transmission	6
In-flight planning/decision	3	Sign	6
Supervision	3	Vehicle	6
Aborted takeoff	3	Airport sign/marker	6
Aircraft weight and balance	3	Other	9
Go-around	3	Residence	6
Flight into known adverse weather	3	Aircraft parked/standing	6
Aircraft preflight	3	Pole	6
Procedures/directives	3	Wire, static	6
Unsuitable terrain or takeoff/landing/taxi area	3	Runway light	6
Climb	3	Building (nonresidential)	6
Maintenance	9	Fence post	6
Airspeed, minimum control speed with the critical engine inopera	3	Hangar/airport building	6
Lift-off	3	Taxiway light	6
Maneuver to avoid obstructions	3	Wall/barricade	6
Airspeed, reference (Vref)	3	Undetermined	6
In-flight weather avoidance assistance	3	Utility pole	6
Relinquishing of control	3	Bird(s)	1
Instructions, written/verbal	3	Aircraft moving on ground	6
Maintenance, inspection	1	Animal(s)	6
Descent	3	GROUND	6
Missed approach	3	Runway	9
Wake turbulence	4	None suitable	3
Ice/frost removal from aircraft	3	Ditch	3
Refueling	3	Water	4
Planned approach	3	Rough/uneven	8
Unstabilized Approach	3	Soft	8
Maintenance, service of aircraft/equipment	2	Mountainous/hilly	6
Maintenance, service bulletin/letter	2	Grass	9
Aircraft handling	3	High vegetation	6
Low pass	3	Open field	9
Rotation	3	Berm	6
Starting procedure	3	Snow covered	4

Spiral	3	Dirt bank/rising embankment	6
Procedures/directives	3	Crop	6
Loading of cargo	3	Snowbank	6
Altitude	3	Other	9
Aircraft service	9	Rising	6
Proper assistance	3	Wet	4
Proper descent rate	3	Muddy	8
Visual separation	3	Roadway/highway	6
Maintenance	2	Swampy	9
Flare	3	Drop-off/descending embankment	6
Design stress limits of aircraft	3	Rock(s)/boulder(s)	6
ATC clearance	3	Short runway/landing area	3
Procedure inadequate	9	Sand bar	6
Spatial disorientation	3	Water, glassy	4
Lack of certification	3	Loose gravel/sandy	9
Lack of total experience in type of aircraft	3	Uphill	6

Note. Code 9 denotes cannot be mapped with Belcastro et al. (2018) factor.

Appendix J

Codes Counts and Normalized Results from the Belcastro et al. (2018) Mapped Codes from ASRS and NTSB Databases for LOC

Table J1

Normalization Data for Parts 121 and 135 and P91 datasets from 2004 to 2020

Year	Part 121	Part 135	Part 91
2004	18882503	2455585	21565890
2005	19390029	2648915	19662170
2006	19263209	2544250	20220709
2007	19637322	2949394	19907774
2008	19126766	1975993	19154513
2009	17626832	1841583	17167888
2010	17750986	1827306	17851337
2011	17962965	1949840[a]	17568252*
2012	17722236	2072373	17285166
2013	17779641	2259169	16168807
2014	17742826	2472131	15988460
2015	17925780	2393048	16806585
2016	18294057	2410858	17690903
2017	18581388	2459228	17810052
2018	19288454	2777012	18336204
2019	19788411	2589781	19131417
2020	8898769	1398482*	17026961*

[a]extrapolated

Table J2

Normalized Results with Mapped Codes for Parts 121 and 135 and P91 datasets from 2004 to 2020

Parts 121&135 Dataset								
DV								
1	2	3	4	5	6	7	8	CODE_TYPE
4.68646E-08	1.41E-07	1.41E-07	7.5E-07	0	0	0	0	1
0	9.07E-08	9.07E-08	3.63E-07	0	0	4.54E-08	0	1
4.58559E-08	4.59E-08	4.59E-08	1.83E-07	9.17E-08	4.59E-08	4.59E-08	0	1
8.85476E-08	8.85E-08	0	1.77E-07	0	0	0	0	1
7.58195E-07	8.51E-05	6.35E-05	7.11E-07	0	2.75E-06	0	0	1
1.23277E-06	8.93E-05	6.09E-05	1.13E-06	0	1.28E-06	0	0	1
8.68309E-07	8.32E-05	6.23E-05	1.07E-06	0	1.69E-06	0	0	1
8.53722E-07	8.47E-05	6.5E-05	1.41E-06	0	8.54E-07	0	0	1
6.56744E-07	8.42E-05	5.99E-05	1.01E-06	0	1.41E-06	0	0	1
5.48935E-07	6.9E-05	5.16E-05	4.49E-07	0	1.5E-06	0	0	1
5.9362E-07	6.89E-05	4.91E-05	6.93E-07	0	1.88E-06	0	0	1
6.39801E-07	6.89E-05	4.51E-05	1.13E-06	0	1.48E-06	0	0	1
6.76168E-07	6.75E-05	4.7E-05	1.06E-06	0	1.74E-06	0	0	1
4.27744E-07	6.3E-05	4.51E-05	8.08E-07	0	1.66E-06	0	0	1
1.81279E-07	5.3E-05	4.11E-05	3.63E-07	0	1.45E-06	0	0	1
0	3.63E-05	2.96E-05	3.57E-07	0	9.83E-07	0	0	1
1.94227E-07	1.67E-05	1.51E-05	2.91E-07	0	9.71E-08	0	0	1
9.37291E-08	4.69E-08	2.34E-07	8.9E-07	0	9.37E-08	0	0	2
1.81497E-07	4.54E-08	1.36E-07	7.26E-07	0	4.54E-08	0	0	2
9.17117E-08	4.59E-08	4.59E-08	4.13E-07	9.17E-08	4.59E-08	4.59E-08	0	2
8.85476E-08	8.85E-08	1.77E-07	2.66E-07	0	8.85E-08	0	0	2
0	2.37E-06	5.69E-07	0	0	4.74E-08	0	0	2
0	1.75E-06	3.6E-07	0	0	5.14E-08	0	0	2
5.1077E-08	2.09E-06	5.11E-07	1.02E-07	0	0	0	0	2
0	2.26E-06	6.53E-07	5.02E-08	0	0	0	0	2
5.05188E-08	1.52E-06	5.05E-07	0	0	0	0	0	2
0	1.45E-06	7.98E-07	0	0	0	0	0	2
0	6.43E-07	4.95E-07	0	0	0	0	0	2
9.84309E-08	7.87E-07	1.48E-07	0	0	0	0	0	2
4.82977E-08	1.06E-06	3.38E-07	0	0	0	0	0	2
0	8.08E-07	3.33E-07	0	0	0	0	0	2
0	2.72E-07	9.06E-08	0	0	0	0	0	2
0	2.23E-07	1.34E-07	0	0	0	0	0	2
0	0	0	0	0	0	0	0	2
1.73399E-06	4.69E-06	2.3E-06	9.37E-08	0	0	2.34E-06	0	3
1.13436E-06	3.31E-06	2.5E-06	1.36E-07	0	1.81E-07	1.81E-06	1.36E-07	3
4.58559E-07	2.11E-06	3.9E-06	8.71E-07	0	1.24E-06	1.51E-06	1.83E-07	3
8.85476E-07	2.57E-06	5.58E-06	4.43E-07	0	8.85E-08	1.86E-06	0	3
6.16033E-07	2.75E-06	7.3E-06	4.88E-06	0	4.26E-07	3.27E-06	3.32E-07	3
2.56826E-07	8.73E-07	5.09E-06	1.8E-06	0	2.88E-06	2.52E-06	8.22E-07	3
0	4.09E-07	4.65E-06	1.58E-06	0	3.93E-06	2.81E-06	1.89E-06	3
0	0	3.62E-06	5.02E-08	0	2.51E-06	1.86E-06	1.76E-06	3
1.51556E-07	1.52E-07	6.72E-06	1.47E-06	0	2.27E-06	2.78E-06	1.52E-06	3
4.49128E-07	1.6E-06	9.68E-06	8.98E-07	0	8.98E-07	2.64E-06	4.49E-07	3
4.94683E-08	9.89E-08	8.76E-06	6.58E-06	0	4.95E-08	2.77E-06	4.95E-08	3
2.46077E-07	7.38E-07	4.82E-06	1.53E-06	0	1.33E-06	2.17E-06	1.23E-06	3
9.65954E-08	9.66E-08	5.41E-06	8.69E-07	0	1.55E-06	1.98E-06	1.11E-06	3
1.42581E-07	0	2.8E-06	4.9E-06	0	1.9E-07	1.66E-06	0	3
4.53197E-08	0	3.44E-06	6.48E-06	0	0	2.18E-06	0	3
2.23432E-07	9.38E-07	5.76E-06	8E-06	0	5.81E-07	2.99E-06	8.94E-08	3
1.94227E-07	4.86E-07	7.57E-06	7.19E-06	0	9.71E-08	2.91E-06	0	3
2.15577E-06	5.2E-06	1.62E-05	1.78E-06	0	1.45E-06	3.75E-07	4.69E-08	4
1.76959E-06	3.99E-06	7.76E-06	1.63E-06	0	8.62E-07	5.9E-07	9.07E-08	4
1.88009E-06	4.49E-06	9.08E-06	1.51E-06	0	9.63E-07	4.13E-07	4.59E-08	4
1.6824E-06	4.52E-06	8.9E-06	1.28E-06	0	1.02E-06	5.76E-07	3.1E-07	4
1.94287E-06	5.88E-06	1.28E-05	1.71E-06	0	1.52E-06	8.06E-07	3.79E-07	4
2.05461E-06	5.34E-06	9.66E-06	2.21E-06	0	1.18E-06	4.11E-07	3.08E-07	4
3.37108E-06	6.54E-06	1.09E-05	2.4E-06	0	1.02E-06	6.64E-07	4.6E-07	4
2.71182E-06	5.88E-06	1.09E-05	2.16E-06	0	5.52E-07	3.01E-07	1.51E-07	4
2.47542E-06	5.51E-06	1.03E-05	2.48E-06	0	1.06E-06	6.57E-07	3.54E-07	4
3.094E-06	7.14E-06	9.78E-06	2E-06	0	1.3E-06	4.99E-07	4.99E-07	4
1.33564E-06	4.16E-06	1.32E-05	3.07E-06	0	9.89E-07	8.41E-07	5.94E-07	4
7.87447E-07	4.82E-06	1.75E-05	2.9E-06	0	1.53E-06	7.38E-07	7.87E-07	4
8.21061E-07	4.4E-06	1.17E-05	2.46E-06	0	8.21E-07	4.83E-07	5.31E-07	4
9.03015E-07	2.9E-06	1.34E-05	2.52E-06	0	1.24E-06	6.18E-07	6.65E-07	4
8.15754E-07	3.99E-06	1.6E-05	3.99E-06	0	1.4E-06	9.06E-07	8.16E-07	4
1.56402E-06	5.23E-06	1.69E-05	3.17E-06	0	2.41E-06	9.83E-07	1.21E-06	4
1.65093E-06	4.18E-06	2.09E-05	4.18E-06	0	3.01E-06	1.07E-06	7.77E-07	4

Part 91 Dataset									
DV	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**CODE_TYPE**
2004	1.85E-05	5.56E-06	7.79E-06	1.03E-05	1.25E-06	5.11E-05	1.95E-06	3.57E-06	1
2005	2.08E-05	5.29E-06	9.36E-06	9.21E-06	1.22E-06	5.97E-05	2.24E-06	3.61E-06	1
2006	1.68E-05	5.69E-06	7.62E-06	9.5E-06	7.42E-07	5.37E-05	1.24E-06	3.07E-06	1
2007	1.84E-05	5.98E-06	8.74E-06	8.64E-06	1.26E-06	5.8E-05	1.86E-06	3.92E-06	1
2008	5.22E-08	1.48E-05	1.41E-05	1.04E-07	0	0	0	0	1
2009	5.82E-08	1.51E-05	1.37E-05	2.33E-07	0	0	0	0	1
2010	0	1.17E-05	1.3E-05	0	0	0	0	0	1
2011	1.71E-07	1.5E-05	1.53E-05	5.69E-08	0	5.69E-08	0	0	1
2012	1.16E-07	1.61E-05	1.54E-05	5.79E-08	0	0	0	0	1
2013	1.24E-07	1.43E-05	1.37E-05	0	0	0	0	0	1
2014	6.25E-08	1.51E-05	1.63E-05	0	0	0	0	0	1
2015	5.95E-08	1.37E-05	1.36E-05	5.95E-08	0	5.95E-08	0	0	1
2016	5.65E-08	1.14E-05	1.19E-05	5.65E-08	0	0	0	0	1
2017	0	1.21E-05	1.25E-05	0	0	5.61E-08	0	0	1
2018	0	9.82E-06	9.93E-06	0	0	0	0	0	1
2019	0	5.33E-06	5.38E-06	0	0	0	0	0	1
2020	0	1.59E-06	1.59E-06	0	0	0	0	0	1
2004	7.42E-07	4.17E-07	6.03E-06	2.74E-06	2.78E-07	2.78E-07	0	0	2
2005	2.19E-06	4.58E-07	9.61E-06	6.15E-06	5.59E-07	1.27E-06	0	3.05E-07	2
2006	1.83E-06	1.09E-06	7.91E-06	4.95E-06	4.95E-07	8.41E-07	4.95E-08	3.96E-07	2
2007	1.31E-06	6.03E-07	6.63E-06	4.97E-06	6.53E-07	1.05E-06	1E-07	3.01E-07	2
2008	8.35E-07	1.43E-05	2.82E-06	3.08E-06	4.18E-07	5.74E-07	1.04E-07	2.61E-07	2
2009	0	2.35E-05	1.07E-05	1.75E-07	0	4.08E-07	0	0	2
2010	2.8E-07	2.12E-05	1.1E-05	1.68E-07	0	1.68E-07	0	0	2
2011	3.42E-07	2.19E-05	1.08E-05	1.14E-07	0	1.71E-07	0	0	2
2012	3.47E-07	2.27E-05	1.18E-05	1.16E-07	0	1.16E-07	0	0	2
2013	2.47E-07	2.3E-05	1.08E-05	1.86E-07	0	2.47E-07	0	0	2
2014	1.25E-07	2.15E-05	9.69E-06	6.25E-08	0	0	0	0	2
2015	3.57E-07	2.02E-05	8.98E-06	5.95E-08	0	4.76E-07	0	0	2
2016	1.7E-07	2.09E-05	8.82E-06	1.13E-07	0	2.26E-07	0	0	2
2017	1.12E-07	1.98E-05	9.88E-06	1.12E-07	0	5.05E-07	0	0	2
2018	1.64E-07	1.57E-05	8.78E-06	1.64E-07	0	3.27E-07	0	0	2
2019	1.05E-07	8.89E-06	6.74E-06	0	0	2.61E-07	0	0	2
2020	0	2.35E-06	4.46E-06	5.87E-08	0	5.87E-08	0	0	2
2004	8.35E-07	1.95E-06	8.39E-06	3.8E-06	0	9.27E-07	2.5E-06	1.39E-07	3
2005	3.56E-07	1.17E-06	4.53E-06	1.83E-06	0	5.09E-07	1.37E-06	1.53E-07	3
2006	6.43E-07	1.63E-06	6.23E-06	2.97E-06	0	9.89E-07	2.27E-06	9.89E-08	3
2007	6.53E-07	1.61E-06	4.47E-06	1.36E-06	0	9.04E-07	1.91E-06	3.01E-07	3
2008	1.31E-06	3.34E-06	1.02E-05	2.92E-06	0	9.92E-07	3.71E-06	1.57E-07	3
2009	1.4E-06	3.84E-06	1.23E-05	4.43E-06	0	1.28E-06	5.42E-06	3.49E-07	3
2010	1.23E-06	3.14E-06	1.24E-05	4.31E-06	0	5.6E-07	5.04E-06	5.6E-08	3
2011	1.25E-06	3.64E-06	1.43E-05	5.01E-06	0	1.59E-06	5.69E-06	2.85E-07	3
2012	1.39E-06	3.76E-06	1.2E-05	3.36E-06	0	1.16E-06	4.92E-06	2.31E-07	3
2013	9.9E-07	3.09E-06	1.24E-05	4.76E-06	0	1.24E-06	4.95E-06	6.18E-08	3
2014	8.76E-07	3.94E-06	1.57E-05	3.57E-06	0	1.25E-06	5.44E-06	2.5E-07	3
2015	4.76E-07	2.74E-06	1.29E-05	3.87E-06	0	1.13E-06	4.4E-06	2.98E-07	3
2016	1.7E-07	9.61E-07	6.27E-06	2.94E-06	0	6.22E-07	2.54E-06	4.52E-07	3
2017	2.25E-07	8.98E-07	5.11E-06	3.31E-06	0	3.37E-07	2.02E-06	1.68E-07	3
2018	1.64E-07	8.18E-07	6E-06	3.6E-06	0	4.91E-07	2.35E-06	2.73E-07	3
2019	2.61E-07	9.93E-07	4.97E-06	2.04E-06	0	8.89E-07	1.99E-06	3.66E-07	3
2020	4.11E-07	1.06E-06	6.46E-06	2.58E-06	0	6.46E-07	2E-06	2.94E-07	3
2004	7.88E-07	1.58E-06	1.08E-05	2.27E-06	0	8.81E-07	6.03E-07	4.64E-08	4
2005	2.54E-07	7.63E-07	5.39E-06	1.32E-06	0	1.02E-07	3.05E-07	0	4
2006	9.89E-07	2.37E-06	8.7E-06	1.78E-06	0	6.43E-07	5.44E-07	4.95E-08	4
2007	4.52E-07	1.16E-06	5.22E-06	7.53E-07	0	7.03E-07	3.52E-07	1.51E-07	4
2008	9.4E-07	2.71E-06	7.67E-06	2.04E-06	0	9.4E-07	1.04E-06	1.04E-07	4
2009	8.15E-07	1.98E-06	6E-06	1.86E-06	0	7.57E-07	6.41E-07	2.33E-07	4
2010	4.48E-07	1.46E-06	6.72E-06	2.46E-06	0	1.68E-07	8.96E-07	0	4
2011	7.4E-07	1.99E-06	6.89E-06	1.25E-06	0	9.68E-07	1.14E-06	1.71E-07	4
2012	9.84E-07	2.2E-06	5.96E-06	1.74E-06	0	1.04E-06	9.84E-07	2.89E-07	4
2013	1.42E-06	2.47E-06	6E-06	2.04E-06	0	5.57E-07	7.42E-07	6.18E-08	4
2014	1E-06	2.56E-06	1.24E-05	2.31E-06	0	8.76E-07	1.38E-06	1.88E-07	4
2015	5.95E-07	2.44E-06	1.31E-05	1.67E-06	0	1.79E-07	1.25E-06	5.95E-08	4
2016	7.35E-07	2.54E-06	1.36E-05	3.17E-06	0	5.09E-07	8.48E-07	1.7E-07	4
2017	6.18E-07	2.7E-06	1.36E-05	3.26E-06	0	3.93E-07	8.42E-07	5.61E-08	4
2018	1.04E-06	4.25E-06	1.72E-05	3.93E-06	0	7.09E-07	9.27E-07	1.09E-07	4
2019	9.93E-07	3.29E-06	1.77E-05	3.45E-06	0	5.75E-07	7.84E-07	1.05E-07	4
2020	6.46E-07	1.76E-06	9.16E-06	2.06E-06	0	6.46E-07	7.05E-07	1.17E-07	4

Appendix K

Original Datasets Descriptive Statistics and Assumptions Tests

Figure K1

Box Plots Showing Parts 121 and 135 Dataset

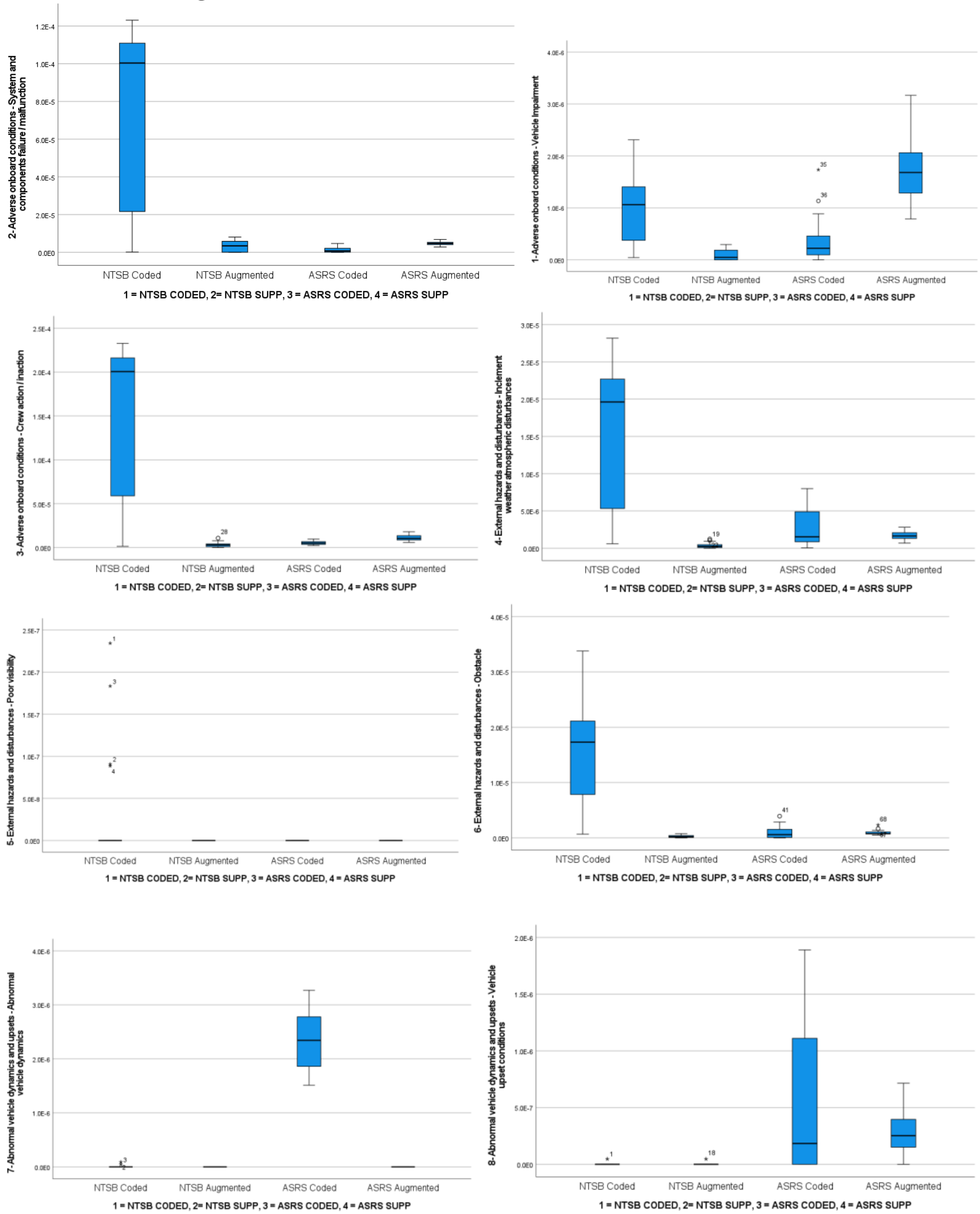


Table K2

Parts 121 and 135 Pearson Correlation Result

Correlations

		1- Adverse onboard conditions - Vehicle Impairment	2- Adverse onboard conditions - System and components failure / malfunction	3- Adverse onboard conditions - Crew action / inaction	4- External hazards and disturbances - Inclement weather atmospheric disturbances	5- External hazards and disturbances - Poor visibility	6- External hazards and disturbances - Obstacle	7- Abnormal vehicle dynamics and upsets - Abnormal vehicle dynamics	8- Abnormal vehicle dynamics and upsets - Vehicle upset conditions
1- Adverse onboard conditions - Vehicle Impairment	Pearson Correlation	1	.315**	.301*	.251*	-.138	.215	-.289*	-.098
	Sig. (2-tailed)		.009	.013	.039	.261	.079	.017	.424
	N	68	68	68	68	68	68	68	68
2- Adverse onboard conditions - System and components failure / malfunction	Pearson Correlation	.315**	1	.995**	.964**	-.122	.929**	-.283*	-.247*
	Sig. (2-tailed)	.009		<.001	<.001	.324	<.001	.019	.042
	N	68	68	68	68	68	68	68	68
3- Adverse onboard conditions - Crew action / inaction	Pearson Correlation	.301*	.995**	1	.967**	-.119	.933**	-.267*	-.229
	Sig. (2-tailed)	.013	<.001		<.001	.333	<.001	.028	.060
	N	68	68	68	68	68	68	68	68
4- External hazards and disturbances - Inclement weather atmospheric disturbances	Pearson Correlation	.251*	.964**	.967**	1	-.115	.938**	-.141	-.233
	Sig. (2-tailed)	.039	<.001	<.001		.350	<.001	.251	.056
	N	68	68	68	68	68	68	68	68
5- External hazards and disturbances - Poor visibility	Pearson Correlation	-.138	-.122	-.119	-.115	1	-.092	-.123	-.107
	Sig. (2-tailed)	.261	.324	.333	.350		.455	.317	.386
	N	68	68	68	68	68	68	68	68
6- External hazards and disturbances - Obstacle	Pearson Correlation	.215	.929**	.933**	.938**	-.092	1	-.233	-.159
	Sig. (2-tailed)	.079	<.001	<.001	<.001	.455		.055	.195
	N	68	68	68	68	68	68	68	68
7- Abnormal vehicle dynamics and upsets - Abnormal vehicle dynamics	Pearson Correlation	-.289*	-.283*	-.267*	-.141	-.123	-.233	1	.490**
	Sig. (2-tailed)	.017	.019	.028	.251	.317	.055		<.001
	N	68	68	68	68	68	68	68	68
8- Abnormal vehicle dynamics and upsets - Vehicle upset conditions	Pearson Correlation	-.098	-.247*	-.229	-.233	-.107	-.159	.490**	1
	Sig. (2-tailed)	.424	.042	.060	.056	.386	.195	<.001	
	N	68	68	68	68	68	68	68	68

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure K3

Scattered Plots Matrix for Parts 121 and 135 Dataset

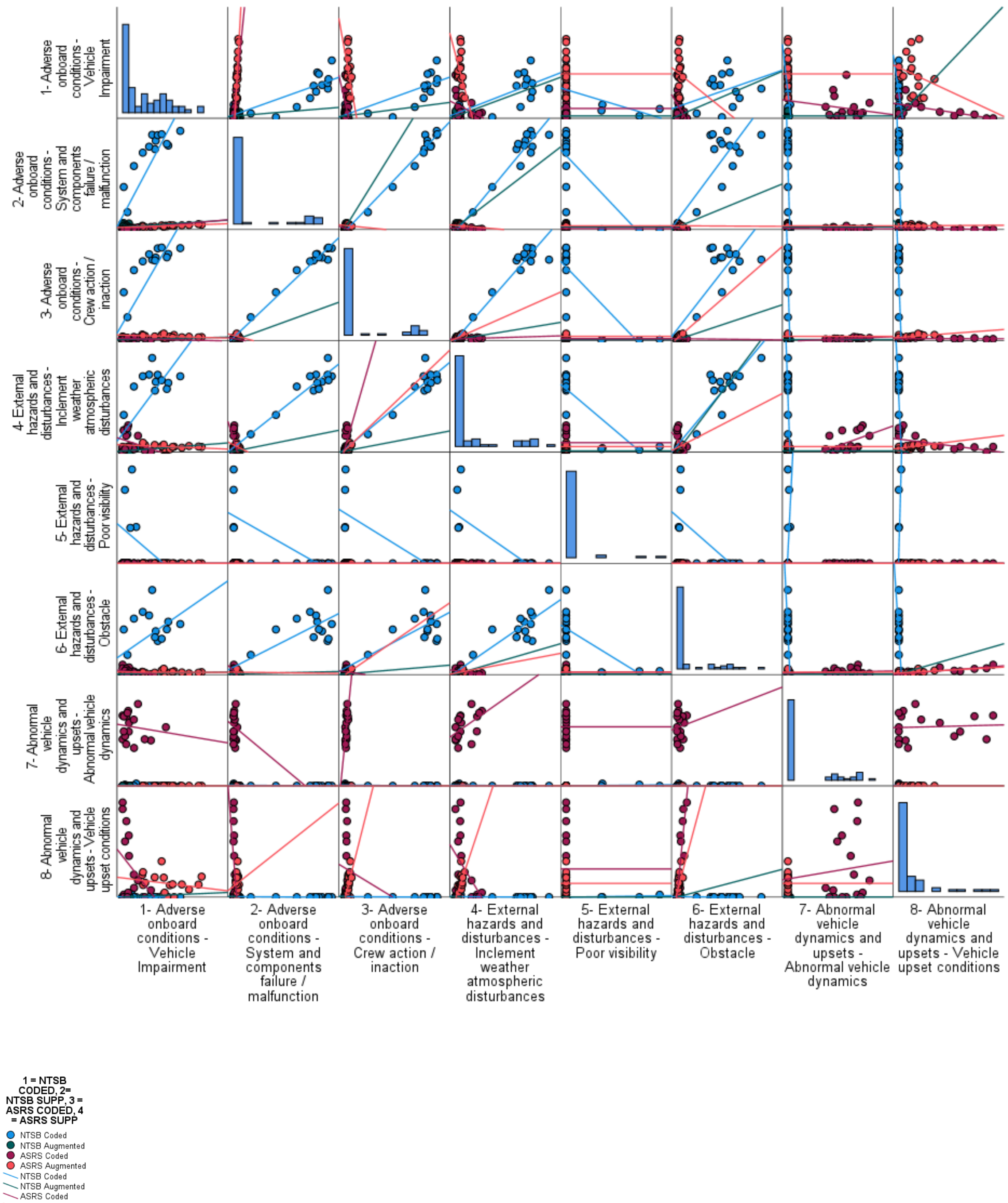


Figure K4

Part 91 Box Plots

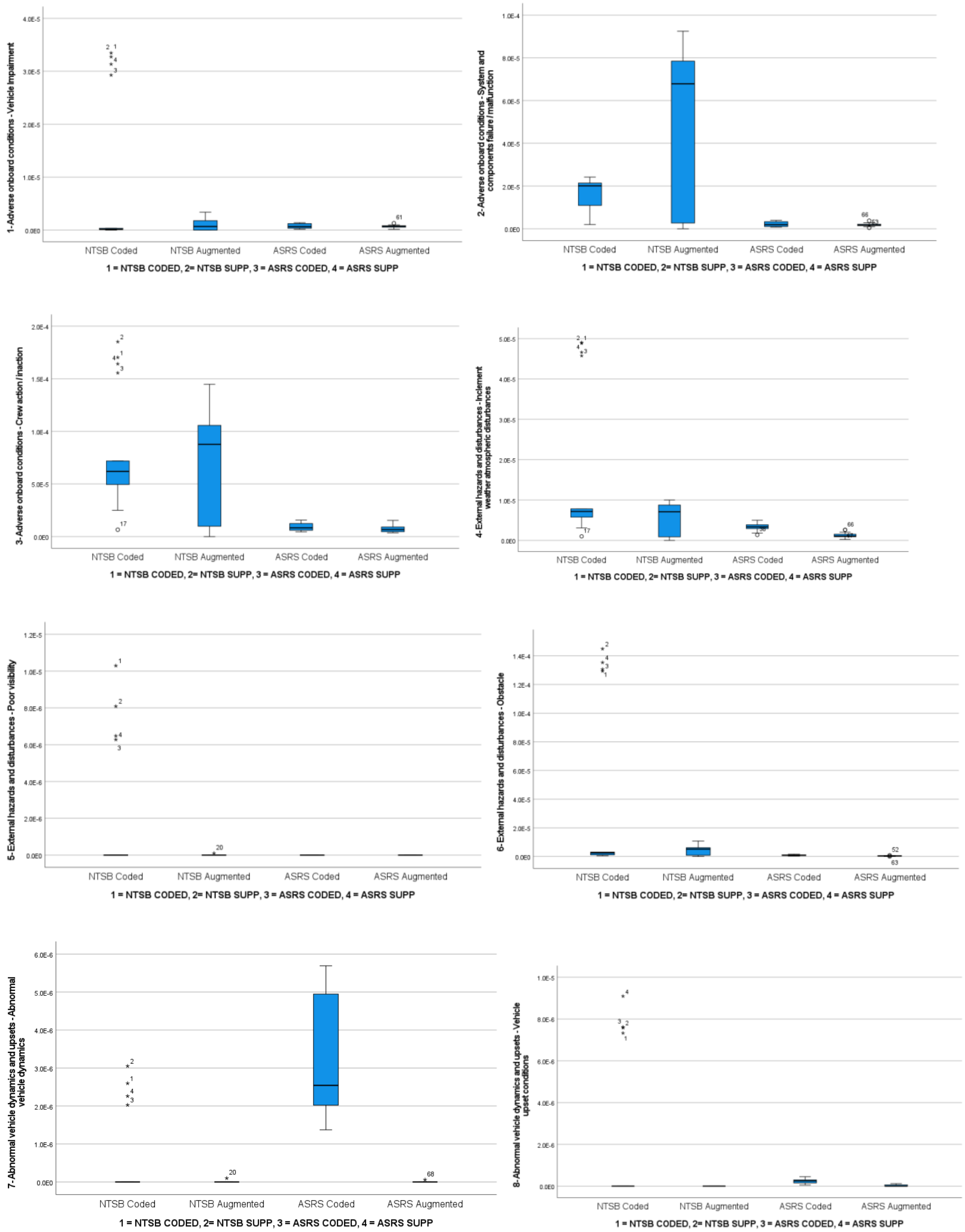


Table K5

Part 91 Pearson Correlation Result

Correlations

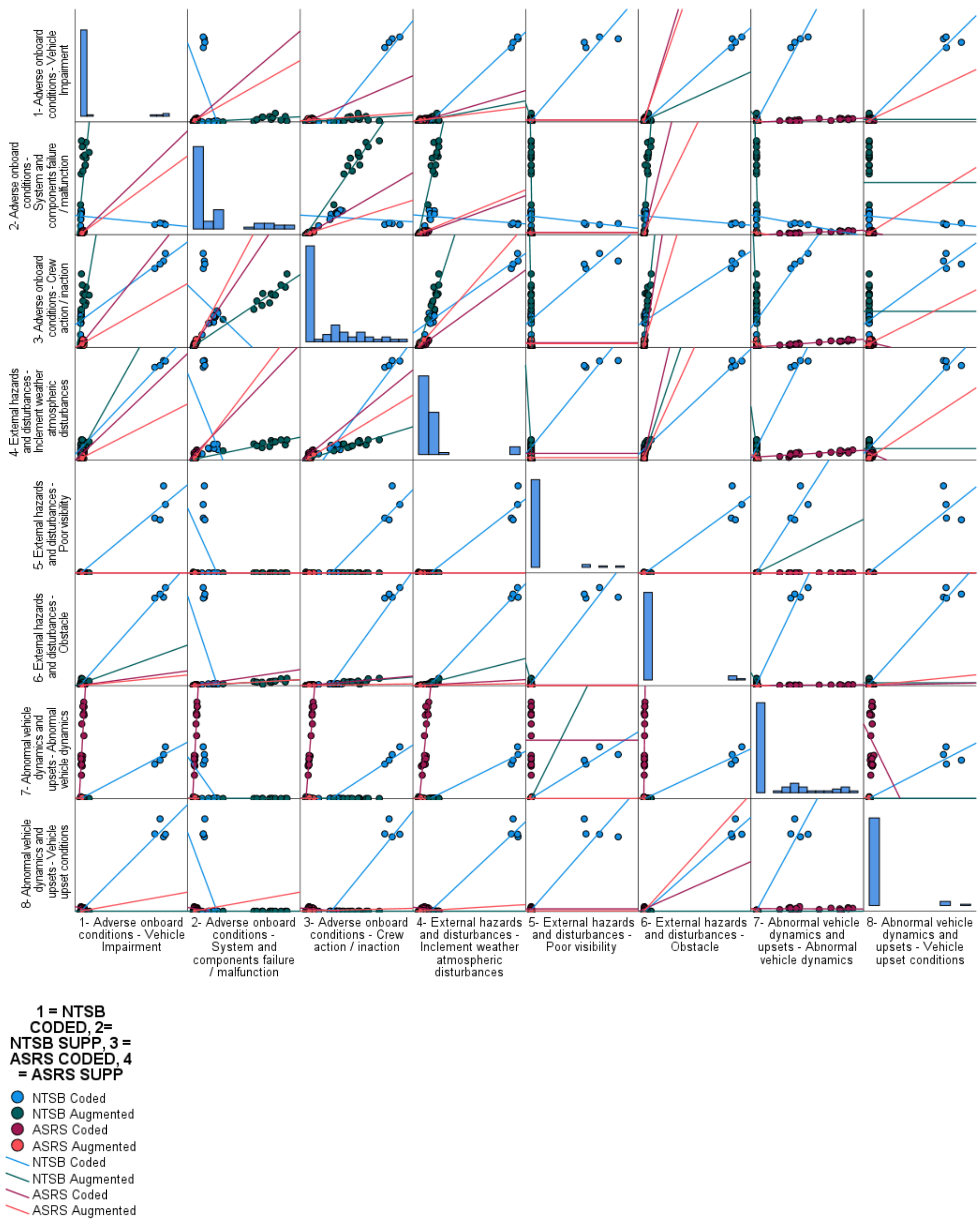
		1- Adverse onboard conditions - Vehicle Impairment	2- Adverse onboard conditions - System and components failure / malfunction	3- Adverse onboard conditions - Crew action / inaction	4- External hazards and disturbances - Inclement weather atmospheric disturbances	5- External hazards and disturbances - Poor visibility	6- External hazards and disturbances - Obstacle	7- Abnormal vehicle dynamics and upsets - Abnormal vehicle dynamics	8- Abnormal vehicle dynamics and upsets - Vehicle upset conditions
1- Adverse onboard conditions - Vehicle Impairment	Pearson Correlation	1	-.027	.668**	.967**	.980**	.995**	.235	.988**
	Sig. (2-tailed)		.828	<.001	<.001	<.001	<.001	.054	<.001
	N	68	68	68	68	68	68	68	68
2- Adverse onboard conditions - System and components failure / malfunction	Pearson Correlation	-.027	1	.679**	.152	-.068	-.003	-.311**	-.090
	Sig. (2-tailed)	.828		<.001	.216	.582	.983	.010	.467
	N	68	68	68	68	68	68	68	68
3- Adverse onboard conditions - Crew action / inaction	Pearson Correlation	.668**	.679**	1	.809**	.637**	.695**	-.099	.623**
	Sig. (2-tailed)	<.001	<.001		<.001	<.001	<.001	.422	<.001
	N	68	68	68	68	68	68	68	68
4- External hazards and disturbances - Inclement weather atmospheric disturbances	Pearson Correlation	.967**	.152	.809**	1	.948**	.976**	.198	.953**
	Sig. (2-tailed)	<.001	.216	<.001		<.001	<.001	.106	<.001
	N	68	68	68	68	68	68	68	68
5- External hazards and disturbances - Poor visibility	Pearson Correlation	.980**	-.068	.637**	.948**	1	.974**	.219	.960**
	Sig. (2-tailed)	<.001	.582	<.001	<.001		<.001	.073	<.001
	N	68	68	68	68	68	68	68	68
6- External hazards and disturbances - Obstacle	Pearson Correlation	.995**	-.003	.695**	.976**	.974**	1	.205	.990**
	Sig. (2-tailed)	<.001	.983	<.001	<.001	<.001		.093	<.001
	N	68	68	68	68	68	68	68	68
7- Abnormal vehicle dynamics and upsets - Abnormal vehicle dynamics	Pearson Correlation	.235	-.311**	-.099	.198	.219	.205	1	.257*
	Sig. (2-tailed)	.054	.010	.422	.106	.073	.093		.034
	N	68	68	68	68	68	68	68	68
8- Abnormal vehicle dynamics and upsets - Vehicle upset conditions	Pearson Correlation	.988**	-.090	.623**	.953**	.960**	.990**	.257*	1
	Sig. (2-tailed)	<.001	.467	<.001	<.001	<.001	<.001	.034	
	N	68	68	68	68	68	68	68	68

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure K6

Scattered Plots Matrix for Part 91 Dataset



Appendix L

Assumptions Test Results on Transformed Parts 121 and 135 and Part 91 Datasets

Table L1

Normality Test Results on Transformed Parts 121 and 135 and Part 91 Datasets

Number of DVs with Shapiro Tests p < .05	Original	Sq root	Cube root	Quartic Root	Log10_PlusOne	Ln_PlusOne	Inv_PlusOne	Sq Root with Data 1-4 removed
Parts 121& 135	14	12	11 [a]	14	17	23	17	N/A
Part 91	19	20	20	20	19	19	19	13 [a]

[a] Lowest Number

Table L2

Parts 121 and 135 Cube Root Transformed Normality Tests Results

Tests of Normality

	1 = NTSB CODED, 2= NTSB SUPP, 3 = ASRS CODED, 4 = ASRS SUPP	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
TX1_CUBE_RT	NTSB Coded	.169	17	.200*	.940	17	.322
	NTSB Augmented	.310	17	<.001	.762	17	<.001
	ASRS Coded	.109	17	.200*	.966	17	.747
	ASRS Augmented	.146	17	.200*	.937	17	.289
TX2_CUBE_RT	NTSB Coded	.299	17	<.001	.701	17	<.001
	NTSB Augmented	.198	17	.077	.878	17	.030
	ASRS Coded	.118	17	.200*	.943	17	.354
	ASRS Augmented	.101	17	.200*	.989	17	.998
TX3_CUBE_RT	NTSB Coded	.326	17	<.001	.686	17	<.001
	NTSB Augmented	.123	17	.200*	.974	17	.890
	ASRS Coded	.077	17	.200*	.975	17	.905
	ASRS Augmented	.136	17	.200*	.973	17	.875
TX4_CUBE_RT	NTSB Coded	.326	17	<.001	.764	17	<.001
	NTSB Augmented	.133	17	.200*	.922	17	.158
	ASRS Coded	.166	17	.200*	.924	17	.170
	ASRS Augmented	.189	17	.108	.940	17	.315
TX5_CUBE_RT	NTSB Coded	.467	17	<.001	.569	17	<.001
	NTSB Augmented	.	17	.	.	17	.
	ASRS Coded	.	17	.	.	17	.
	ASRS Augmented	.	17	.	.	17	.
TX6_CUBE_RT	NTSB Coded	.264	17	.003	.828	17	.005
	NTSB Augmented	.189	17	.108	.848	17	.010
	ASRS Coded	.112	17	.200*	.959	17	.615
	ASRS Augmented	.186	17	.122	.903	17	.076
TX7_CUBE_RT	NTSB Coded	.520	17	<.001	.398	17	<.001
	NTSB Augmented	.	17	.	.	17	.
	ASRS Coded	.142	17	.200*	.954	17	.516
	ASRS Augmented	.	17	.	.	17	.
TX8_CUBE_RT	NTSB Coded	.537	17	<.001	.262	17	<.001
	NTSB Augmented	.537	17	<.001	.262	17	<.001
	ASRS Coded	.189	17	.108	.885	17	.038
	ASRS Augmented	.274	17	.001	.823	17	.004

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table L3

Parts 91 Square Root Transformed Normality Tests Results

Tests of Normality

	1 = NTSB CODED, 2= NTSB SUPP, 3 = ASRS CODED, 4 = ASRS SUPP	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
TX1_SQRT	NTSB Coded	.429	17	<.001	.600	17	<.001
	NTSB Augmented	.175	17	.176	.889	17	.044
	ASRS Coded	.146	17	.200*	.916	17	.126
	ASRS Augmented	.133	17	.200*	.965	17	.729
TX2_SQRT	NTSB Coded	.239	17	.011	.855	17	.013
	NTSB Augmented	.311	17	<.001	.744	17	<.001
	ASRS Coded	.182	17	.136	.868	17	.020
	ASRS Augmented	.113	17	.200*	.976	17	.910
TX3_SQRT	NTSB Coded	.266	17	.002	.882	17	.034
	NTSB Augmented	.273	17	.001	.778	17	.001
	ASRS Coded	.208	17	.048	.890	17	.046
	ASRS Augmented	.220	17	.028	.904	17	.079
TX4_SQRT	NTSB Coded	.391	17	<.001	.717	17	<.001
	NTSB Augmented	.282	17	<.001	.748	17	<.001
	ASRS Coded	.140	17	.200*	.961	17	.649
	ASRS Augmented	.128	17	.200*	.963	17	.681
TX5_SQRT	NTSB Coded	.468	17	<.001	.563	17	<.001
	NTSB Augmented	.537	17	<.001	.262	17	<.001
	ASRS Coded	.	17	.	.	17	.
	ASRS Augmented	.	17	.	.	17	.
TX6_SQRT	NTSB Coded	.434	17	<.001	.608	17	<.001
	NTSB Augmented	.233	17	.015	.881	17	.033
	ASRS Coded	.157	17	.200*	.962	17	.665
	ASRS Augmented	.176	17	.166	.931	17	.227
TX7_SQRT	NTSB Coded	.469	17	<.001	.557	17	<.001
	NTSB Augmented	.537	17	<.001	.262	17	<.001
	ASRS Coded	.225	17	.022	.870	17	.022
	ASRS Augmented	.537	17	<.001	.262	17	<.001
TX8_SQRT	NTSB Coded	.469	17	<.001	.547	17	<.001
	NTSB Augmented	.	17	.	.	17	.
	ASRS Coded	.146	17	.200*	.952	17	.485
	ASRS Augmented	.343	17	<.001	.742	17	<.001

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table L4

Parts 91 Square Root Transformed with Data Items 1-4 Removed Normality Tests Results

Tests of Normality

	1 = NTSB CODED, 2= NTSB SUPP, 3 = ASRS CODED, 4 = ASRS SUPP	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
TX1_SQRT	NTSB Coded	.188	13	.200*	.919	13	.241
	NTSB Augmented	.201	16	.083	.885	16	.046
	ASRS Coded	.146	17	.200*	.916	17	.126
	ASRS Augmented	.133	17	.200*	.965	17	.729
TX2_SQRT	NTSB Coded	.314	13	.001	.715	13	<.001
	NTSB Augmented	.323	16	<.001	.735	16	<.001
	ASRS Coded	.182	17	.136	.868	17	.020
	ASRS Augmented	.113	17	.200*	.976	17	.910
TX3_SQRT	NTSB Coded	.290	13	.004	.729	13	.001
	NTSB Augmented	.282	16	.001	.768	16	.001
	ASRS Coded	.208	17	.048	.890	17	.046
	ASRS Augmented	.220	17	.028	.904	17	.079
TX4_SQRT	NTSB Coded	.261	13	.016	.756	13	.002
	NTSB Augmented	.291	16	<.001	.739	16	<.001
	ASRS Coded	.140	17	.200*	.961	17	.649
	ASRS Augmented	.128	17	.200*	.963	17	.681
TX5_SQRT	NTSB Coded	.	13	.	.	13	.
	NTSB Augmented	.	16	.	.	16	.
	ASRS Coded	.	17	.	.	17	.
	ASRS Augmented	.	17	.	.	17	.
TX6_SQRT	NTSB Coded	.156	13	.200*	.898	13	.127
	NTSB Augmented	.242	16	.013	.872	16	.029
	ASRS Coded	.157	17	.200*	.962	17	.665
	ASRS Augmented	.176	17	.166	.931	17	.227
TX7_SQRT	NTSB Coded	.	13	.	.	13	.
	NTSB Augmented	.	16	.	.	16	.
	ASRS Coded	.225	17	.022	.870	17	.022
	ASRS Augmented	.537	17	<.001	.262	17	<.001
TX8_SQRT	NTSB Coded	.	13	.	.	13	.
	NTSB Augmented	.	16	.	.	16	.
	ASRS Coded	.146	17	.200*	.952	17	.485
	ASRS Augmented	.343	17	<.001	.742	17	<.001

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Figure L5

Outliers Analysis for Parts 121 and 135 Transformed Dataset

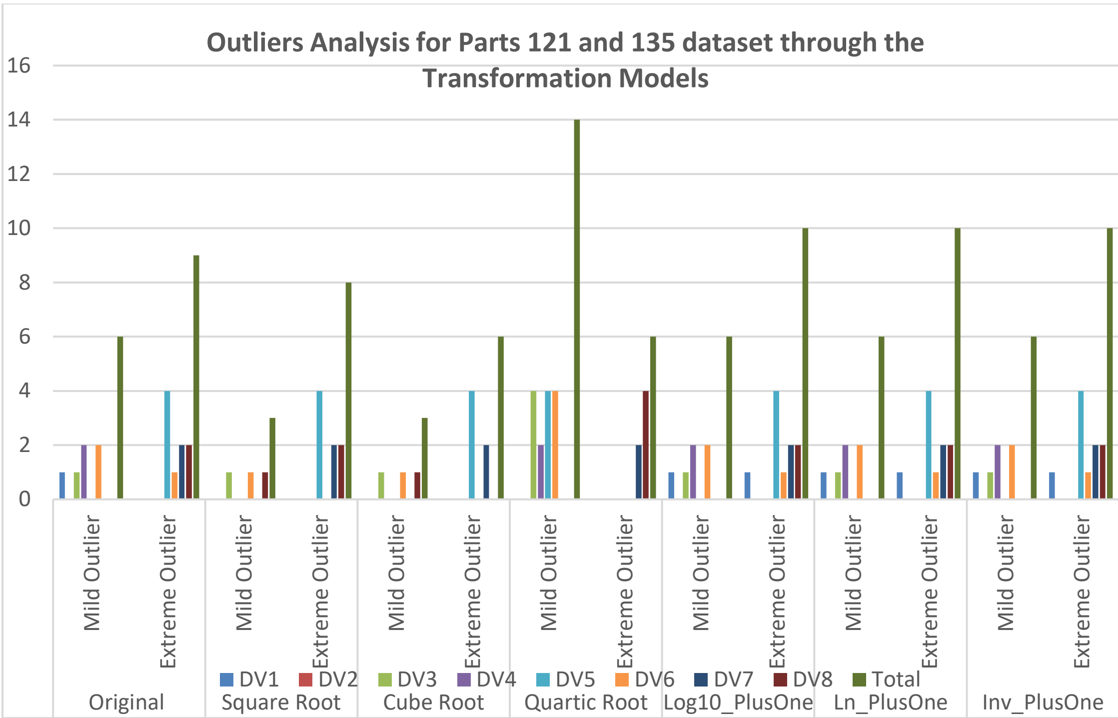


Figure L6

Outliers Analysis for Part 91 Transformed Dataset

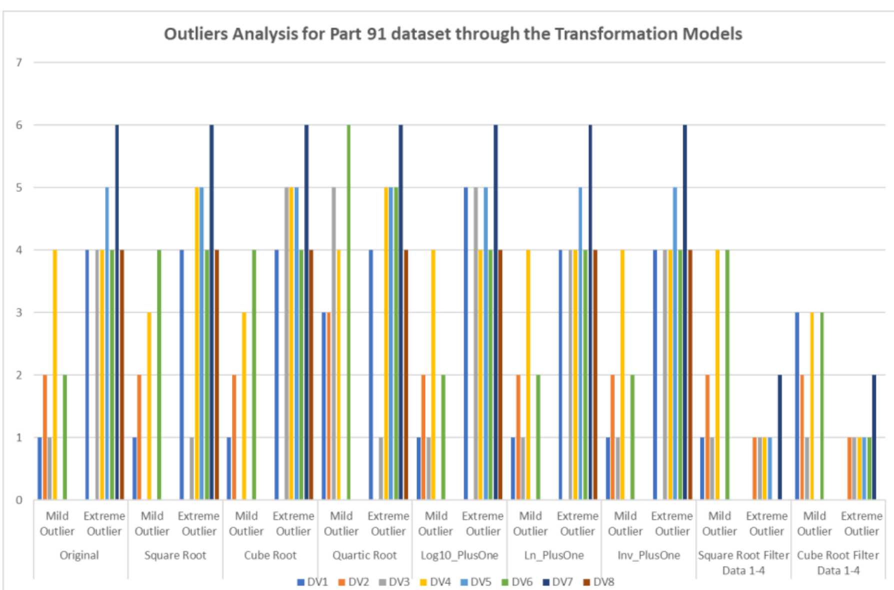


Figure L7

Box Plot of Cube Root Transformed Parts 121 and 135 Dataset

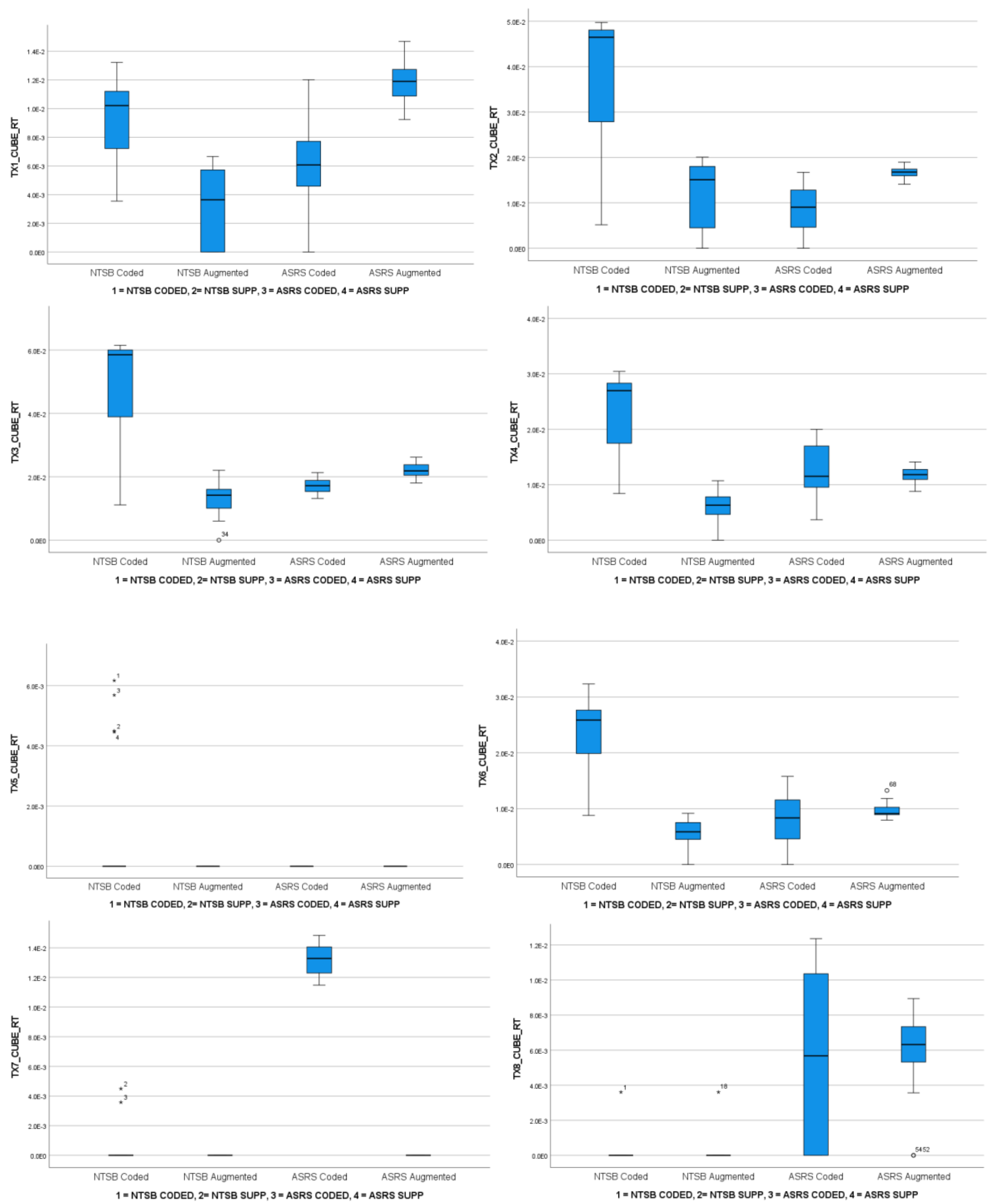


Figure L8

Scatterplot Matrix of Cube Root Transformed Parts 121 and 135 Dataset

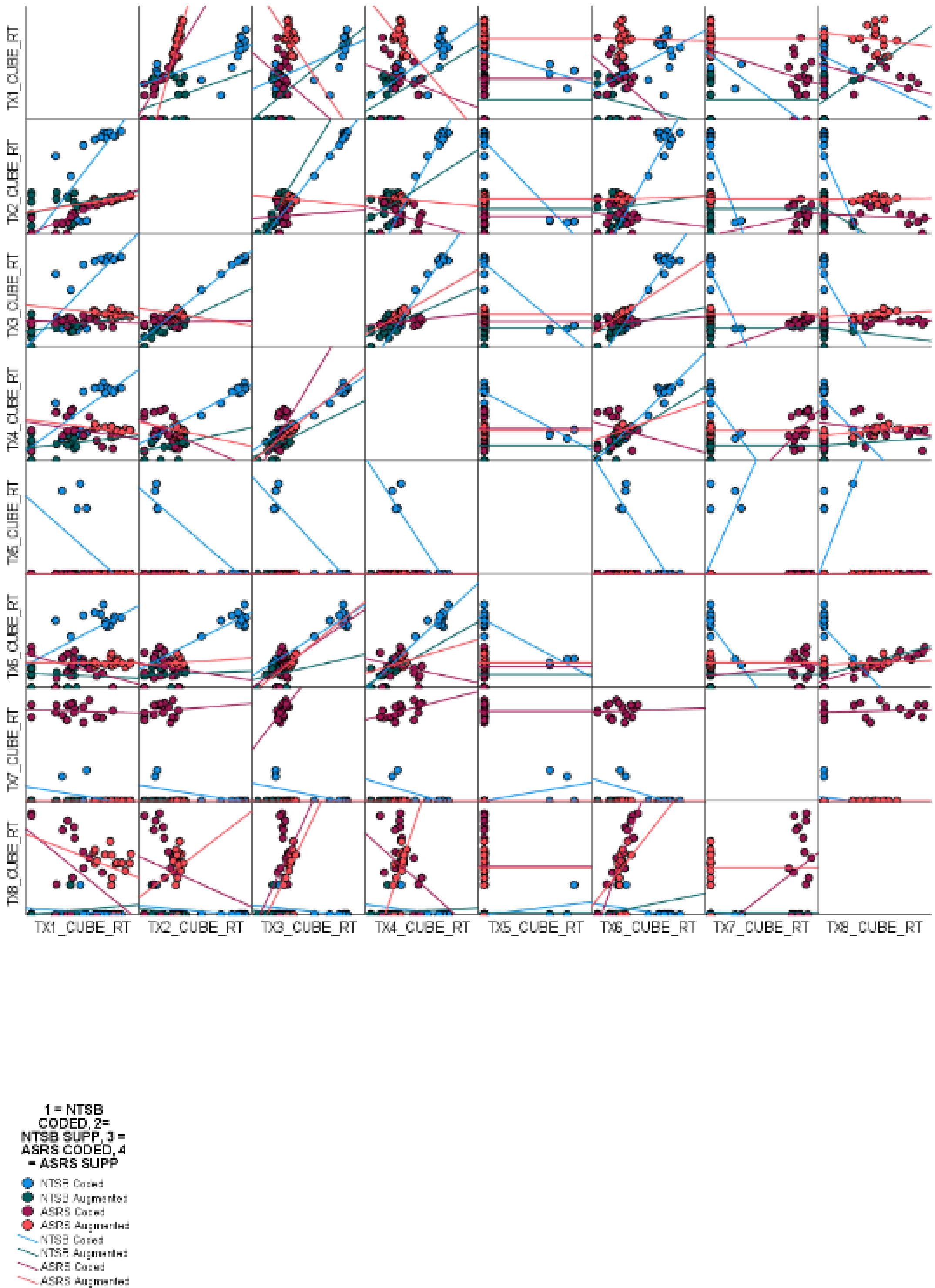


Table L9

Multicollinearity Test on Parts 121 and 135 Dataset by Pearson Correlation

Correlations

		TX1_CUBE_RT	TX2_CUBE_RT	TX3_CUBE_RT	TX4_CUBE_RT	TX5_CUBE_RT	TX6_CUBE_RT	TX7_CUBE_RT	TX8_CUBE_RT
TX1_CUBE_RT	Pearson Correlation	1	.482**	.454**	.421**	-.053	.341**	-.203	.131
	Sig. (2-tailed)		<.001	<.001	<.001	.670	.004	.097	.288
	N	68	68	68	68	68	68	68	68
TX2_CUBE_RT	Pearson Correlation	.482**	1	.952**	.788**	-.212	.833**	-.410**	-.338**
	Sig. (2-tailed)	<.001		<.001	<.001	.082	<.001	<.001	.005
	N	68	68	68	68	68	68	68	68
TX3_CUBE_RT	Pearson Correlation	.454**	.952**	1	.900**	-.178	.910**	-.274*	-.237
	Sig. (2-tailed)	<.001	<.001		<.001	.146	<.001	.024	.052
	N	68	68	68	68	68	68	68	68
TX4_CUBE_RT	Pearson Correlation	.421**	.788**	.900**	1	-.091	.835**	-.061	-.175
	Sig. (2-tailed)	<.001	<.001	<.001		.462	<.001	.619	.152
	N	68	68	68	68	68	68	68	68
TX5_CUBE_RT	Pearson Correlation	-.053	-.212	-.178	-.091	1	-.033	-.064	-.125
	Sig. (2-tailed)	.670	.082	.146	.462		.788	.605	.311
	N	68	68	68	68	68	68	68	68
TX6_CUBE_RT	Pearson Correlation	.341**	.833**	.910**	.835**	-.033	1	-.229	-.076
	Sig. (2-tailed)	.004	<.001	<.001	<.001	.788		.060	.536
	N	68	68	68	68	68	68	68	68
TX7_CUBE_RT	Pearson Correlation	-.203	-.410**	-.274*	-.061	-.064	-.229	1	.420**
	Sig. (2-tailed)	.097	<.001	.024	.619	.605	.060		<.001
	N	68	68	68	68	68	68	68	68
TX8_CUBE_RT	Pearson Correlation	.131	-.338**	-.237	-.175	-.125	-.076	.420**	1
	Sig. (2-tailed)	.288	.005	.052	.152	.311	.536	<.001	
	N	68	68	68	68	68	68	68	68

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Note. Relationships with $p < .05$ and $r > 0.8$:

TX3/TX2, TX6/TX2, TX4/TX3, TX6/TX3, TX6/TX4

Figure L10

Scatterplot Matrix of Square Root Transformed Part 91 Dataset

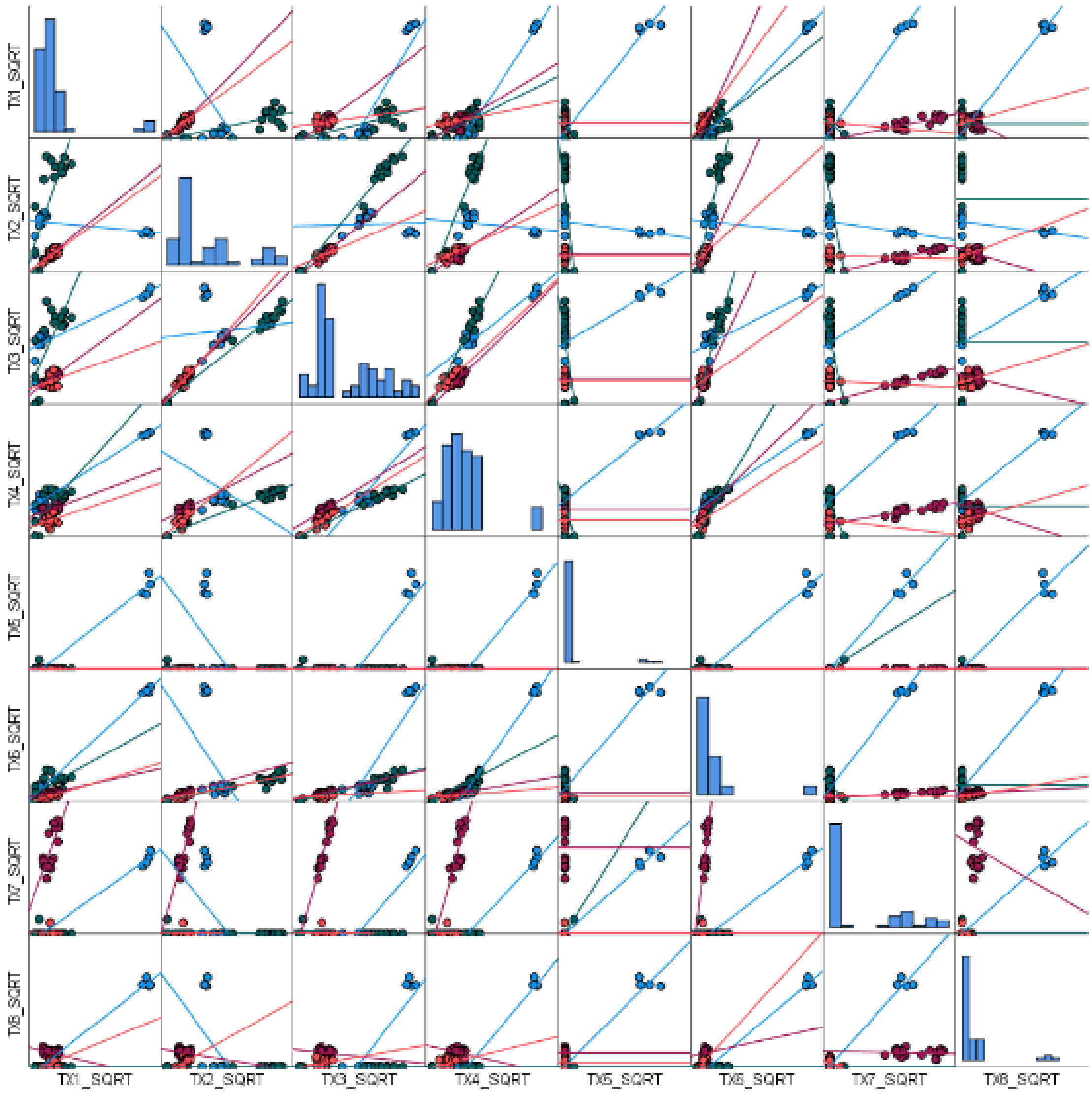


Table L11

Multicollinearity Test on Part 91 Dataset by Pearson Correlation

Correlations

		TX1_SQRT	TX2_SQRT	TX3_SQRT	TX4_SQRT	TX5_SQRT	TX6_SQRT	TX7_SQRT	TX8_SQRT
TX1_SQRT	Pearson Correlation	1	.121	.568**	.840**	.937**	.938**	.358**	.915**
	Sig. (2-tailed)		.326	<.001	<.001	<.001	<.001	.003	<.001
	N	68	68	68	68	68	68	68	68
TX2_SQRT	Pearson Correlation	.121	1	.828**	.470**	-.004	.261*	-.330**	-.137
	Sig. (2-tailed)	.326		<.001	<.001	.973	.032	.006	.264
	N	68	68	68	68	68	68	68	68
TX3_SQRT	Pearson Correlation	.568**	.828**	1	.863**	.509**	.713**	-.109	.374**
	Sig. (2-tailed)	<.001	<.001		<.001	<.001	<.001	.376	.002
	N	68	68	68	68	68	68	68	68
TX4_SQRT	Pearson Correlation	.840**	.470**	.863**	1	.812**	.923**	.264*	.760**
	Sig. (2-tailed)	<.001	<.001	<.001		<.001	<.001	.029	<.001
	N	68	68	68	68	68	68	68	68
TX5_SQRT	Pearson Correlation	.937**	-.004	.509**	.812**	1	.951**	.302*	.936**
	Sig. (2-tailed)	<.001	.973	<.001	<.001		<.001	.012	<.001
	N	68	68	68	68	68	68	68	68
TX6_SQRT	Pearson Correlation	.938**	.261*	.713**	.923**	.951**	1	.253*	.888**
	Sig. (2-tailed)	<.001	.032	<.001	<.001	<.001		.037	<.001
	N	68	68	68	68	68	68	68	68
TX7_SQRT	Pearson Correlation	.358**	-.330**	-.109	.264*	.302*	.253*	1	.536**
	Sig. (2-tailed)	.003	.006	.376	.029	.012	.037		<.001
	N	68	68	68	68	68	68	68	68
TX8_SQRT	Pearson Correlation	.915**	-.137	.374**	.760**	.936**	.888**	.536**	1
	Sig. (2-tailed)	<.001	.264	.002	<.001	<.001	<.001	<.001	
	N	68	68	68	68	68	68	68	68

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix M

Parts 121 and 135 Cube Root Transformed Dataset MANOVA Results

Table M1

Parts 121 and 135 Cube Root Transformed Dataset Estimated Means and Standard Deviation Results

Dependent Variable	1 = NTSB CODED, 2= NTSB SUPP, 3 = ASRS CODED, 4 = ASRS SUPP	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
TX1_CUBE_RT	NTSB Coded	.00917901	.00066630	.0078479	.011
	NTSB Augmented	.00289090	.00066630	.0015598	.004
	ASRS Coded	.00604622	.00066630	.0047151	.007
	ASRS Augmented	.01187062	.00066630	.0105395	.013
TX2_CUBE_RT	NTSB Coded	.03592771	.00244664	.0310400	.041
	NTSB Augmented	.01193318	.00244664	.0070455	.017
	ASRS Coded	.00835683	.00244664	.0034691	.013
	ASRS Augmented	.01666097	.00244664	.0117732	.022
TX3_CUBE_RT	NTSB Coded	.04664863	.00257695	.0415006	.052
	NTSB Augmented	.01288629	.00257695	.0077382	.018
	ASRS Coded	.01711046	.00257695	.0119624	.022
	ASRS Augmented	.02206582	.00257695	.0169178	.027
TX4_CUBE_RT	NTSB Coded	.02279122	.00123346	.0203271	.025
	NTSB Augmented	.00578678	.00123346	.0033227	.008
	ASRS Coded	.01217923	.00123346	.0097151	.015
	ASRS Augmented	.01159063	.00123346	.0091265	.014
TX6_CUBE_RT	NTSB Coded	.02264259	.00117101	.0203032	.025
	NTSB Augmented	.00512176	.00117101	.0027824	.007
	ASRS Coded	.00821750	.00117101	.0058781	.011
	ASRS Augmented	.00967379	.00117101	.0073344	.012
TX7_CUBE_RT	NTSB Coded	.00047489	.00020405	.0000673	.001
	NTSB Augmented	.00000000	.00020405	-.0004076	.000
	ASRS Coded	.01323719	.00020405	.0128296	.014
	ASRS Augmented	.00000000	.00020405	-.0004076	.000
TX8_CUBE_RT	NTSB Coded	.00021208	.00066242	-.0011112	.002
	NTSB Augmented	.00021208	.00066242	-.0011112	.002
	ASRS Coded	.00587685	.00066242	.0045535	.007
	ASRS Augmented	.00575695	.00066242	.0044336	.007

Table M2

Parts 121 and 135 Cube Root Transformed MANOVA Post Hoc Results

Dependent Variable		(I) Code Type	(J) Code Type	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
TX1_CUBE_RT	Tukey HSD	NTSB Coded	NTSB Augmented	6.3000E-003*	9.42285E-004	<.001	3.8000E-003	8.8000E-003
			ASRS Coded	3.1000E-003*	9.42285E-004	.008	6.4719E-004	5.6000E-003
			ASRS Augmented	-2.7000E-003*	9.42285E-004	.029	-5.2000E-003	-2.0602E-004
		NTSB Augmented	NTSB Coded	-6.3000E-003*	9.42285E-004	<.001	-8.8000E-003	-3.8000E-003
			ASRS Coded	-3.2000E-003*	9.42285E-004	.007	-5.6000E-003	-6.6972E-004
			ASRS Augmented	-9.0000E-003*	9.42285E-004	<.001	-1.1500E-002	-6.5000E-003
		ASRS Coded	NTSB Coded	-3.1000E-003*	9.42285E-004	.008	-5.6000E-003	-6.4719E-004
			NTSB Augmented	3.2000E-003*	9.42285E-004	.007	6.6972E-004	5.6000E-003
			ASRS Augmented	-5.8000E-003*	9.42285E-004	<.001	-8.3000E-003	-3.3000E-003
		ASRS Augmented	NTSB Coded	2.7000E-003*	9.42285E-004	.029	2.0602E-004	5.2000E-003
			NTSB Augmented	9.0000E-003*	9.42285E-004	<.001	6.5000E-003	1.1500E-002
			ASRS Coded	5.8000E-003*	9.42285E-004	<.001	3.3000E-003	8.3000E-003
TX2_CUBE_RT	Tukey HSD	NTSB Coded	NTSB Augmented	2.4000E-002*	3.46000E-003	<.001	1.4900E-002	3.3100E-002
			ASRS Coded	2.7600E-002*	3.46000E-003	<.001	1.8400E-002	3.6700E-002
			ASRS Augmented	1.9300E-002*	3.46000E-003	<.001	1.0100E-002	2.8400E-002
		NTSB Augmented	NTSB Coded	-2.4000E-002*	3.46000E-003	<.001	-3.3100E-002	-1.4900E-002
			ASRS Coded	3.6000E-003	3.46000E-003	.730	-5.6000E-003	1.2700E-002
			ASRS Augmented	-4.7000E-003	3.46000E-003	.525	-1.3900E-002	4.4000E-003
		ASRS Coded	NTSB Coded	-2.7600E-002*	3.46000E-003	<.001	-3.6700E-002	-1.8400E-002
			NTSB Augmented	-3.6000E-003	3.46000E-003	.730	-1.2700E-002	5.6000E-003
			ASRS Augmented	-8.3000E-003	3.46000E-003	.087	-1.7400E-002	8.2296E-004
		ASRS Augmented	NTSB Coded	-1.9300E-002*	3.46000E-003	<.001	-2.8400E-002	-1.0100E-002
			NTSB Augmented	4.7000E-003	3.46000E-003	.525	-4.4000E-003	1.3900E-002
			ASRS Coded	8.3000E-003	3.46000E-003	.087	-8.2296E-004	1.7400E-002
TX3_CUBE_RT	Tukey HSD	NTSB Coded	NTSB Augmented	3.3800E-002*	3.64000E-003	<.001	2.4100E-002	4.3400E-002
			ASRS Coded	2.9500E-002*	3.64000E-003	<.001	1.9900E-002	3.9200E-002
			ASRS Augmented	2.4600E-002*	3.64000E-003	<.001	1.5000E-002	3.4200E-002
		NTSB Augmented	NTSB Coded	-3.3800E-002*	3.64000E-003	<.001	-4.3400E-002	-2.4100E-002
			ASRS Coded	-4.2000E-003	3.64000E-003	.655	-1.3800E-002	5.4000E-003
			ASRS Augmented	-9.2000E-003	3.64000E-003	.066	-1.8800E-002	4.3368E-004
			NTSB Coded	-2.9500E-002*	3.64000E-003	<.001	-3.9200E-002	-1.9900E-002

		ASRS Coded	NTSB Augmented	4.2000E-003	3.64000E-003	.655	-5.4000E-003	1.3800E-002
			ASRS Augmented	-5.0000E-003	3.64000E-003	.529	-1.4600E-002	4.7000E-003
		ASRS Augmented	NTSB Coded	-2.4600E-002*	3.64000E-003	<.001	-3.4200E-002	-1.5000E-002
			NTSB Augmented	9.2000E-003	3.64000E-003	.066	-4.3368E-004	1.8800E-002
			ASRS Coded	5.0000E-003	3.64000E-003	.529	-4.7000E-003	1.4600E-002
TX4_CUBE_RT	Tukey HSD	NTSB Coded	NTSB Augmented	1.7000E-002*	1.74000E-003	<.001	1.2400E-002	2.1600E-002
			ASRS Coded	1.0600E-002*	1.74000E-003	<.001	6.0000E-003	1.5200E-002
			ASRS Augmented	1.1200E-002*	1.74000E-003	<.001	6.6000E-003	1.5800E-002
		NTSB Augmented	NTSB Coded	-1.7000E-002*	1.74000E-003	<.001	-2.1600E-002	-1.2400E-002
			ASRS Coded	-6.4000E-003*	1.74000E-003	.003	-1.1000E-002	-1.8000E-003
			ASRS Augmented	-5.8000E-003*	1.74000E-003	.008	-1.0400E-002	-1.2000E-003
		ASRS Coded	NTSB Coded	-1.0600E-002*	1.74000E-003	<.001	-1.5200E-002	-6.0000E-003
			NTSB Augmented	6.4000E-003*	1.74000E-003	.003	1.8000E-003	1.1000E-002
			ASRS Augmented	5.8860E-004	1.74000E-003	.987	-4.0000E-003	5.2000E-003
		ASRS Augmented	NTSB Coded	-1.1200E-002*	1.74000E-003	<.001	-1.5800E-002	-6.6000E-003
			NTSB Augmented	5.8000E-003*	1.74000E-003	.008	1.2000E-003	1.0400E-002
			ASRS Coded	-5.8860E-004	1.74000E-003	.987	-5.2000E-003	4.0000E-003
TX6_CUBE_RT	Tukey HSD	NTSB Coded	NTSB Augmented	1.7500E-002*	1.66000E-003	<.001	1.3200E-002	2.1900E-002
			ASRS Coded	1.4400E-002*	1.66000E-003	<.001	1.0100E-002	1.8800E-002
			ASRS Augmented	1.3000E-002*	1.66000E-003	<.001	8.6000E-003	1.7300E-002
		NTSB Augmented	NTSB Coded	-1.7500E-002*	1.66000E-003	<.001	-2.1900E-002	-1.3200E-002
			ASRS Coded	-3.1000E-003	1.66000E-003	.251	-7.5000E-003	1.3000E-003
			ASRS Augmented	-4.6000E-003*	1.66000E-003	.038	-8.9000E-003	-1.8360E-004
		ASRS Coded	NTSB Coded	-1.4400E-002*	1.66000E-003	<.001	-1.8800E-002	-1.0100E-002
			NTSB Augmented	3.1000E-003	1.66000E-003	.251	-1.3000E-003	7.5000E-003
			ASRS Augmented	-1.5000E-003	1.66000E-003	.816	-5.8000E-003	2.9000E-003
		ASRS Augmented	NTSB Coded	-1.3000E-002*	1.66000E-003	<.001	-1.7300E-002	-8.6000E-003
			NTSB Augmented	4.6000E-003*	1.66000E-003	.038	1.8360E-004	8.9000E-003
			ASRS Coded	1.5000E-003	1.66000E-003	.816	-2.9000E-003	5.8000E-003
TX7_CUBE_RT	Tukey HSD	NTSB Coded	NTSB Augmented	4.7489E-004	2.88567E-004	.361	-2.8631E-004	1.2000E-003
			ASRS Coded	-1.2800E-002*	2.88567E-004	<.001	-1.3500E-002	-1.2000E-002
			ASRS Augmented	4.7489E-004	2.88567E-004	.361	-2.8631E-004	1.2000E-003
		NTSB Augmented	NTSB Coded	-4.7489E-004	2.88567E-004	.361	-1.2000E-003	2.8631E-004
			ASRS Coded	-1.3200E-002*	2.88567E-004	<.001	-1.4000E-002	-1.2500E-002
			ASRS Augmented	0.0000E+000	2.88567E-004	1.000	-7.6119E-004	7.6119E-004
		ASRS Coded	NTSB Coded	1.2800E-002*	2.88567E-004	<.001	1.2000E-002	1.3500E-002
			NTSB Augmented	1.3200E-002*	2.88567E-004	<.001	1.2500E-002	1.4000E-002
			ASRS Augmented	1.3200E-002*	2.88567E-004	<.001	1.2500E-002	1.4000E-002

		ASRS Augmented	NTSB Coded	-4.7489E-004	2.88567E-004	.361	-1.2000E-003	2.8631E-004
			NTSB Augmented	0.0000E+000	2.88567E-004	1.000	-7.6119E-004	7.6119E-004
			ASRS Coded	-1.3200E-002*	2.88567E-004	<.001	-1.4000E-002	-1.2500E-002
TX8_CUBE_RT	Tukey HSD	NTSB Coded	NTSB Augmented	0.0000E+000	9.36797E-004	1.000	-2.5000E-003	2.5000E-003
			ASRS Coded	-5.7000E-003*	9.36797E-004	<.001	-8.1000E-003	-3.2000E-003
			ASRS Augmented	-5.5000E-003*	9.36797E-004	<.001	-8.0000E-003	-3.1000E-003
		NTSB Augmented	NTSB Coded	0.0000E+000	9.36797E-004	1.000	-2.5000E-003	2.5000E-003
			ASRS Coded	-5.7000E-003*	9.36797E-004	<.001	-8.1000E-003	-3.2000E-003
			ASRS Augmented	-5.5000E-003*	9.36797E-004	<.001	-8.0000E-003	-3.1000E-003
		ASRS Coded	NTSB Coded	5.7000E-003*	9.36797E-004	<.001	3.2000E-003	8.1000E-003
			NTSB Augmented	5.7000E-003*	9.36797E-004	<.001	3.2000E-003	8.1000E-003
			ASRS Augmented	1.1989E-004	9.36797E-004	.999	-2.4000E-003	2.6000E-003
		ASRS Augmented	NTSB Coded	5.5000E-003*	9.36797E-004	<.001	3.1000E-003	8.0000E-003
			NTSB Augmented	5.5000E-003*	9.36797E-004	<.001	3.1000E-003	8.0000E-003
			ASRS Coded	-1.1989E-004	9.36797E-004	.999	-2.6000E-003	2.4000E-003

Note. Based on observed means.

The error term is Mean Square(Error) = 7.46E-006.

*. The mean difference is significant at the .05 level.

Figure M3

Parts 121 and 135 Cube Root Transformed MANOVA Estimated Marginal Means Plots

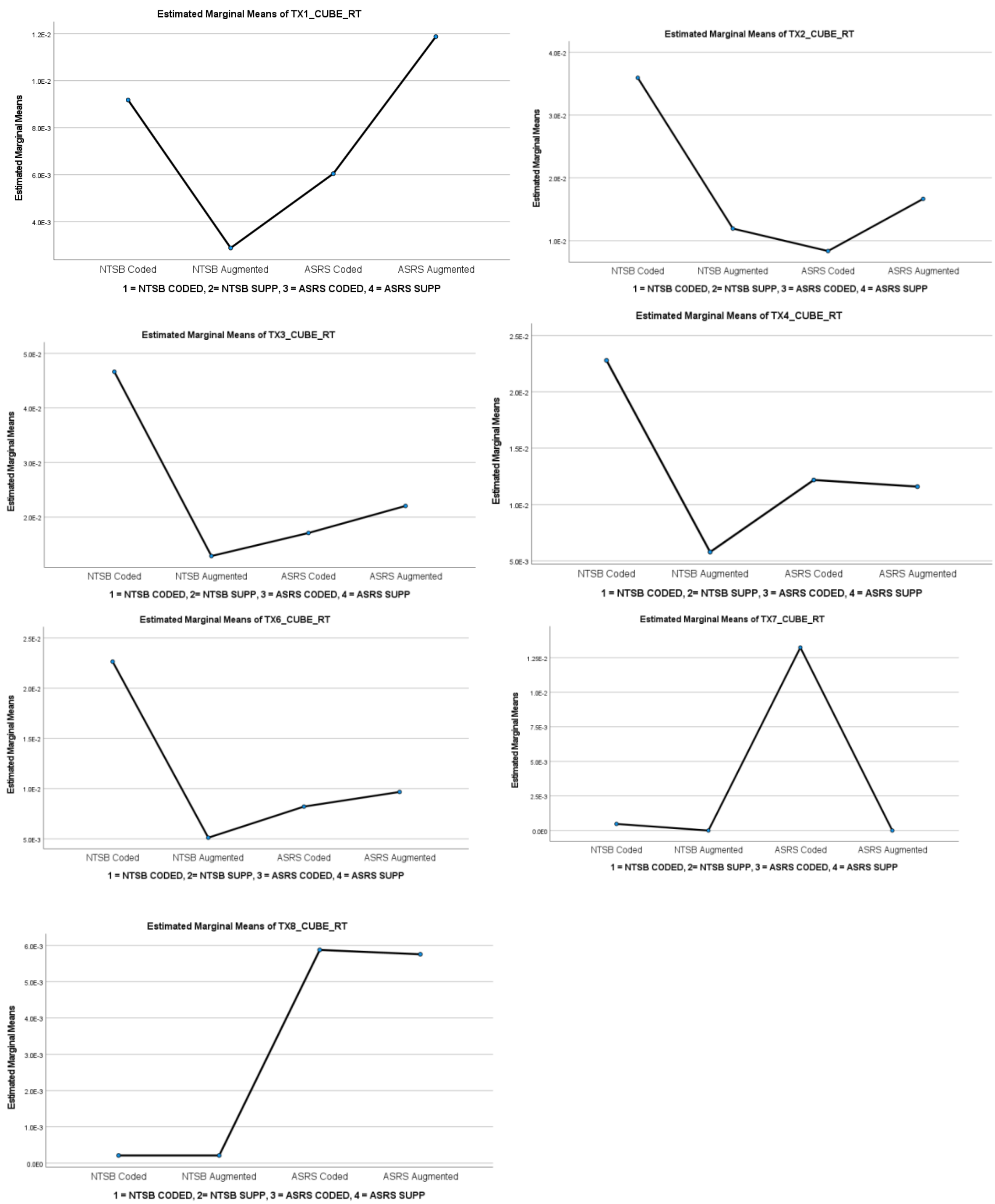


Table M4

Part 91 Square Root Transformed Estimated Means and Standard Deviation Results

Dependent Variable	1 = NTSB CODED, 2= NTSB SUPP, 3 = ASRS CODED, 4 = ASRS SUPP	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
TX1_SQRT	NTSB Coded	1.000E-003	2.754E-004	7.632E-004	2.000E-003
	NTSB Augmented	7.531E-004	2.672E-004	2.191E-004	1.000E-003
	ASRS Coded	8.185E-004	2.672E-004	2.845E-004	1.000E-003
	ASRS Augmented	8.144E-004	2.672E-004	2.804E-004	1.000E-003
TX2_SQRT	NTSB Coded	4.000E-003	5.143E-004	3.000E-003	5.000E-003
	NTSB Augmented	6.000E-003	4.990E-004	5.000E-003	7.000E-003
	ASRS Coded	1.000E-003	4.990E-004	4.549E-004	2.000E-003
	ASRS Augmented	1.000E-003	4.990E-004	3.311E-004	2.000E-003
TX3_SQRT	NTSB Coded	8.000E-003	6.702E-004	7.000E-003	1.000E-002
	NTSB Augmented	7.000E-003	6.502E-004	6.000E-003	9.000E-003
	ASRS Coded	3.000E-003	6.502E-004	2.000E-003	4.000E-003
	ASRS Augmented	3.000E-003	6.502E-004	1.000E-003	4.000E-003
TX4_SQRT	NTSB Coded	3.000E-003	2.831E-004	3.000E-003	4.000E-003
	NTSB Augmented	2.000E-003	2.747E-004	1.000E-003	3.000E-003
	ASRS Coded	2.000E-003	2.747E-004	1.000E-003	2.000E-003
	ASRS Augmented	1.000E-003	2.747E-004	5.338E-004	2.000E-003
TX6_SQRT	NTSB Coded	3.000E-003	5.318E-004	2.000E-003	4.000E-003
	NTSB Augmented	2.000E-003	5.160E-004	7.232E-004	3.000E-003
	ASRS Coded	9.380E-004	5.160E-004	-9.304E-005	2.000E-003
	ASRS Augmented	5.934E-004	5.160E-004	-4.377E-004	2.000E-003
TX7_SQRT	NTSB Coded	2.921E-004	9.420E-005	1.039E-004	4.804E-004
	NTSB Augmented	1.850E-005	9.139E-005	-1.641E-004	2.011E-004
	ASRS Coded	2.000E-003	9.139E-005	2.000E-003	2.000E-003
	ASRS Augmented	1.426E-005	9.139E-005	-1.684E-004	1.969E-004
TX8_SQRT	NTSB Coded	5.330E-004	1.419E-004	2.494E-004	8.166E-004
	NTSB Augmented	0.000E+000	1.377E-004	-2.751E-004	2.751E-004
	ASRS Coded	4.654E-004	1.377E-004	1.902E-004	7.405E-004
	ASRS Augmented	1.239E-004	1.377E-004	-1.512E-004	3.991E-004

Table M5

Part 91 Square Root Transformed MANOVA Post Hoc Results

Dependent Variable		(I) IV Group	(J) IV Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
TX1_ SQRT	Tukey HSD	NTSB Coded	NTSB Augmented	5.61E-004	3.84E-004	4.70E-001	-4.52E-004	0.00E+000
			ASRS Coded	4.95E-004	3.84E-004	5.70E-001	-5.18E-004	0.00E+000
			ASRS Augmented	4.99E-004	3.84E-004	5.70E-001	-5.14E-004	0.00E+000
		NTSB Augmented	NTSB Coded	-5.61E-004	3.84E-004	4.70E-001	0.00E+000	4.52E-004
			ASRS Coded	-6.55E-005	3.78E-004	1.00E+000	0.00E+000	9.32E-004
			ASRS Augmented	-6.13E-005	3.78E-004	1.00E+000	0.00E+000	9.36E-004
		ASRS Coded	NTSB Coded	-4.95E-004	3.84E-004	5.70E-001	0.00E+000	5.18E-004
			NTSB Augmented	6.55E-005	3.78E-004	1.00E+000	-9.32E-004	0.00E+000
			ASRS Augmented	4.13E-006	3.78E-004	1.00E+000	-9.93E-004	0.00E+000
		ASRS Augmented	NTSB Coded	-4.99E-004	3.84E-004	5.70E-001	0.00E+000	5.14E-004
			NTSB Augmented	6.13E-005	3.78E-004	1.00E+000	-9.36E-004	0.00E+000
			ASRS Coded	-4.13E-006	3.78E-004	1.00E+000	0.00E+000	9.93E-004
			NTSB Augmented	6.13E-005	3.78E-004	8.70E-001	-6.94E-004	8.17E-004
			ASRS Coded	-4.13E-006	3.78E-004	9.90E-001	-7.59E-004	7.51E-004
TX2_ SQRT	Tukey HSD	NTSB Coded	NTSB Augmented	0.00E+000*	7.17E-004	3.00E-002	0.00E+000	-1.70E-004
			ASRS Coded	0.00E+000*	7.17E-004	0.00E+000	6.54E-004	0.00E+000
			ASRS Augmented	0.00E+000*	7.17E-004	0.00E+000	7.78E-004	0.00E+000
		NTSB Augmented	NTSB Coded	0.00E+000*	7.17E-004	3.00E-002	1.70E-004	0.00E+000
			ASRS Coded	0.00E+000*	7.06E-004	<.001	0.00E+000	1.00E-002
			ASRS Augmented	0.00E+000*	7.06E-004	<.001	0.00E+000	1.00E-002
		ASRS Coded	NTSB Coded	0.00E+000*	7.17E-004	0.00E+000	0.00E+000	-6.54E-004

			NTSB Augmented	0.00E+000*	7.06E-004	<.001	-1.00E-002	0.00E+000
			ASRS Augmented	1.24E-004	7.06E-004	1.00E+000	0.00E+000	0.00E+000
		ASRS Augmented	NTSB Coded	0.00E+000*	7.17E-004	0.00E+000	0.00E+000	-7.78E-004
			NTSB Augmented	0.00E+000*	7.06E-004	<.001	-1.00E-002	0.00E+000
			ASRS Coded	-1.24E-004	7.06E-004	1.00E+000	0.00E+000	0.00E+000
			NTSB Augmented	0.00E+000*	7.06E-004	<.001	-1.00E-002	0.00E+000
			ASRS Coded	-1.24E-004	7.06E-004	8.60E-001	0.00E+000	0.00E+000
TX3_ SQRT	Tukey HSD	NTSB Coded	NTSB Augmented	9.99E-004	9.34E-004	7.10E-001	0.00E+000	0.00E+000
			ASRS Coded	1.00E-002*	9.34E-004	<.001	0.00E+000	1.00E-002
			ASRS Augmented	1.00E-002*	9.34E-004	<.001	0.00E+000	1.00E-002
		NTSB Augmented	NTSB Coded	-9.99E-004	9.34E-004	7.10E-001	0.00E+000	0.00E+000
			ASRS Coded	0.00E+000*	9.20E-004	<.001	0.00E+000	1.00E-002
			ASRS Augmented	0.00E+000*	9.20E-004	<.001	0.00E+000	1.00E-002
		ASRS Coded	NTSB Coded	-1.00E-002*	9.34E-004	<.001	-1.00E-002	0.00E+000
			NTSB Augmented	0.00E+000*	9.20E-004	<.001	-1.00E-002	0.00E+000
			ASRS Augmented	2.86E-004	9.20E-004	9.90E-001	0.00E+000	0.00E+000
		ASRS Augmented	NTSB Coded	-1.00E-002*	9.34E-004	<.001	-1.00E-002	0.00E+000
			NTSB Augmented	0.00E+000*	9.20E-004	<.001	-1.00E-002	0.00E+000
			ASRS Coded	-2.86E-004	9.20E-004	9.90E-001	0.00E+000	0.00E+000
			NTSB Augmented	0.00E+000*	9.20E-004	<.001	-1.00E-002	0.00E+000
			ASRS Coded	-2.86E-004	9.20E-004	7.60E-001	0.00E+000	0.00E+000
TX4_ SQRT	Tukey HSD	NTSB Coded	NTSB Augmented	0.00E+000*	3.94E-004	2.00E-002	1.41E-004	0.00E+000
			ASRS Coded	0.00E+000*	3.94E-004	0.00E+000	3.78E-004	0.00E+000
			ASRS Augmented	0.00E+000*	3.94E-004	<.001	0.00E+000	0.00E+000
			NTSB Coded	0.00E+000*	3.94E-004	2.00E-002	0.00E+000	-1.41E-004

		NTSB Augmented	ASRS Coded	2.38E-004	3.88E-004	9.30E-001	-7.87E-004	0.00E+000
			ASRS Augmented	9.58E-004	3.88E-004	8.00E-002	-6.69E-005	0.00E+000
		ASRS Coded	NTSB Coded	0.00E+000*	3.94E-004	0.00E+000	0.00E+000	-3.78E-004
			NTSB Augmented	-2.38E-004	3.88E-004	9.30E-001	0.00E+000	7.87E-004
			ASRS Augmented	7.21E-004	3.88E-004	2.60E-001	-3.04E-004	0.00E+000
		ASRS Augmented	NTSB Coded	0.00E+000*	3.94E-004	<.001	0.00E+000	0.00E+000
			NTSB Augmented	-9.58E-004	3.88E-004	8.00E-002	0.00E+000	6.69E-005
			ASRS Coded	-7.21E-004	3.88E-004	2.60E-001	0.00E+000	3.04E-004
			NTSB Augmented	-9.58E-004*	3.88E-004	2.00E-002	0.00E+000	-1.82E-004
			ASRS Coded	-7.21E-004	3.88E-004	7.00E-002	0.00E+000	5.57E-005
TX6_ SQRT	Tukey HSD	NTSB Coded	NTSB Augmented	0.00E+000	7.41E-004	1.90E-001	-4.54E-004	0.00E+000
			ASRS Coded	0.00E+000*	7.41E-004	1.00E-002	3.62E-004	0.00E+000
			ASRS Augmented	0.00E+000*	7.41E-004	0.00E+000	7.07E-004	0.00E+000
		NTSB Augmented	NTSB Coded	0.00E+000	7.41E-004	1.90E-001	0.00E+000	4.54E-004
			ASRS Coded	8.16E-004	7.30E-004	6.80E-001	0.00E+000	0.00E+000
			ASRS Augmented	0.00E+000	7.30E-004	3.90E-001	-7.65E-004	0.00E+000
		ASRS Coded	NTSB Coded	0.00E+000*	7.41E-004	1.00E-002	0.00E+000	-3.62E-004
			NTSB Augmented	-8.16E-004	7.30E-004	6.80E-001	0.00E+000	0.00E+000
			ASRS Augmented	3.45E-004	7.30E-004	9.60E-001	0.00E+000	0.00E+000
		ASRS Augmented	NTSB Coded	0.00E+000*	7.41E-004	0.00E+000	0.00E+000	-7.07E-004
			NTSB Augmented	0.00E+000	7.30E-004	3.90E-001	0.00E+000	7.65E-004
			ASRS Coded	-3.45E-004	7.30E-004	9.60E-001	0.00E+000	0.00E+000
			NTSB Augmented	0.00E+000	7.30E-004	1.20E-001	0.00E+000	2.97E-004
			ASRS Coded	-3.45E-004	7.30E-004	6.40E-001	0.00E+000	0.00E+000
TX7_ SQRT	Tukey HSD	NTSB Coded	NTSB Augmented	2.74E-004	1.31E-004	1.70E-001	-7.27E-005	6.20E-004

			ASRS Coded	0.00E+000*	1.31E-004	<.001	0.00E+000	0.00E+000
			ASRS Augmented	2.78E-004	1.31E-004	1.60E-001	-6.85E-005	6.24E-004
		NTSB Augmented	NTSB Coded	-2.74E-004	1.31E-004	1.70E-001	-6.20E-004	7.27E-005
			ASRS Coded	0.00E+000*	1.29E-004	<.001	0.00E+000	0.00E+000
			ASRS Augmented	4.24E-006	1.29E-004	1.00E+000	-3.37E-004	3.45E-004
		ASRS Coded	NTSB Coded	0.00E+000*	1.31E-004	<.001	0.00E+000	0.00E+000
			NTSB Augmented	0.00E+000*	1.29E-004	<.001	0.00E+000	0.00E+000
			ASRS Augmented	0.00E+000*	1.29E-004	<.001	0.00E+000	0.00E+000
		ASRS Augmented	NTSB Coded	-2.78E-004	1.31E-004	1.60E-001	-6.24E-004	6.85E-005
			NTSB Augmented	-4.24E-006	1.29E-004	1.00E+000	-3.45E-004	3.37E-004
			ASRS Coded	0.00E+000*	1.29E-004	<.001	0.00E+000	0.00E+000
TX8_ SQRT	Tukey HSD	NTSB Coded	NTSB Augmented	5.33E-004*	1.98E-004	4.00E-002	1.12E-005	0.00E+000
			ASRS Coded	6.76E-005	1.98E-004	9.90E-001	-4.54E-004	5.89E-004
			ASRS Augmented	4.09E-004	1.98E-004	1.70E-001	-1.13E-004	9.31E-004
		NTSB Augmented	NTSB Coded	-5.33E-004*	1.98E-004	4.00E-002	0.00E+000	-1.12E-005
			ASRS Coded	-4.65E-004	1.95E-004	9.00E-002	-9.79E-004	4.84E-005
			ASRS Augmented	-1.24E-004	1.95E-004	9.20E-001	-6.38E-004	3.90E-004
		ASRS Coded	NTSB Coded	-6.76E-005	1.98E-004	9.90E-001	-5.89E-004	4.54E-004
			NTSB Augmented	4.65E-004	1.95E-004	9.00E-002	-4.84E-005	9.79E-004
			ASRS Augmented	3.41E-004	1.95E-004	3.10E-001	-1.72E-004	8.55E-004
		ASRS Augmented	NTSB Coded	-4.09E-004	1.98E-004	1.70E-001	-9.31E-004	1.13E-004
			NTSB Augmented	1.24E-004	1.95E-004	9.20E-001	-3.90E-004	6.38E-004
			ASRS Coded	-3.41E-004	1.95E-004	3.10E-001	-8.55E-004	1.72E-004

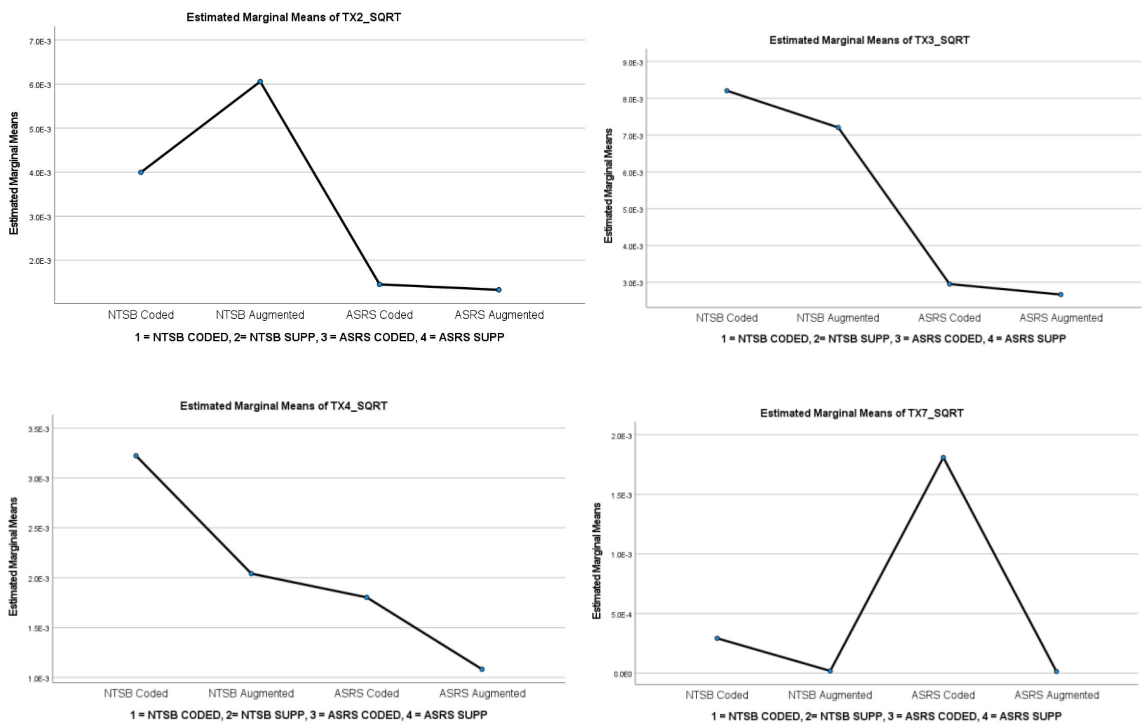
Based on observed means.

The error term is Mean Square(Error) = 3.222E-7.

*. The mean difference is significant at the .05 level.

Figure M6

Part 91 Square Root Transformed MANOVA Estimated Marginal Means Plots



Note. Only plots with significant univariate difference are captured

Appendix N

Parts 121 and 135 Discriminant Analysis Results

Table N1

Equality of Group Means Result

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
TX1_CUBE_RT	.386	33.997	3	64	<.001
TX2_CUBE_RT	.458	25.214	3	64	<.001
TX3_CUBE_RT	.383	34.427	3	64	<.001
TX4_CUBE_RT	.393	32.980	3	64	<.001
TX6_CUBE_RT	.329	43.492	3	64	<.001
TX7_CUBE_RT	.020	1028.316	3	64	<.001
TX8_CUBE_RT	.472	23.869	3	64	<.001

Table N2

Pooled Within-Groups Matrices Result

Pooled Within-Groups Matrices

		TX1_CUBE_RT	TX2_CUBE_RT	TX3_CUBE_RT	TX4_CUBE_RT	TX6_CUBE_RT	TX7_CUBE_RT	TX8_CUBE_RT
Correlation	TX1_CUBE_RT	1.000	.512	.355	.231	.057	-.178	-.261
	TX2_CUBE_RT	.512	1.000	.931	.672	.659	-.429	-.105
	TX3_CUBE_RT	.355	.931	1.000	.797	.755	-.440	-.033
	TX4_CUBE_RT	.231	.672	.797	1.000	.588	-.227	-.150
	TX6_CUBE_RT	.057	.659	.755	.588	1.000	-.383	.365
	TX7_CUBE_RT	-.178	-.429	-.440	-.227	-.383	1.000	.060
	TX8_CUBE_RT	-.261	-.105	-.033	-.150	.365	.060	1.000

Table N3

Standard Canonical Discriminant Function Coefficients Result

Standardized Canonical Discriminant Function Coefficients

	Function		
	1	2	3
TX1_CUBE_RT	.074	1.124	.219
TX2_CUBE_RT	-.347	-1.720	.331
TX3_CUBE_RT	.905	1.024	-.160
TX4_CUBE_RT	-.396	.273	.192
TX6_CUBE_RT	.237	.418	-1.251
TX7_CUBE_RT	1.161	-.046	-.100
TX8_CUBE_RT	-.128	.296	.974

Table N4

Structure Matrix Result

Structure Matrix

	Function		
	1	2	3
TX7_CUBE_RT	.889*	-.164	.284
TX1_CUBE_RT	-.033	.625*	.068
TX6_CUBE_RT	-.029	.408	-.634*
TX3_CUBE_RT	-.035	.365	-.553*
TX4_CUBE_RT	.005	.405	-.521*
TX2_CUBE_RT	-.060	.256	-.461*
TX8_CUBE_RT	.075	.259	.395*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

*. Largest absolute correlation between each variable and any discriminant function

Table N5

Classification Results

Classification Results[a,c]

		1 = NTSB CODED, 2= NTSB SUPP, 3 = ASRS CODED, 4 = ASRS SUPP	Predicted Group Membership				Total
			NTSB Coded	NTSB Augmented	ASRS Coded	ASRS Augmented	
Original	Count	NTSB Coded	13	4	0	0	17
		NTSB Augmented	0	17	0	0	17
		ASRS Coded	0	0	17	0	17
		ASRS Augmented	0	0	0	17	17
	%	NTSB Coded	76.5	23.5	.0	.0	100.0
		NTSB Augmented	.0	100.0	.0	.0	100.0
		ASRS Coded	.0	.0	100.0	.0	100.0
		ASRS Augmented	.0	.0	.0	100.0	100.0
Cross-validated[b]	Count	NTSB Coded	13	4	0	0	17
		NTSB Augmented	1	16	0	0	17
		ASRS Coded	0	0	17	0	17
		ASRS Augmented	0	0	0	17	17
	%	NTSB Coded	76.5	23.5	.0	.0	100.0
		NTSB Augmented	5.9	94.1	.0	.0	100.0
		ASRS Coded	.0	.0	100.0	.0	100.0
		ASRS Augmented	.0	.0	.0	100.0	100.0

a. 94.1% of original grouped cases correctly classified.

b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

c. 92.6% of cross-validated grouped cases correctly classified.

Figure N6

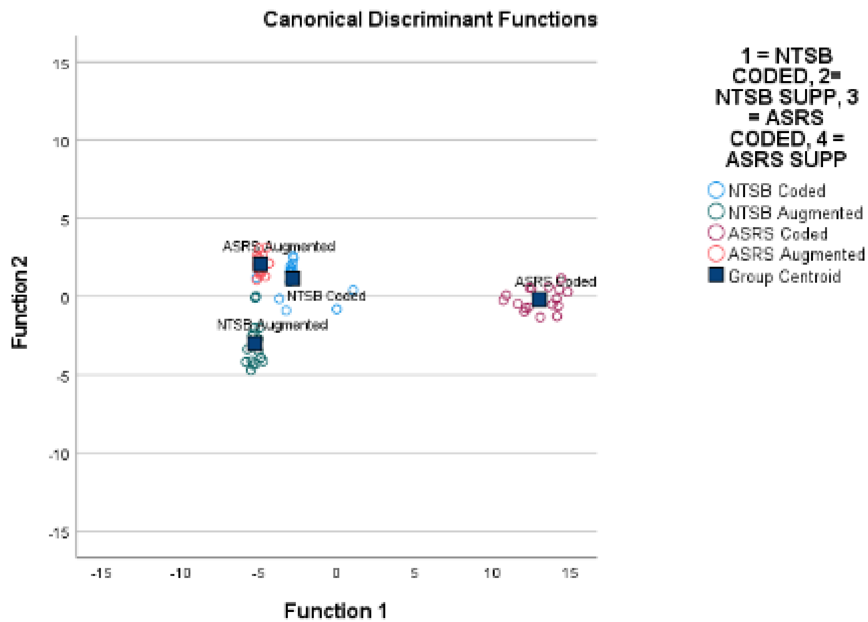
Territorial Map Result

```
                                            Territorial Map
                          (Assuming all functions but the first two are zero)
Canonical Discriminant
Function 2
      -16.0     -12.0      -8.0      -4.0        .0       4.0       8.0      12.0      16.0
          +---------+---------+---------+---------+---------+---------+---------+---------+
    16.0 +                                                 41    13                        +
         I                                                41     13                        I
         I                                                41     13                        I
         I                                               41      13                        I
         I                                              41       13                        I
         I                                             41        13                        I
    12.0 +         +         +         +         +    41   +    13   +         +          +
         I                                            41        13                         I
         I                                           41         13                         I
         I                                          41          13                         I
         I                                          41          13                         I
         I                                         41           13                         I
     8.0 +         +         +         +         41        +   13    +         +          +
         I                                       41             13                         I
         I                                       41            13                          I
         I                                      41             13                          I
         I                                     41              13                          I
         I                                    41               13                          I
     4.0 +         +         +         +    41  +         +  13     +         +          +
         I                                   41                13                          I
         I                                  41                 13                          I
         I                            *    41                 13                           I
         I                                 *1                 13                           I
         I4444444444444444                41                  13                           I
      .0 +222222222222222244444444444444441       +         + 13      +         + *        +
         I                2222222222222222111                 13                           I
         I                                222111              13                           I
         I                                   222111           13                           I
         I                                      222111       13                            I
         I                           *             222111    13                            I
    -4.0 +         +         +         +         +   22211 +13       +         +          +
         I                                               221113                            I
         I                                                 2223                            I
         I                                                   23                            I
         I                                                   23                            I
         I                                                    23                           I
    -8.0 +         +         +         +         +         + 23      +         +          +
         I                                                    23                           I
         I                                                     23                          I
         I                                                     23                          I
         I                                                     23                          I
         I                                                     23                          I
   -12.0 +         +         +         +         +         +   23    +         +          +
         I                                                      23                         I
         I                                                      23                         I
         I                                                      23                         I
         I                                                       23                        I
         I                                                       23                        I
   -16.0 +                                                       23                        +
          +---------+---------+---------+---------+---------+---------+---------+---------+
      -16.0     -12.0      -8.0      -4.0        .0       4.0       8.0      12.0      16.0
                                   Canonical Discriminant Function 1
```

```
Symbols used in territorial map

Symbol  Group  Label
------  -----  --------------------

   1       1   NTSB Coded
   2       2   NTSB Augmented
   3       3   ASRS Coded
   4       4   ASRS Augmented
   *           Indicates a group centroid
```

Figure N7

Graph Showing Canonical Discriminant Functions



Appendix O

Part 91 Discriminant Analysis Results

Table O1

Equality of Group Means Result

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
TX1_SQRT	.924	1.746	3	64	.167
TX2_SQRT	.507	20.774	3	64	<.001
TX3_SQRT	.515	20.079	3	64	<.001
TX4_SQRT	.654	11.294	3	64	<.001
TX6_SQRT	.775	6.192	3	64	<.001
TX7_SQRT	.221	75.039	3	64	<.001
TX8_SQRT	.840	4.073	3	64	.010

Table O2

Pooled Within-Groups Matrices Result

Pooled Within-Groups Matrices

		TX1_SQRT	TX2_SQRT	TX3_SQRT	TX4_SQRT	TX6_SQRT	TX7_SQRT	TX8_SQRT
Correlation	TX1_SQRT	1.000	.135	.641	.895	.968	.847	.939
	TX2_SQRT	.135	1.000	.814	.467	.132	-.073	-.097
	TX3_SQRT	.641	.814	1.000	.880	.655	.430	.466
	TX4_SQRT	.895	.467	.880	1.000	.913	.716	.803
	TX6_SQRT	.968	.132	.655	.913	1.000	.825	.957
	TX7_SQRT	.847	-.073	.430	.716	.825	1.000	.825
	TX8_SQRT	.939	-.097	.466	.803	.957	.825	1.000

Table O3

Standard Canonical Discriminant Function Coefficients Result

Standardized Canonical Discriminant Function Coefficients

	Function		
	1	2	3
TX1_SQRT	-1.885	.835	-3.241
TX2_SQRT	2.718	3.514	.697
TX3_SQRT	-3.195	-3.211	.358
TX4_SQRT	1.480	-2.294	-.862
TX6_SQRT	-1.963	2.245	3.517
TX7_SQRT	1.945	.068	.460
TX8_SQRT	2.633	.458	-.037

Table O4

Structure Matrix Result

Structure Matrix

	Function		
	1	2	3
TX7_SQRT	.392*	-.273	.073
TX8_SQRT	.025	-.234*	.069
TX1_SQRT	-.022	-.142*	.071
TX2_SQRT	-.100	.191	.583*
TX3_SQRT	-.112	-.120	.575*
TX4_SQRT	-.036	-.266	.387*
TX6_SQRT	-.049	-.189	.261*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

*. Largest absolute correlation between each variable and any discriminant function

Table O5

Classification Results

Classification Results[a]

		1 = NTSB CODED, 2= NTSB SUPP, 3 = ASRS CODED, 4 = ASRS SUPP	Predicted Group Membership				Total
			NTSB Coded	NTSB Augmented	ASRS Coded	ASRS Augmented	
Original	Count	NTSB Coded	16	1	0	0	17
		NTSB Augmented	0	16	0	1	17
		ASRS Coded	0	0	17	0	17
		ASRS Augmented	0	0	0	17	17
	%	NTSB Coded	94.1	5.9	.0	.0	100.0
		NTSB Augmented	.0	94.1	.0	5.9	100.0
		ASRS Coded	.0	.0	100.0	.0	100.0
		ASRS Augmented	.0	.0	.0	100.0	100.0

a. 97.1% of original grouped cases correctly classified.

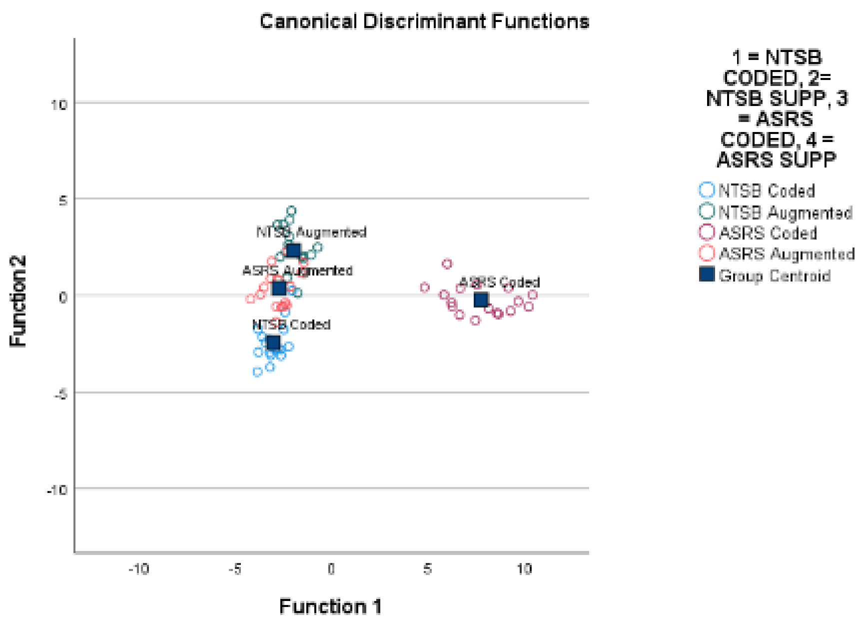
Figure O6

Territorial Map Result

```
                                            Territorial Map
                          (Assuming all functions but the first two are zero)
Canonical Discriminant
Function 2
      -16.0     -12.0      -8.0      -4.0        .0       4.0       8.0      12.0      16.0
          +---------+---------+---------+---------+---------+---------+---------+---------+
    16.0 +                                                         23                      +
         I                                                         23                      I
         I                                                        23                       I
         I                                                        23                       I
         I                                                       23                        I
         I                                                       23                        I
    12.0 +         +         +         +         +         +   23    +         +          +
         I                                                      23                         I
         I                                                      23                         I
         I                                                     23                          I
         I                                                     23                          I
         I                                                    23                           I
     8.0 +         +         +         +         +         + 23      +         +          +
         I                                                   23                            I
         I                                                   23                            I
         I22                                                23                             I
         I442222                                            23                             I
         I  444422222                                       23                             I
     4.0 +      444442222     +         +         +        23        +         +          +
         I           44442222                              23                              I
         I               444422222                        23                               I
         I                   444442222       *            23                               I
         I44                      44442222               23                                I
         I1144444444444444444         44442*222          23                                I
      .0 +  11111111111111114444444444444444444442222      23 +        *+         +          +
         I                  111111111111111444444442222223                                 I
         I                                 111111111111123                                 I
         I                                              13                                 I
         I                                *             13                                 I
         I                                               13                                I
    -4.0 +         +         +         +         +      13 +         +         +          +
         I                                               13                                I
         I                                                13                               I
         I                                                13                               I
         I                                                13                               I
         I                                                 13                              I
    -8.0 +         +         +         +         +        13         +         +          +
         I                                                 13                              I
         I                                                  13                             I
         I                                                  13                             I
         I                                                  13                             I
         I                                                   13                            I
   -12.0 +         +         +         +         +         +13       +         +          +
         I                                                   13                            I
         I                                                    13                           I
         I                                                    13                           I
         I                                                    13                           I
         I                                                     13                          I
   -16.0 +                                                     13                          +
          +---------+---------+---------+---------+---------+---------+---------+---------+
      -16.0     -12.0      -8.0      -4.0        .0       4.0       8.0      12.0      16.0
                                   Canonical Discriminant Function 1


Symbols used in territorial map

Symbol  Group  Label
------  -----  --------------------

   1        1  NTSB Coded
   2        2  NTSB Augmented
   3        3  ASRS Coded
   4        4  ASRS Augmented
   *           Indicates a group centroid
```

Figure O7

Graph Showing Canonical Discriminant Functions



Appendix P

Qualitative Analysis Results

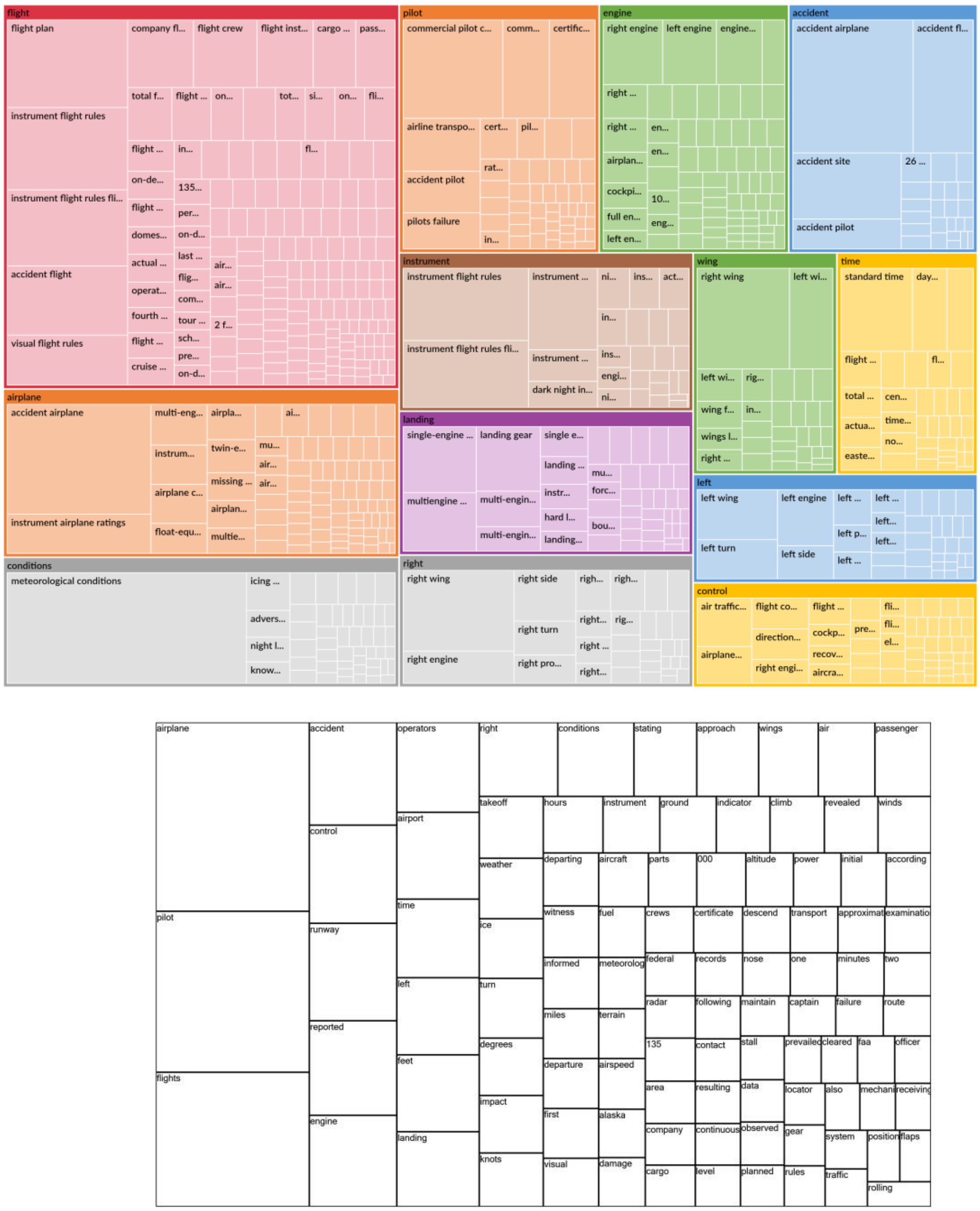
Table P1

Sources of Extracts from NTSB, AIDS, and ASRS Narratives

Database	Narratives Source
NTSB database	First run: combined: summary of event (Narr_accf), factual report of event (narr_accp) and cause of event (narr_cause)
ASRS database	Combined Report 1 & 2: Synopsis and narrative
AIDS database	Report Narrative

Figure P2

Qualitative Analysis NVivo® Tree Map and Hierarchy Chart for NTSB Part 121 & 135 Classified Dataset

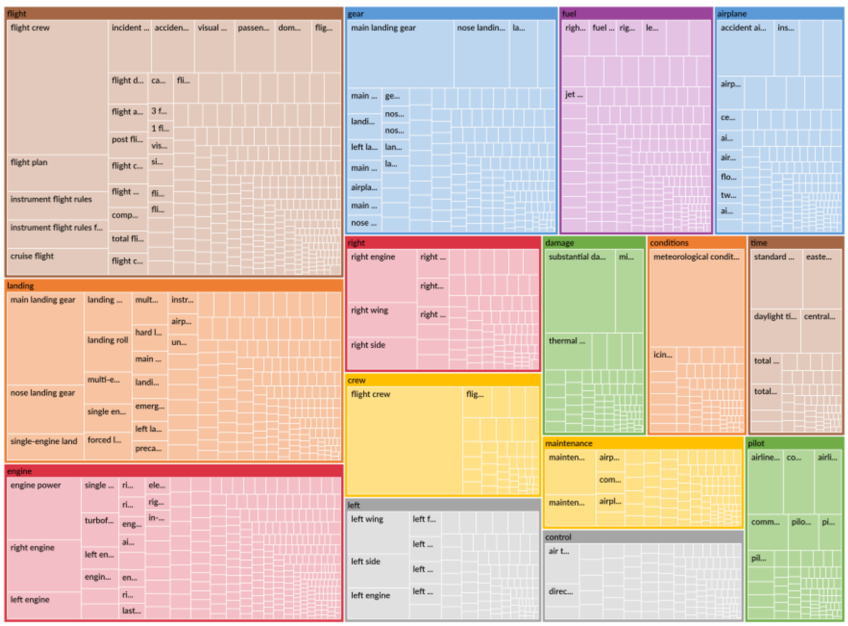


Appendix O-3

Qualitative Analysis NVivo® Tree Map for NTSB Part 121 & 135 Augmented Dataset

Figure P3

Qualitative Analysis NVivo® Tree Map for NTSB Part 121 & 135 Augmented Dataset



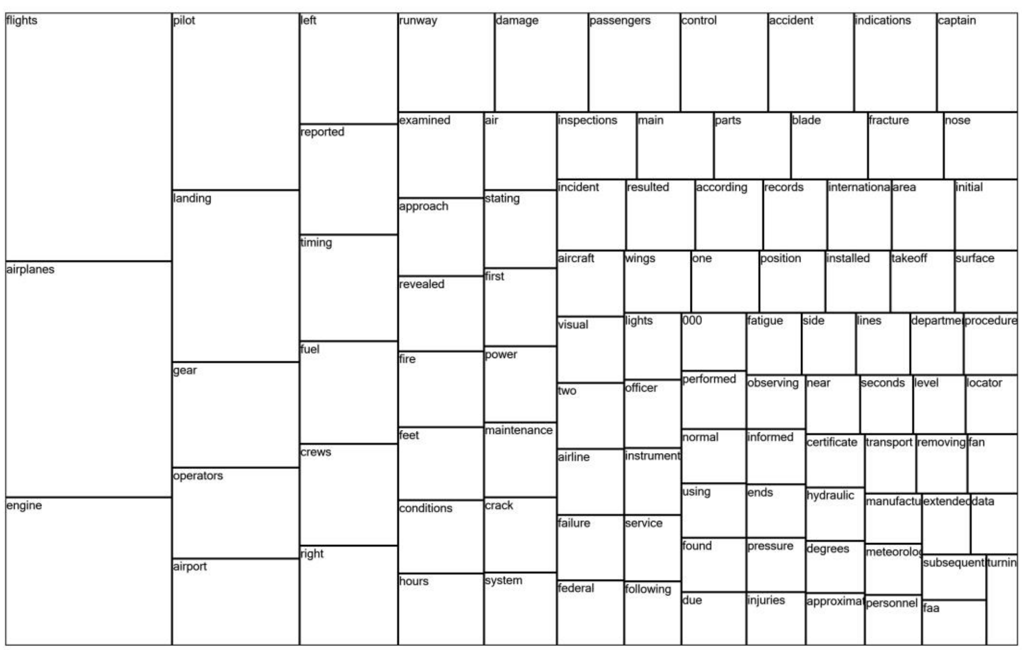
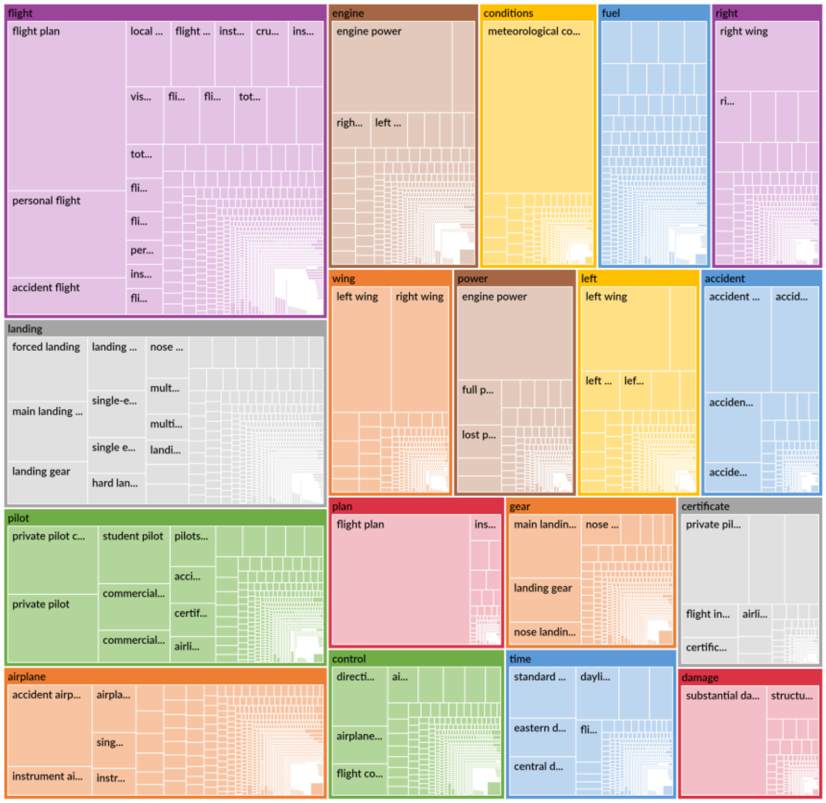


Figure P4

Qualitative Analysis NVivo® Tree Map for NTSB Part 91 Classified Dataset



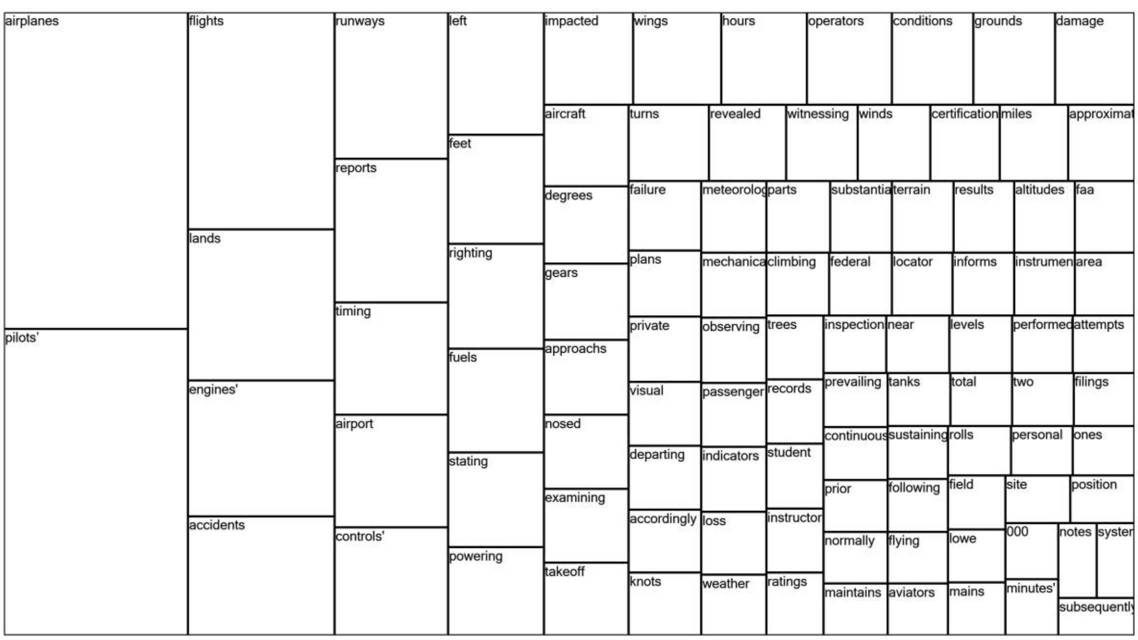
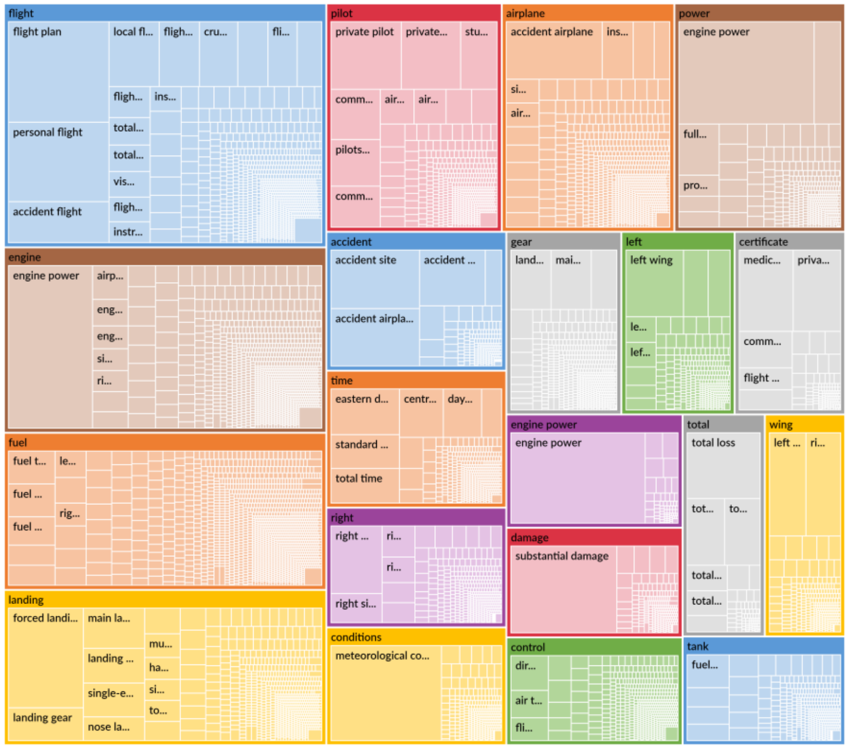


Figure P5

Qualitative Analysis NVivo® Tree Map for NTSB Part 91 Augmented Dataset



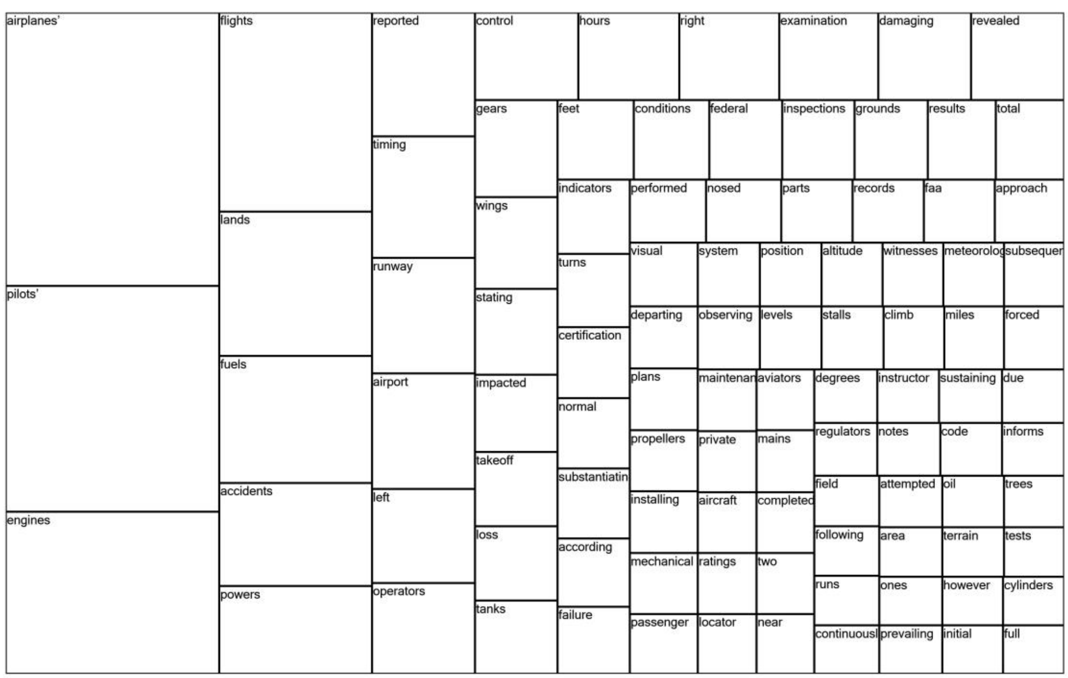


Figure P6

Qualitative Analysis NVivo® Tree Map for ASRS Parts 121 and 135 Classified Dataset

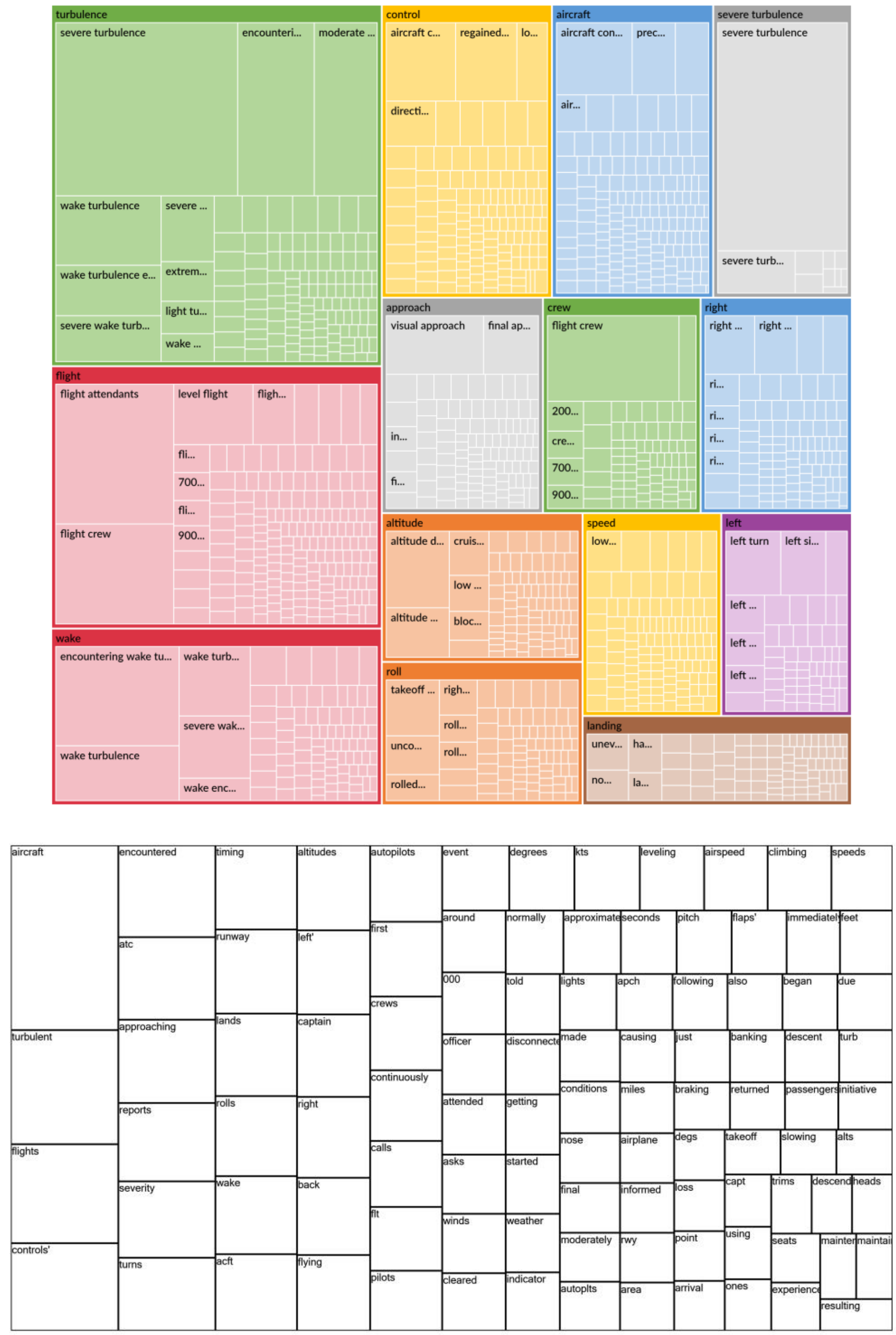


Figure P7

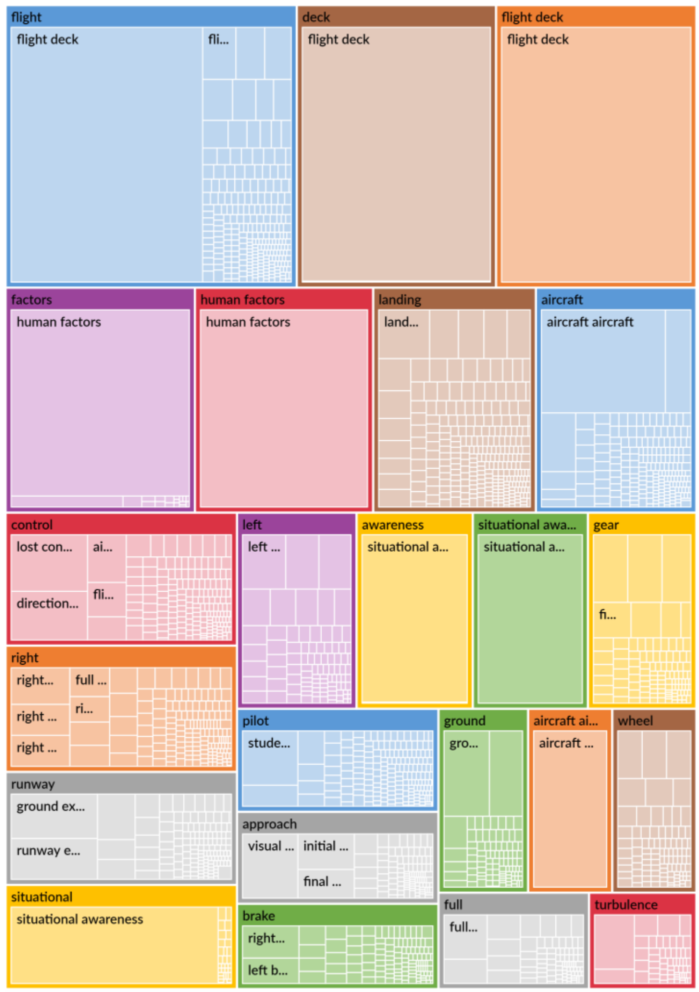
Qualitative Analysis NVivo® Tree Map for ASRS Parts 121 and 135 Augmented Dataset

Figure P8

Qualitative Analysis NVivo® Tree Map for ASRS Parts 91 Classified Dataset

Figure P9

Qualitative Analysis NVivo® Tree Map for ASRS Parts 91 Augmented Dataset



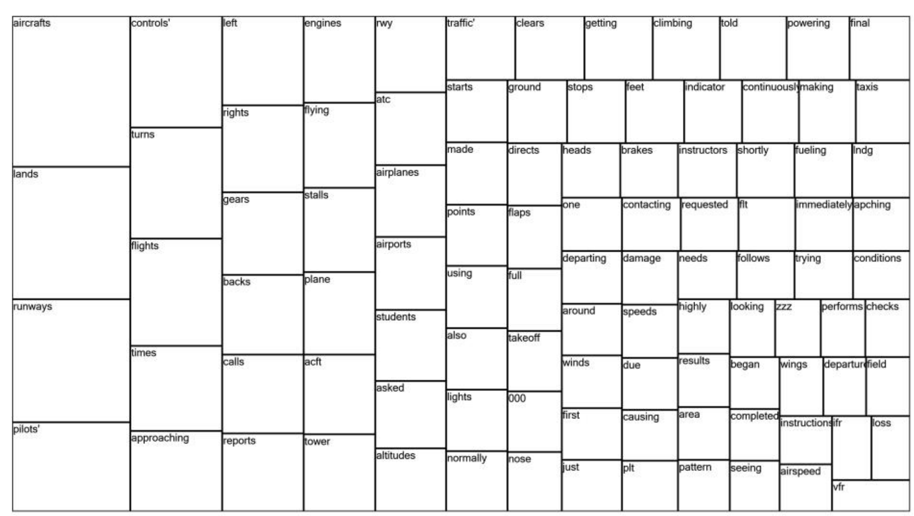


Figure P10

Qualitative Analysis NVivo® Tree Map for AIDS Parts 121 and 135 Classified Dataset

Figure P11

Qualitative Analysis NVivo® Tree Map for AIDS Parts 121 and 135 Augmented Dataset

Figure P12

Qualitative Analysis NVivo® Tree Map for AIDS Part 91 Classified Dataset

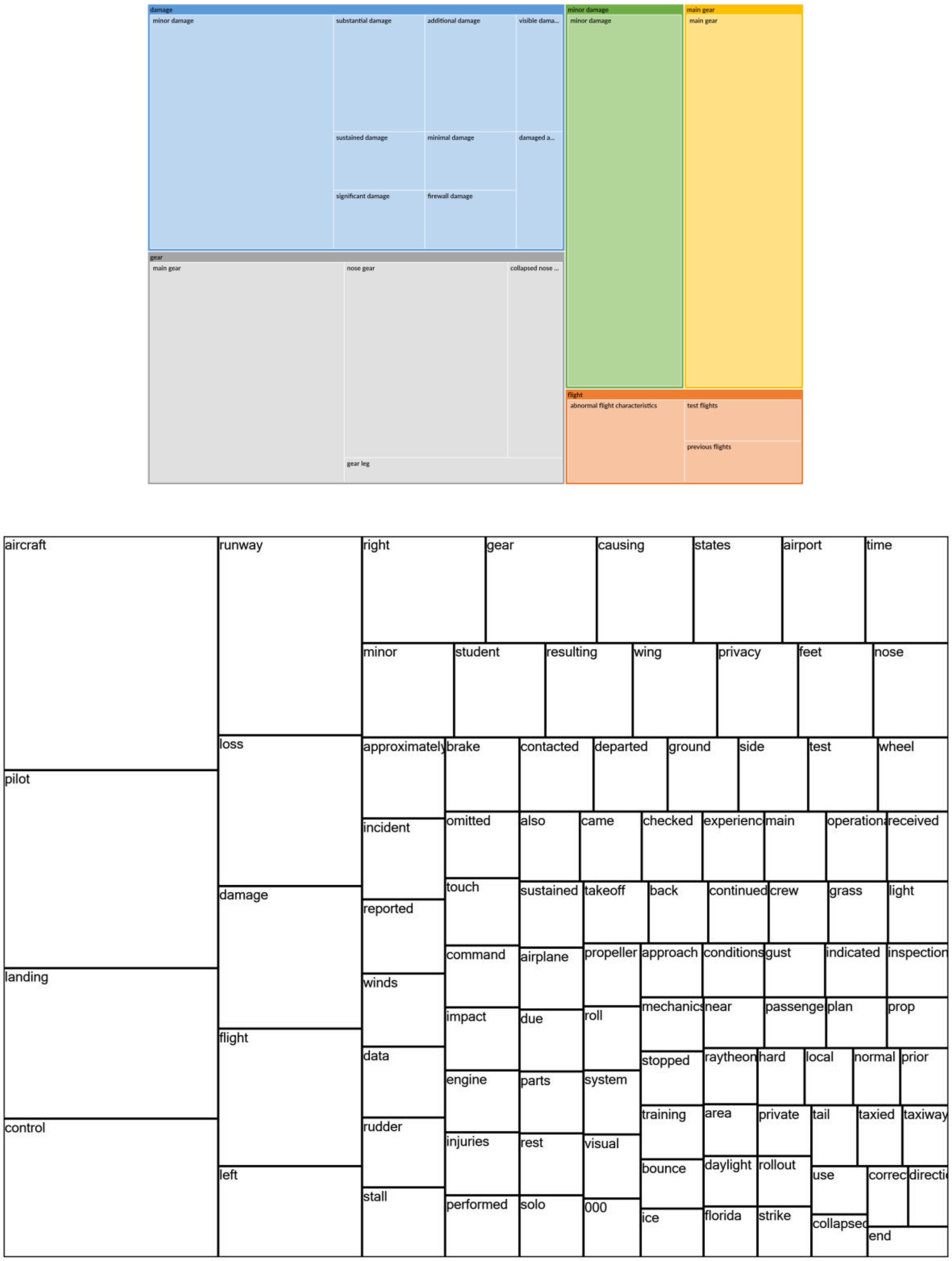
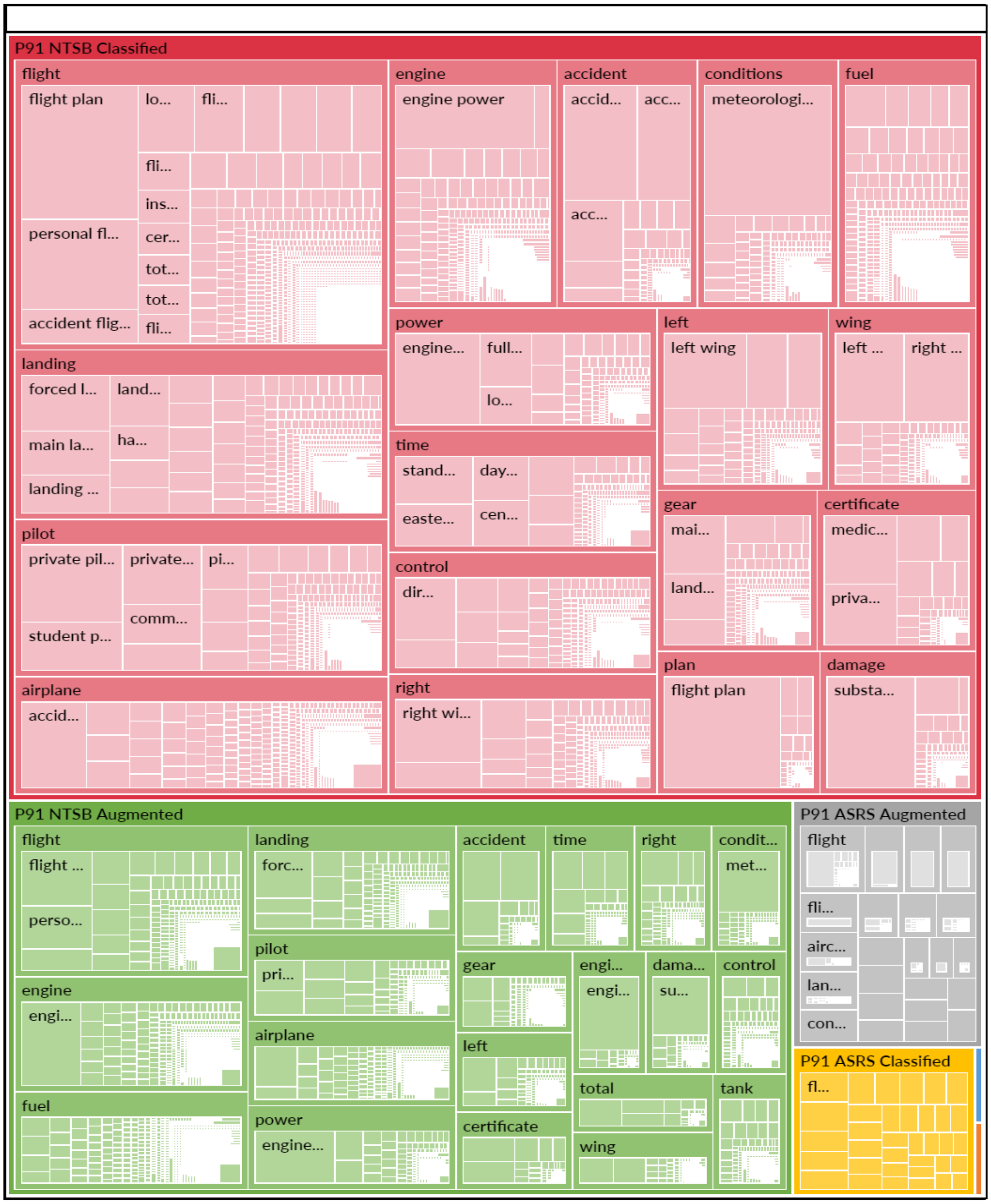


Figure P13

Qualitative Analysis NVivo® Tree Map for AIDS Part 91 Augmented Dataset

Figure P14

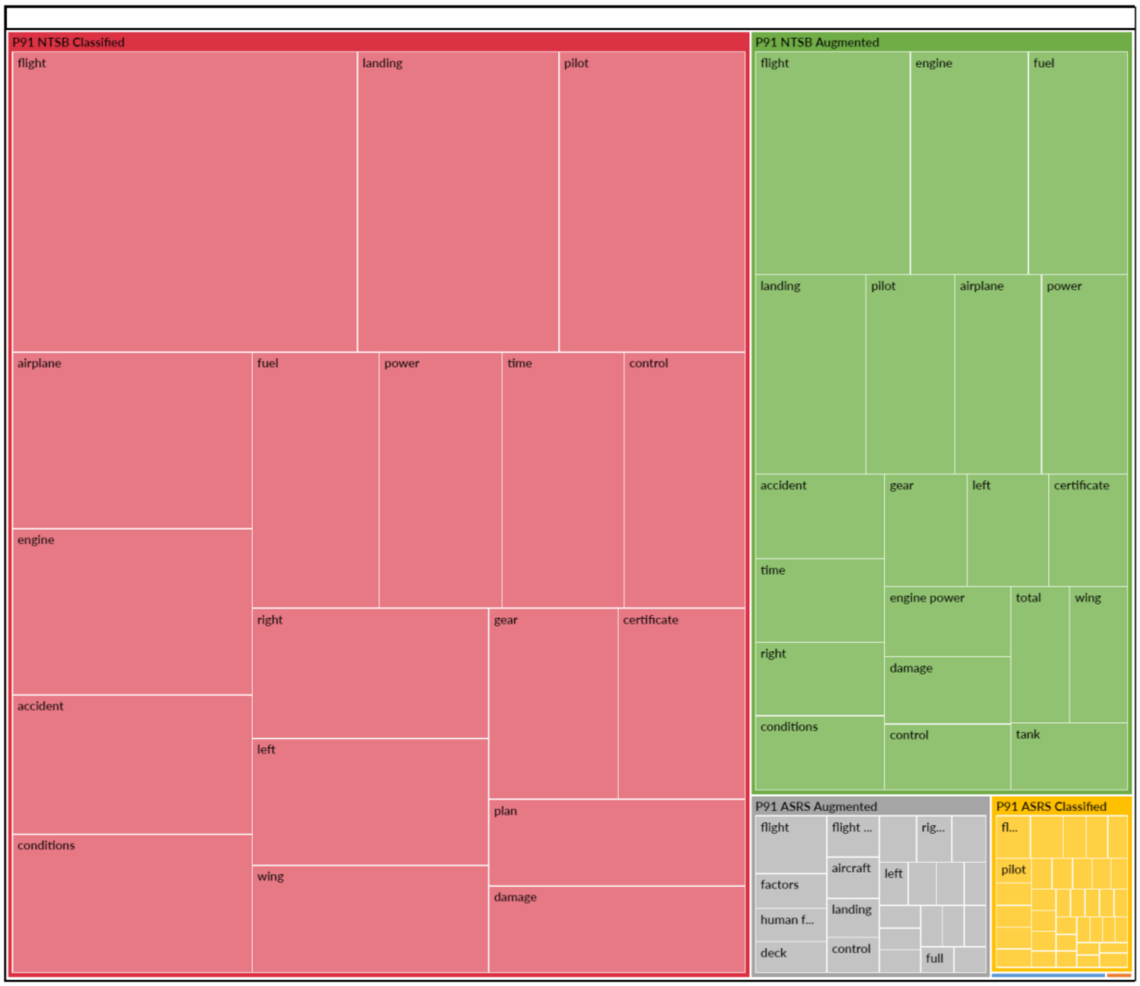
Qualitative Analysis NVivo® Hierarchy Chart for Combined ASRS, NTSB and AIDS Part 91 Datasets – Level three diagram



Note. The bottom right blue and orange color represent the AIDS dataset with minimal data size.

Figure P15

Qualitative Analysis NVivo® Hierarchy Chart for Combined ASRS, NTSB and AIDS Part 91 Datasets – Level Two Diagram



Note. The bottom right blue and orange color represent the AIDS dataset with minimal data size.

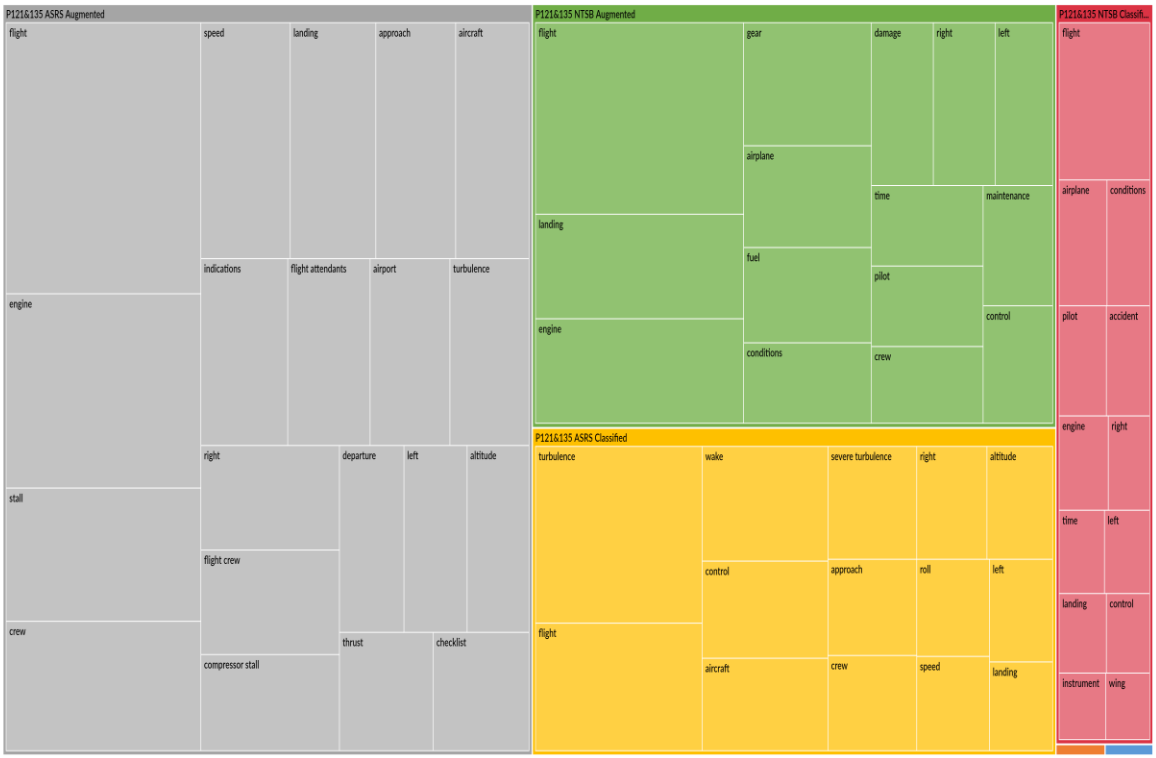
Figure P16

Qualitative Analysis NVivo® Hierarchy Chart for Combined ASRS, NTSB and AIDS Parts 121 and 135 dataset – Level Three Diagram

Note. The bottom right blue and orange color represent the AIDS dataset with minimal data size.

Figure P17

Qualitative Analysis NVivo® Hierarchy Chart for Combined ASRS, NTSB and AIDS Parts 121 and 135 dataset – Level Two Diagram



Note. The bottom right blue and orange color represent the AIDS dataset with minimal data size.

Figure P18

GPower® Graph Showing MANOVA Sample Size Calculation Example Word Tree on NTSB Parts 121 and 135 Classified Dataset

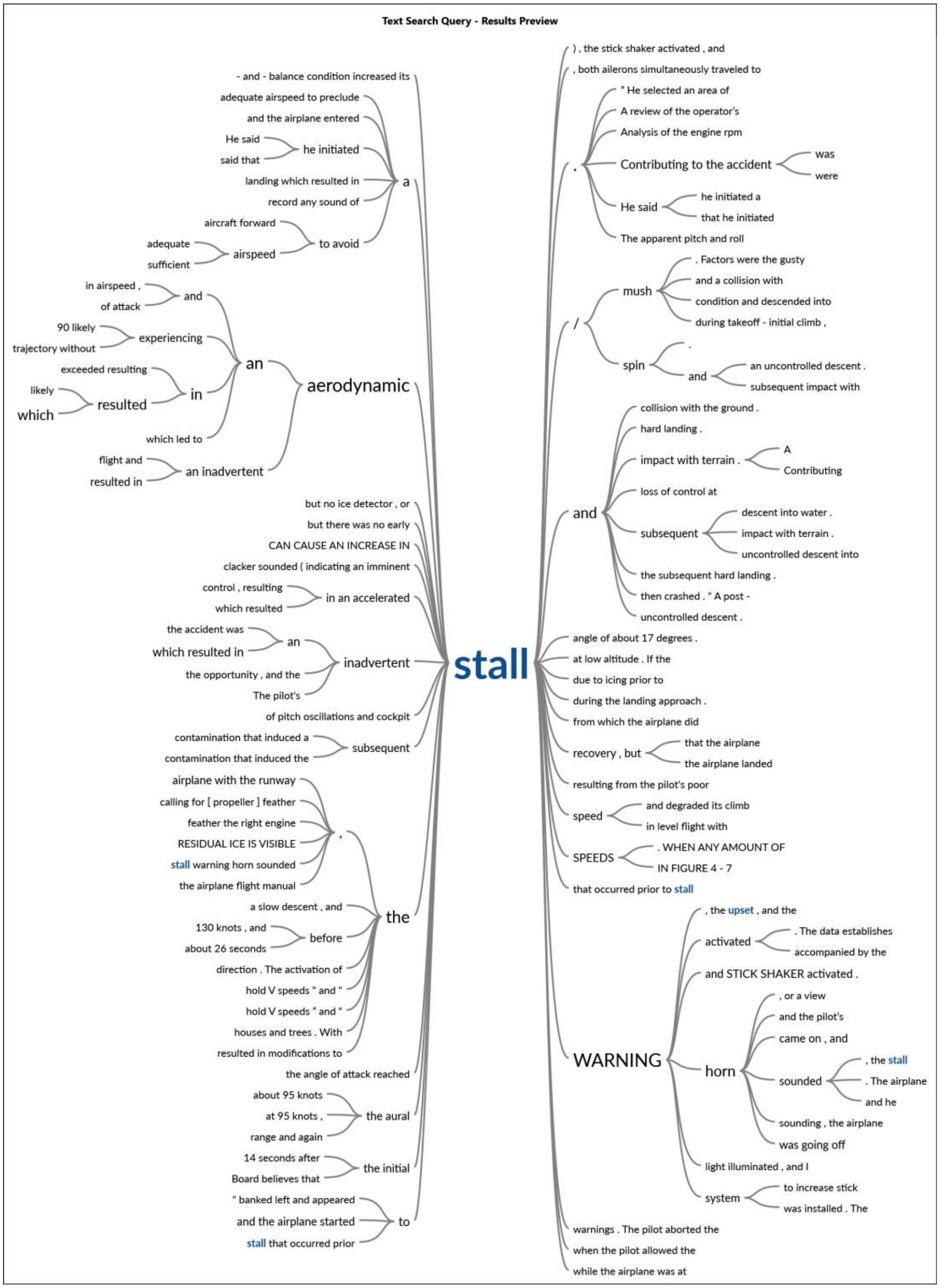


Figure P19

Example Cluster Analysis on ASRS Parts 121 and 135 Classified (Left) and Augmented (Right)