Airborne Incidents

An Econometric Analysis of Severity

December 19, 2014

Final Report

Produced for:

Office of Safety

Air Traffic Organization

Federal Aviation Administration

Washington DC

Produced by:

Lee Biernbaum, Megan Price and Jacob Wishart

U.S. Department of Transportation

Office of the Secretary of Transportation

Volpe National Transportation Systems Center

Cambridge, MA

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT-VNTSC-FAA-15-03		

4. Title and Subtitle	5. Report Date / 6. Performing Organization Code
Airborne Incidents: An Econometric Analysis of Severity	5. Report Date **December 19, 2014** 6. Performing Organization Code **OST-R/Volpe/V321**

7. Author(s)	8. Performing Organization Report No.
Lee Biernbaum, Megan Price, Jacob Wishart	**DOT-VNTSC-FAA-15-03**

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS) / 11. Contract or Grant No.
U.S. Department of Transportation **Volpe National Transportation Systems Center** **55 Broadway** **Cambridge, MA 02142**	10. Work Unit No. (TRAIS) 11. Contract or Grant No. **FA01B1**

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered / 14. Sponsoring Agency Code
U.S. Department of Transportation **Federal Aviation Administration** **Air Traffic Organization, Office of Safety and Technical Training** **800 Independence Avenue, SW** **Washington, DC 20591**	13. Type of Report and Period Covered **Final Report, December 2014** 14. Sponsoring Agency Code **FAA/AJI-33**

15. Supplementary Notes

16. Abstract

Airborne loss of separation incidents occur when an aircraft breaches the defined separation limit (vertical and/or horizontal) with another aircraft or terrain imposed by Air Traffic Control. Identifying conditions that lead to more severe loss of separation incidents can lead to policy implications and future areas of research. Previous research focused on qualitative approaches to analyzing such events, and tended to examine only the frequency of events. This report puts the severity of a loss of separation incident front and center and uses econometric techniques to examine the relationship between severity and conditional factors during the incident.

The report utilizes report data from the Air Traffic Safety Action Program (ATSAP), with a concentration on terminal airspace incidents. A number of other FAA data sources were merged to provide a robust set of information at the time of event in terms of facility, weather, and other operational characteristics. The primary focus of this research was on the use of discrete choice, multinomial logit models to better understand the relationship between these different set of factors at the time of the event and the severity outcome.

17. Key Words	18. Distribution Statement
Airborne loss of separation, air traffic control, incident severity, safety analysis, discrete choice, multinomial logit, ordinal logit, terminal airspace safety	**No restrictions**

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	**Unclassified**	**163**	

EXECUTIVE SUMMARY

Study Purpose

The Volpe team examined loss of separation incidents in terminal airspace to determine what factors are associated with increased incident severity. The study utilized robust quantitative statistical models that provide an unbiased look at what types of factors are most often associated with severe, and in particular, catastrophic events. The resulting findings have potential implications for future policies aimed at reducing the odds of an airborne incident becoming severe.

This is the only time known to Volpe that the Air Traffic Safety Action Program (ATSAP) database, controlled by the National Air Traffic Controllers Union (NACTA), has been opened up to researchers for an econometric analysis. It provides a wealth of information on the specific factors that were present when an airborne incident occurred, as well as the severity of the resulting incident. This was combined with data on facility information (NFDC and DTRB), daily operations (OPSENT) and weather in order to provide a more complete view of the circumstances surrounding each incident.

Key Results

These results are a small subset of the results found in the study with high statistical robustness. They highlight the types of findings the study generated.

Aircraft

- **Single Engine Props:** Single engine props are **1.7 times more likely to be associated with severe incidents** than are single-aisle jets. They are **3.6 times more likely to be catastrophic** in Tower facilities.
- **Experimental Aircraft:** In Tower facilities, incidents with experimental aircraft are **6.2 times more likely to be severe, and 21 times more likely to be catastrophic**. In TRACON facilities, they are **22 times more likely to be catastrophic**.
- **Visual Approaches:** Incidents with visual approaches are **2.6 times more likely to be associated with severe incidents** in Tower facilities than incidents with instrument approaches.

Airspace and Pilot Actions

- **Aircraft/Pilot action complexity factor**: In Tower facilities, these are **1.6 times more likely to be associated with severe incidents**, and **2.2 times more likely to be catastrophic.** This variable indicates if the complexity of the aircraft performance or pilot action was a significant factor during the loss of separation incident.
- **Airspace Type D:** Type D airspace is **2.3 times more likely to be associated with severe incidents** for all facility types, and is **3.4 times more likely to be catastrophic.**
- **Pilot Evasive Actions**: There is a **300% percentage point decrease in the probability of a severe incident** for all facility types if the pilot takes action to avoid a potentially dangerous situation.

Control Status

- **Training is in Progress:** Incidents with training in progress are **1.4 times more likely to be severe** than incidents without training in progress in Tower facilities.

Communication

- The following communication variables are “causal factors”, meaning that they are entered into ATSAP if it is believed that they were a contributing factor to an incident.
- **Flight Plan/PDC Processing Problem:** In Tower facilities, flight plan/PDC processing problems are overwhelmingly low in severity.
- **Radar Misidentification:** Incidents with radar misidentification are **3 times more likely to be severe** than incidents without Radar Misidentification in TRACON facilities.
- **Acknowledgement Problems:** When an acknowledgement problem is cited as a causal factor, incidents are **1.7 times more likely to be severe** in both Tower and TRACON facilities.
- **Loss of Communication:** Incidents with a loss of communication are **1.4 times more likely to be severe** in Tower facilities; in TRACON facilities they are **2 times more likely to be severe, and are 5 times more likely to be catastrophic**.

Methodology

The Volpe team used Multinomial Logit (MNL) techniques in order to tease out associations between severity levels and factors present in the ATSAP database. MNLs were pioneered by Dr. Daniel McFadden (Nobel Prize winner in 2000) in the 1980s and are now ubiquitous in discrete choice modeling. Since other ATSAP variables are held constant, the individual impact of each variable becomes apparent in this type of modeling environment.

The ATSAP database breaks down incidents into Minimal, Minor, Major, Hazardous and Catastrophic severity categories. For this analysis, Major, Hazardous and Catastrophic incidents were considered “severe”. The Volpe team placed a particular focus on catastrophic events, because although rare, these are the incidents most likely to be involved in a crash, and thus have the most direct relationship to safety outcomes. The model output is the form of relative risk ratios, which indicate how much more likely a variable’s inclusion makes the outcome more likely to be severe as opposed to not-severe, or at a given severity level as compared to a minimal severity level.

ATSAP consists entirely of situations in which incidents occurred, thus all comparisons are in reference to a “typical incident”. Since Volpe does not have data on normal operations (flights that do not result in incidents), the results cannot be interpreted as identifying factors that leads to a flight being more (or less) likely to be involved in an incident in the first place.

TABLE OF ACRONYMS

Acronym	Definition
ATC	Air Traffic Control
ATO	Air Traffic Organization
ATSAP	Air Traffic Safety Action Program
ARTCC	Air Route Traffic Control Centers
CNAC	The Center for Naval Analyses Corporation
DTRB	The Digital Terminal Resource Book
FCT	Federally Contracted Towers
GA	General Aviation
ICAO	International Civil Aviation Organization
IFR	Instrument Flight Rules
IIA	Independence of Irrelevant Alternatives
METAR	From the French Mètéorologique Aviation Régulière. Hourly weather reports automatically generated
MLE	Maximum Likelihood Estimation
MNL	Multinomial Logit
NACTA	National Air Traffic Controllers Association
NAS	National Airspace System
NFDC	The National Flight Data Center
OE	Operator Error
OI	Operation Incident
OED	Operational Error/Deviation
OLS	Ordinary Least Squares
OPSNET	Operations Network Database
PPO	Partial Proportional Odds
RNAV	Area Navigation
RO	Routine Operation
RRR	Relative Risk Ratios
SATORI	Systematic Air Traffic Operations Research Initiative
SAR	System Analysis Recordings
SID/STAR	Standard Instrument Departure/Standard Terminal Arrival Route
SMS	Safety Management System
STARS	Standard Terminal Automation Replacement System
TCAS-RA	Traffic Collision Avoidance System – Resolution Advisory
TFM	Traffic Management Initiatives
TRACON	Terminal Radar Approach Control
VFR	Visual Flight Rules

TABLE OF CONTENTS

TABLE OF TABLES

Table Page

TABLE OF FIGURES

1. INTRODUCTION

The motivation of this report is to examine the underlying factors that could contribute to the severity of an airborne loss of separation incident in terminal airspace. An airborne loss of separation incident is defined as a situation where two (or more) aircraft breach the defined separation limit (vertical and/or horizontal) imposed by Air Traffic Control (ATC). Breaching this defined separation limit can lead to aircraft getting dangerously close to another, with the most severe outcome resulting in either mid-air or terrain collisions. In order to reduce the likelihood of severe loss of separation incidents, understanding what the main factors and components that drive these incidents is vital.

The report's findings rely heavily on detailed data examination and statistical techniques. Previous quantitative research is relative thin in the area of loss of separation severity modeling, and tends to focus on the frequency of loss of separation incidents instead. Additionally, research presented in this report focuses on loss of separation incidents at an aggregate level, not on a case by case basis as is frequently the case in other research. Doing so will allow for a better understanding of the broader trends and patterns that will help inform both policy-making and guide future research.

1.1. Background

In accordance with standard established by the International Civil Aviation Organization (ICAO), the FAA developed the Safety Management System (SMS) in 2005. SMS ranks loss of separation incidents according to their severity into 5 categories, with 'Minimal' being least severe and 'Catastrophic' being the most severe, resulting in a collision with another aircraft, object, or terrain. In 2008, the Air Traffic Safety Action Program (ATSAP) was implemented as a voluntary self-reporting system for air traffic controllers to report safety and operational concerns including loss of separation incidents. The research detailed in this report uses ATSAP reports dating from 2008 to January 2013 as the source of incidents analyzed. Each incident reported in ATSAP contains a severity level based on the SMS criteria and Table 1 details the distribution of severity levels by calendar year.

Table 1 - Loss of Separation Severity Levels across Calendar Years (Terminal Airspace Only)

Year	Minimal	Minor	Major	Hazardous	Catastrophic	Total
2008	249	46	15	2	0	312
2009	2,662	923	267	39	24	3,915
2010	3,013	1,618	554	108	51	5,344
2011	2,734	1,707	613	104	55	5,213
2012	4,104	1,379	621	87	55	6,246
2013	230	49	34	2	4	319
Total	12,992	5,722	2,104	342	189	21,349

Calendar year 2008 was the roll out period for ATSAP and 2013 only contains one month of data, which explains the relatively few number of reported loss of separation incidents for these CYs. The general trend in reporting, however, has been upwards, with 2012 seeing the most reported incidents. This should not be confused with an overall increase in loss of separation incidents, but rather it is a function of the increase use of the ATSAP reporting system.

1.2. Research Methodology

The goal of this research is to use statistical and econometric methods to identify trends in airborne loss of separation incidents. The dataset is comprised of reported loss of separation incidents in terminal airspace. Special care was given to what facility type (tower or TRACON) the incident occurred at, as influences of events may vary significantly by facility type. In addition to the main ATSAP dataset, data on flight operations, weather, and facility characteristics were gathered and merged for each incident reported in ATSAP.

The analysis focuses on statistical models, in which multiple variables were included together to allow for their interactions with severity levels to be better understood. The modeling effort focused on discrete choice models, namely through the use of multinomial logit estimations. Estimated multinomial logit results are presented in tandem with graphs and figures showing changes in likelihood of severity outcomes for key variables.

1.3. Overview of the Document

This document is split into six main sections. The first section contains a brief literature review of previous airborne loss of separation severity research and discrete choice modeling of severity in other modes of transportation, followed by a methodology review. The third section is a detailed discussion of the four main input datasets. The fourth section provides details on statistical procedures undertaken in this report, a discussion on how to interpret results from statistical models and graphics, and the research team's choices regarding variable selection and grouping. The fifth section contains the data exploration and modeling broken down into six subsections for each respective category analyzed in this report. Each subsection is further broken into a data exploration portion and a statistical modeling portion. The data exploration portion employs a variety of one- and two-way descriptive statistics and cross tabulations of variables, coupled with figures and graphs detailing the nature of data with respect to severity levels. Due to the categorical nature of much of the data, cross-tabulations and other analysis of variance tests against severity are an effective, albeit preliminary means of exploring these variables. The final section contains a "best-prediction" logit model comprised of variables from each subsection of the data exploration.

Due to the length of the appendices, a second standalone document contains all appendices. Appendix A contains the full, comprehensive list of causal factor variables analyzed. Appendix B contains the literature review from the previous Runway Incursions report, while Appendices C-H contains tables, figures, and alternate model results for each data categories.

2. LITERATURE AND METHODOLOGY REVIEW

The current field of research on airborne loss of separation occurrence and severity can be broken-down into two broad sets. The first set includes qualitative studies that attempt to identify specific factors (elements of airspace sectors, human/causal, etc.) that contribute to the occurrence and severity of operation incidents (OI).[1] The second set uses predictive models to estimate the correlations or relationships between potentially relevant factors and OI occurrence and severity. The first set of studies could provide guidance on potential explanatory variables or concepts that should be explored further. The second set provides background into the methodology that has been previously employed, and what lessons can be taken away when developing a model of airborne loss of separation severity.

2.1. Previous Airborne Operational Error Research

2.1.1. Causal Factors

Previous research on OI severity has tended to focus on human or causal factors as a means to explain the reasons for and suggest ways to mitigate OIs. A frequently cited study by Schroeder and Nye examined various measures of controller workload at the time operational errors/deviations (OED) occurred.[2] Using data from the 1985-88 air route traffic control centers (ARTCC) operational errors/deviations, the authors found relatively strong correlation between the 5 different causal factor categories (i.e. high correlation between causal factor category A and B, B and C, C and A, etc.). However, the correlation between any one of the causal factor categories and the reported traffic complexity or the number of operations within the sector was much smaller. The authors make note that this weak relationship between causal factors and varying workload was not all surprising and was consistent with other studies done at the time suggesting that air traffic controllers are capable of handling variable work scenarios without committing causal factor errors.[3]

[1] Note that the operational incident term is relatively new. All the previous airborne research presented here will refer to such events as operational errors (OE), as that was the term in use when the authors published their reports. An Operational Error/Deviation (OED) is outdated terminology for an OE.

[2] Schroeder and Nye (1993)

[3] Vortac et al. (1992)

In a related paper, Rodgers and Nye used data from the FAA's Operational Error Data Base, sampled from 1988-1991 and grouped operational errors into three levels of severity (minor, moderate, and severe). They then examined how the number of aircraft being controlled, traffic complexity, aircraft stage of flight, and causal factors involved in the error varied across the different levels of severity.[4] Through the use of Chi-squared tests of significance, the authors determined that there were no statistical relationships between the number of aircraft, traffic complexity, or phase of flight with major or moderate OE severity. Although this result seems counterintuitive, a simple Chi-squared test of significance may not be rigorous enough in determining the actual relationship between any two independent variables. More sophisticated statistical techniques for untangling the relationship between OE severity and these independent variables may have yielded different results. The author, however, found that the misuse of conflict alert systems, readback communication errors, and inter-facility coordination errors were all causal factors associated with a higher percentage of moderately severe errors.

Pounds and Ferrante detail potential contributing causal factors in OE events and highlight strategies for reducing them.[5] The authors describe prevention programs for readback/hearback and position relief briefings errors. Facility programs designed to prevent hearback/readback errors included "tape talks" where facility staff specialists review voice recordings of controllers on duty to assess communication performance, along with a mandatory briefing video produced by the Air Traffic Investigation Division Staff highlighting how different noise distractions can contribute to readback/hearback errors.

The authors also examined trends in TRACON facility OE event data sampled from 1997-2000, which revealed that OEs often occurred within 10 minutes of a controller taking over another controller's position.[6] The authors show that approximately 9% of OE occurred within the first 5 minutes, and 18% within the first 10 minutes of a controller switch. In order to address to this perceived problem, all managers were required to validate position relief checklists, in addition to ensuring the briefing was recorded.

[4] Rodgers and Nye (1993)

[5] Pounds and Ferrante (2003)

[6] Data provided by the Air Traffic Evaluations and Investigations Staff, AAT-20.

2.1.2. Age and Experience

The relationship between age and/or air traffic controller experience and the frequency of OEs has been considered in a number of studies. Of the more statistically comprehensive studies, the Center for Naval Analyses Corporation (CNAC) combined agency personnel records and controller experience at the time of the error with en route OE data for the period January 1991 to July 1995 from the FAA Operational Error/Deviation System.[7] Controllers were grouped by experience, and the total number of controllers with errors was divided by the total number of controllers with the same experience to estimate the "likelihood" of an OE for each experience group. The CNAC study found that experience and the likelihood of an OE were significantly related, with the likelihood of an OE declining rapidly in the first few years of experience.

Broach re-analyzed the CNAC dataset to look at both controller age and experience.[8] OE dates were not available in the CNAC data, so calculating the age of the controller at the time of the event was not possible. Thus, controller age was based on the beginning of the 5-year period (January 1991) and grouped by one-year increments from ages 18 to 48. The likelihood of OEs was then regressed on both age and experience, with the estimated coefficient on age being positive and the estimated coefficient on experience being negative. This would suggest that age could increase the likelihood of an OE, while experience has the counter-effect of reducing the likelihood of an OE.

Following the previous study, Broach and Schroader used 7 years of OE data from the OEDS ranging from October 1, 1996 to September 30, 2003 to estimate the effects of age and experience on the frequency of OEs.[9] Age was then split into two groups, with controllers falling into either the older controller (age greater than 56) or younger controller (age 55 and less) category. Experience was measured as tenure at a facility, and similar to age, six tenure groups were created starting with less than three years, followed by an interval 6 years wide, then 5 years wide up to greater than 25 years. Broach and Schroader then estimated a Poisson regression to analyze the number of OEs as a function of controllers' age (split into the two groups) and experience (tenure) groups.

They found that there was no statistically significant difference between the groups, casting doubt on the overall effect of age on the likelihood of an OE event. However, the authors did not investigate the effects of experience on OE events across different tenure groups, leaving this relationship as an open question.

---

[7] Center for Naval Analyses Corporation (1995)

[8] Broach (1999)

[9] Broach and Schroader (2005)

It should be noted that these studies only dealt with the *frequency* of OE events, and not the *severity* associated with these events. It could still be the case that older controllers or less experienced controllers are involved in more severe OE events. Also not discussed was how different age group cutoffs could change the outcome. By only providing output for one set of age groups, the authors fail to be robust for the alternative age group scenarios. Moreover, these studies only focused on en route centers and not facilities functioning in terminal airspace. While these airspaces are not completely unrelated, there is no evidence to the contrary that the age and experience relationships with OE severity should be same.

2.2. Previous Airborne Incident Severity Predictive Modeling Research

Two papers published by the FAA attempt to predict OE events through logistic regression analysis. Pfleiderer and Manning examined prediction and classification of OEs and routine operations (ROs) with a two stepwise logistic regression analysis.[10] The authors used sector characteristics variables to estimate two separate logistic models, one for high-altitude sectors and one for low-altitude sectors. The OE events were constructed from Systematic Air Traffic Operations Research Initiative (SATORI) re-creations, while RO events were derived from System Analysis Recordings (SAR).[11] The central goal of the paper was to determine how well the logistic regression model could accurately distinguish between the OE and RO events. The high-altitude model accurately predicted 80% of the cases between in-sample OE and RO events, while the low-altitude model accurately predicted 79% of the cases between in-sample OE and RO events.

This study, however suffers from a number of drawbacks. First, the author’s correctly note that the coefficient estimates on key sector characteristic variables in a stepwise logistic regression framework do not provide causation to OE and RO events. Second, the dataset consisted of OE events dating from 9/17/2001 to 12/10/2013 and RO events from 2/25/2005 to 3/3/2005. This time differential between the sampled OE and RO events creates multiple issues with the dataset having uncontrolled, systematic differences between the two groups. Moreover, there was a clear violation of the assumption that all OE and RO events were independent of one another because of how OE and RO events were paired (by sector, day of week, and time of day). This is because logistic regression analysis assumes that all observations are independent of one another. A random draw from the sample of OE and RO events would have ensured this assumption.

[10] Pfleiderer and Manning (2007)

[11] The FAA Civil Aerospace Medical Institute (CAMI) developed SATORI to allow for the recreation of air traffic control operational incidents in a format much like the one displayed to air traffic controllers. SATORI recreations are generally used a training tool after an OI has occurred.

In light of these modeling concerns in their previous study, Pfleiderer et al. conducted a very similar study using logistic regression analyses to determine whether a set of sector characteristics could distinguish between OE and RO events.[12] OE data was pulled from a sample of Systematic Air Traffic Operations Research Initiative (SATORI) re-creations of OE events at the Indianapolis ARTCC between 9/17/2001 and 12/10/2003, and RO data was derived from System Analysis Recordings (SARs) taped between 5/8/2003 and 5/10/2003. Again, a backward stepwise elimination was used to reduce the sectors variables down to the “best” statistical fit of the model. These “best” fit models were then used to predict between RO and OE events, with the low- and high-altitude models accurately classifying 75% and 79% of events, respectively.

The authors claim that their models classification rates achieved through the use of select sector characteristics support the hypothesis that elements of the sector environment contributes to the occurrence of OE events. The authors do not take the extra step to isolate which variables directly contribute to OE, or rather, which variables they believe have a causal relationship with OE events. While desirable, this would have been inappropriate given their use of backward stepwise regression. This method eliminates independent variables that may not necessarily improve the model fit but are still important in terms of their relationship with the dependent variable. Moreover, the order in which variables are removed can impact which remaining variables are significant. In other words, the backward stepwise elimination method has the potential to create omitted variable bias, and thus the interpretation of the remaining variables will be biased either positively or negatively, depending on the relationship.

2.3. Severity Research on Other Modes

While the general focus of previous airborne incident severity research has focused narrowly on data based on en-route airspace, the question of factors contributing to automobile crash severity has been examined extensively. This highway literature can provide important insight into how to approach modeling airborne incident severity. In addition, reviewing crash severity literature can illuminate those areas were airborne incidents are similar to and diverge from the highway crash literature and will require careful consideration. Even when factors diverge, reviewing similar methodology and modeling techniques could prove fruitful.

[12] Pfleiderer et al. (2009).

Schneider IV and Savolainen examined statewide motorcycle data from the state of Ohio to identify factors associated with the level of injury sustained by motorcyclists involved in crashes.[13] Multinomial logit models were developed for different types of motorcyclist crashes (single-vehicle and multivehicle motorcycle crashes at both intersection and nonintersecting locations), with the results suggesting that crash factors varied by crash type and location and that severe injuries were more likely at high speeds or when drugs and/or alcohol were present. While drugs and alcohol have no clear translation in airborne incidents as such violations are rare, crash location and speeds are more easily associated with airborne factors such as phase of flight and flight speeds. This study, however, also suffers from a number of reporting and methodology drawbacks. The authors *only* present statistically significant coefficients and elasticities from the multinomial logit models. This is due to the fact that the authors restrict their models to only include those variables that are statistically significant. Not only does this have the potential to bias the estimations due to omitted variable bias, but also non-statistically significant findings are important in their own right. Reporting statistically insignificant results is a critical step in the research process and this paper will avoid omitting these results.

In a similar study, Schneider IV et al. examined the factors contributing to driver injury severity along horizontal curves in Texas.[14] A multinomial logit approach was used and separate models were developed for three different curve radii (small, medium and large). The authors found that not wearing a seatbelt greatly increased the chance of a fatality. The same is true for the presence of alcohol and drugs. Those factors have no clear analogues in the airborne incident framework. The authors also examined environmental factors and found that clear weather and daylight increase the chance of a less severe accident. Weather may also play a significant role in airborne incident severity. Another factor the authors considered was vehicle type. Certain vehicle types (motorcycles) were associated with higher probabilities of more severe injuries while others (semi- and pickup trucks) were not. This has the potential to translate directly in terms of examining the impacts of aircraft type or certification (commercial versus general aviation aircraft (GA)) on the airborne incident severity.

Kockelman and Kweon also examined the factors contributing to driver injury severity.[15] The authors used an ordered probit methodology and focused on different types of crashes: single versus two vehicle crashes. O'Donnell and Connor provide an example of using an ordered logit to examine automobile accident injury severity.[16] Both of these papers are discussed at length in the Runway Incursion Report.[17]

---

[13] Schneider IV and Savolainen (2011)

[14] Schneider IV*, et al.* (2009).

[15] Kockelman and Kweon (2002).

[16] O'Donnell and Connor (1996).

[17] Biernbaum and Hagemann (2012).

As an extension of the ordered logit model, partial proportional odds (PPOs) models have been used in recent automobile crash studies. Using the example of crash severity, the main advantage of the PPOs models have over traditional ordered models is that restriction that coefficient values remain the same across severity levels is relaxed, and are allowed to vary. Yasmin et al. examine driver actions that contributed to the severity of a crash involving one or more emergency vehicles.[18] Using a PPO model to overcome the proportional odds assumption and crash data from the Province of Alberta, Canada, from the periods 1999 to 2008, the authors found that drivers' violations (deliberate deviations of driving laws) contributed significantly to increasing the severity of the crash. Kaplan and Prato also use a PPO model to study the link between crash severity and crash avoidance maneuvers.[19] Using crash data from 2005 to 2009 from the General Estimates System crash database, the author's main findings were that most drivers failed to act when facing critical events and drivers rarely performed crash avoidance maneuvers that reduced the severity of the crash.

Both of these papers highlight the ordered-response nature of crash severity and the applicability of ordered logit models, which has direct comparisons to the ordered nature of airborne incident severity. These papers also correctly identify the restrictive nature of the proportional odds assumption in traditional ordered models, and use PPOs as a means to gain more complete and accurate models of severity. Further discussion of ordered and PPOs models are presented in the Methodology Review section of this report.

While this research is suggestive of methodologies and factors to consider for airborne incidents, there are glaring differences between motor vehicle crash injury severity and airborne incident severity. Both methods are conditional on the incident occurring in the first place, but the nature of the incident varies considerably between the two. That is, all automobile models consider the severity of a *crash*, whereas the airborne models consider the severity of a non-crash, which is itself a loose proxy for the *likelihood* of a crash at all. It is important to keep these differences in mind when using injury severity literature to inform a study on airborne incidents. While the underlying methodology and statistically modeling will not vary greatly, the interpretation of the coefficients will be slightly different.

2.4. Conclusion

These papers present a summary of the types of methodologies that may be used to understand airborne incident severity. Yet, the papers have some flaws worth noting with the intention that the same flaws are avoided during the modeling process for the current research. Certain papers suffered from reporting deficiencies, such as not reporting all coefficients. Other papers suffered from methodological limitations in their selected modeling techniques.

[18] Yasmin et al. (2012)

[19] Kaplan and Prato (2012)

Methodology drawbacks aside, it should be restated that this paper’s underlining topic is inherently different than other papers discussed here. Past studies on airborne incidents tended to focus on the number of severe events that occurred in en-route airspace, not the severity of those incidents. In other words, past airborne studies were *frequency* based, while the research presented here will be *severity* based. While it would be ideal to be able to considered both the frequency side (what is the likelihood of an event occurring given a random sample of events) and the severity level side (given that there was loss of separation, what is the likelihood it was more or less severe incident) of airborne incidents, data limitations force this study to focus solely on factors contributing to the severity of a loss of separation incident. With that said, the severity of airborne incidents is a relatively unexplored area in terms of statistical modeling that can benefit greatly from the thorough research presented in this report.

3. MODELING METHODS

3.1. Methodology Background

While analysts use a variety of modeling methods, the purpose of this research is to engage in statistical analysis using regression models. Within regression models, though, a wide range of specifications are possible; selecting an appropriate model (or series of appropriate models) requires an understanding of the different assumptions underlying each model. These underlying assumptions can also impact the interpretation of model results, which can in turn affect policy recommendations. This section will review basic regressions as well as discrete choice models.

3.1.1. Regression as a Concept

The most basic regression framework is ordinary least squares (OLS) regression. Given a dependent variable **Y** and a set of independent variables **X**, the basic structure can be described as:

$$\mathbf{Y} = \beta\mathbf{X} + \varepsilon$$

where β is a set of coefficients that can be estimated that captures the effects of variables, and ε is a random disturbance term that includes "unobserved variables" that are not captured in X. In this framework, β represents the marginal impact of an increase in X on Y. If β is positive, then increased X is associated with increased Y; if β is negative, then increased X is associated with decreased Y. It is also important to note that this framework merely describes the relationship between X and Y and says nothing of causation in either direction.

In the context of regression analysis, OLS regression is applicable to a wide range of situations. For example, it can be used to explore the relationship between income and demographic factors or the health impacts of various policy decisions. It allows the researcher to decompose the effects of exogenous variables, controlling for their differing impacts on the dependent variable. OLS regression is extremely flexible in terms of the relationships between variables that can be captured. The X described above can include just a few variables, or many with interactions between them. OLS regression is also simple to implement.

Despite its many advantages, OLS regression has some serious shortfalls when trying to describe data such as airborne incident severity. By definition, the severity measure in ATSAP of an airborne incident falls into one of five categories: Catastrophic to Minimal. The convention in this case is to number the categories 1 through 5, with Catastrophic being the highest number (thus positive β suggest increasing severity). However, it becomes quickly apparent that OLS does not bound the estimation in any way. That is, given the right confluence of negative βs, OLS may predict a score less than one (or perhaps even a negative score).

Figure 1 below presents this distinction graphically. The figure depicts a hypothetical sample of heights and weights and plots the relationship between them. Notice that various intermediate values of height are shown and that the values of height are not restricted in any fashion. These data are appropriate for analyzing with OLS regression.

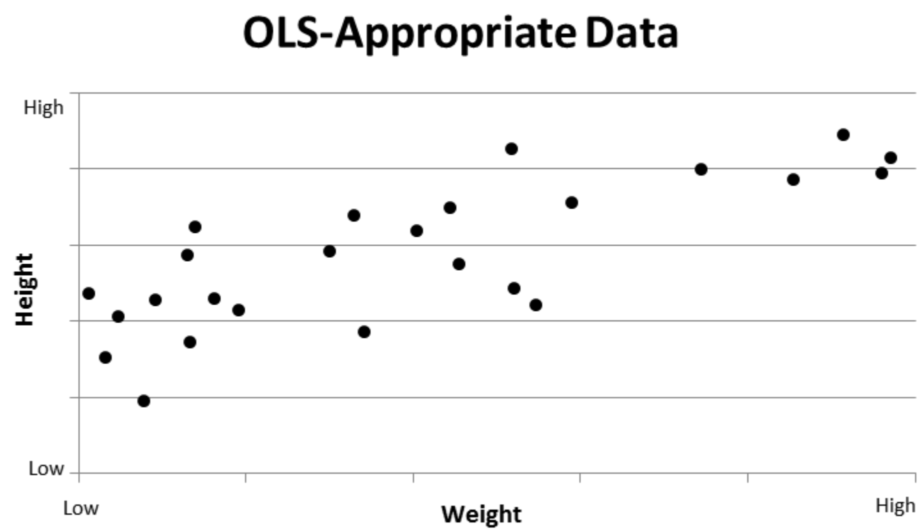

Figure 1 - Example OLS Data

The following figure, Figure 2, depicts data that is not appropriate for analyzing with OLS and is categorical in nature. Notice that the heart attack risk group outcome is restricted to only three values: low medium high and intermediate values are not possible.

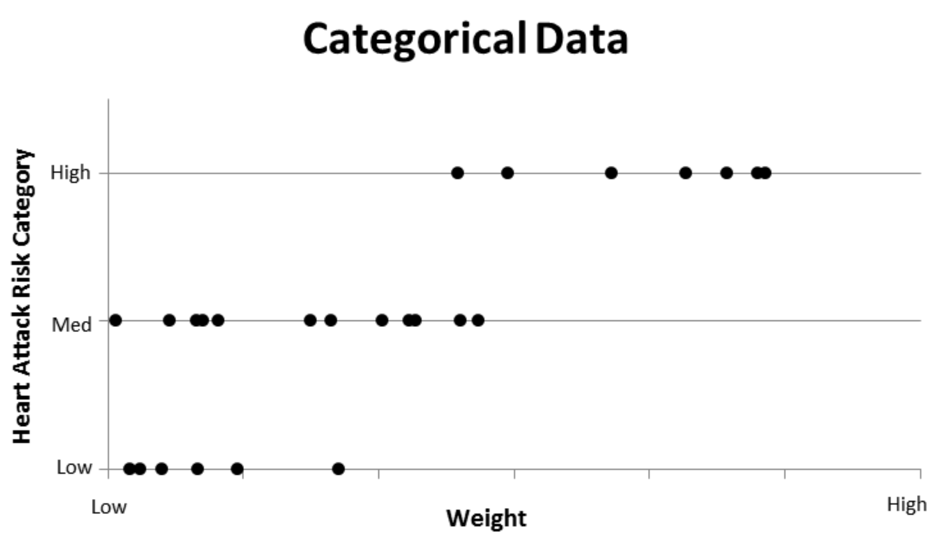

Figure 2 - Example Categorical Data

In addition to the problems relating to boundedness and integer values mentioned above, OLS has an additional, and perhaps more important, failing in relation to incident severity data. Airborne incident severity data has the property that it is merely ordinal, not cardinal. That is, incident severity data has some sort of ranking but the ranking does not describe the distance between ranks. An incident with a severity level of Catastrophic is more severe than a Hazardous incident, which is in turn more severe than a Major incident. However, a Catastrophic incident may be *much* more severe than a Hazardous incident compared to the difference between a Hazardous incident and a Major incident. OLS regression will not acknowledge this aspect of the data; OLS will treat the change between any two categories as equal, which makes it a suboptimal choice for analyzing airborne incident data.

3.1.2. Alternatives to Linear Regression

Data like airborne incident severity falls into a category that can be described as "discrete choice" data. The data points are placed into distinct categories, often of a qualitative nature. An entire class of models has been developed to analyze discrete choice data and overcome the limitations of OLS regression discussed above.

Discrete choice models have been developed to look at binary choice and to analyze sets with more than two choices. These multi-choice models come in a variety of flavors such as ordered (which recognizes an inherent ordering in the categories) and multinomial (which do not recognize any ranking among choices). There are additional extensions to the multinomial model framework that seek to relax several of the constraints imposed by the standard multinomial model; for more information, see Appendix D.

Beyond the world of OLS and its extensions, a major tool for (frequentist) econometrics is maximum likelihood estimation (MLE). MLE can be used to estimate a plethora of different model types and all of the models discussed later in this report are estimated using MLE techniques. The focus of MLE is the likelihood function, $\mathcal{L}$:[20]

$$f(y_1, \dots, y_n | \beta) \equiv \mathcal{L}(\beta | y)$$

for a sample of n observations, each with a value of y, noted as $y_1 \dots y_n$. This equation represents the likelihood of observing the data, y, given parameters β. For this particular application, the likelihood function, f or $\mathcal{L}$, represents the distribution of airborne incident severities. This formulation can be extended to include other conditioning variables X:[21]

$$f(y_1, \dots, y_n | \beta, X) \equiv \mathcal{L}(\beta | y, X)$$

[20] Greene (2003).

[21] Ibid.

On the above equation, Greene notes:

> the likelihood function is written in this fashion to highlight our interest in the parameters and the information about them that is contained in the observed data. However, it is understood that the likelihood function is not meant to represent a probability density…, the parameters are assumed to be fixed constants which we hope to learn about from the data[22]

This likelihood function can be thought of as the data generation process. Suppose y is the probability of rain today. Then X will be variables that may influence that, such as temperature, humidity, and atmospheric pressure. β characterizes the impact of those variables on y. The likelihood can also be thought of as the probability of observing that set of y, given X and β. Maximum likelihood estimation, true to its name, seeks to choose a β to maximize the above expression (the probability of observing that set of y given X and β.)

β is of fundamental interest to the econometrician and policy-maker. β captures the effects of the various exogenous variables X on the dependent variable y. It is from this information that informed policy decisions can be made.

3.1.3. Discrete Choice Models

As noted earlier, airborne incident severity rankings fall into a category known as discrete choice data. A variety of models have been developed to analyze these types of data. Each of the potential models has underlying assumptions and characteristics that may influence the applicability of that model to the analysis of airborne incident severity. This report will focus mainly on the different extension of the logit model, which is the discrete choice technique utilized in this research.[23]

Ordered and multinomial models address choice sets with multiple alternatives (3 or more alternatives). However, the main difference is that ordered models recognize an inherent ordering of the choices while multinomial models assume there is no underlying order to the choices. Of course, situations such as airborne incident severity are clearly presented as ordered, but multinomial models can also be used to examine ordered data, providing some potential benefits as well as drawbacks. The following section discusses multinomial, ordered, and partial proportional odds logit models.

[22] Ibid., p. 468-469.

[23] For an at length discussion of the differences between logit and probit models, please refer to Biernbaum and Hagemann (2012), and Greene (2003) for a more technical approach.

3.1.3.1. Multinomial Logit Models

In the multinomial logit (MNL) framework, the random disturbances for different choices are assumed to be uncorrelated.[24] In other words, the unobserved variables that influence the probability of choice A are entirely unrelated to the unobserved variables that influence the probability of choice B. This property may not hold in reality, resulting in biased estimates from the model.

> A direct result of the assumption regarding the correlation of the random disturbances is what is called the independence of irrelevant alternatives (IIA) property. Specifically, the ratio of any two choice probabilities is independent of the probabilities of any other possible choices.[25]

There exists a battery of post-estimation statistical tests for testing the IAA assumptions. Long and Freese present evidence of this from various simulation studies that show these tests are not useful, in practice, for assessing violations to the IAA assumption.[26] They further argue that the multinomial model best work under conditions where the alternatives are unrelated and not just substitutes. This is clearly not always the case, but In terms of this research, airborne severity categories are by definition different from each other. There remain questions on how severity categories could be better defined and structured, but for the point of this study the IAA assumptions will by definition hold.

Another property of the MNL relates to parameter estimation. Specifically, "estimable parameters relating to variables that do not vary across outcome alternatives can, at most, be estimated in *I*-1 of the functions determining the discrete outcome (*I* is the total number of discrete outcomes)."[27] One potential way to address this is to normalize the coefficients for one outcome (the "base" outcome). Thus, parameters for variables that do not vary across categories can be estimated for the remaining categories. The coefficients are then interpreted as a change relative to the base outcome.

[24] Greene (2003)., p. 724

[25] Ibid., p. 724

[26] Long and Freese (2006)., p. 243-246

[27] Ibid., p. 318.

Conditional on the IIA assumption holding, the main advantage MNL models have over the ordered family is the ability to relax the ordering of choices. This allows for each severity level to vary independently from each other, and allows for a more in-depth examination of each severity classification. As noted above, MNL models see extensive use in practice. Islam and Mannering provide a good example of a multinomial logit being used to examine injury severity.[28] Dow and Endersby provide an example of a multinomial logit looking at voter behavior in comparison to a multinomial probit model.[29] Finally, Schneider IV et al. also examine injury severity using a multinomial logit framework.[30] Additional discussion of the theoretical aspects of the multinomial logit specifications can be found in Washington et al. and Greene.[31] There are extensions to the MNL model that seek to relax some of these restrictions, such as IIA. Two of the most common extensions are nested logit and random parameter models. A brief discussion of these extensions can be found in Appendix D.

3.1.3.2. Ordered Logit and Partial Proportional Odds Models

Ordered models place a strong constraint on the estimated coefficients: the parallel lines assumption. The estimated coefficients in an ordered model are assumed to be invariant by the choice outcome. This assumption is often violated, causing the interpretation of ordered model results to be inconsistent. Washington et al. provides a good example: consider accident severity data that has severity rankings of property damage only (i.e., no injuries), injury, and fatality. Additionally, suppose the effect of airbag deployment was of interest. An ordered model constrains the coefficient to either "increase the probability of a fatality (and decrease the probability of no injury) or decrease the probability of fatality (and increase the probability of no injury)."[32] This may not be the case in reality. Airbag deployment may reduce the probability of a fatality and of no injury, while increasing the probability of an injury.

[28] Islam and Mannering (2006).

[29] Dow and Endersby (2004).

[30] Schneider IV*, et al.* (2009).

[31] Washington*, et al.* (2011)., Greene (2003).

[32] Washington*, et al.* (2011), p. 358.

Traditionally, a multinomial specification is the solution to allow the flexibility for such effects.[33] Recent methodology advances in generalized ordered models and partial proportional odds models now allow for such flexibility for varying coefficient estimations in ordered models. Generalized ordered models will vary all estimated coefficient for each outcome, while partial proportional odds (PPOs) models allow the econometrician to control which variables to vary per outcome.[34] Using the Washington et al. example again, a partial proportional odds model would have allowed the direction and probability of the coefficients on airbags to reflect the fact that they airbag deployment may reduce the probability of a fatality and of property damage, while increasing the probability of injury.

While traditional ordered models do not allow for this sort of complexity, they do provide more intuitive coefficient interpretation over multinomial models. If the coefficient is positive, increasing the value of the explanatory variable unambiguously increases the probability of being in the highest category and the probability of being in the lowest category decreases, though intermediate categories have a more subtle relationship.[35] Thus, a tradeoff must be made between accounting for additional accuracy in modeling complex relationships between severity levels and providing results that are useful and practical to policy-makers. Moreover, this distinction only exists in the event that the effect of an explanatory variable is not the same across severity levels.

Given that airborne severity levels are potentially ordinal in nature, Washington et al. note that "if an unordered model (such as the multinomial logit model) is used to model ordered data, the model parameter estimates remain consistent but there is a loss of efficiency."[36] In other words, the multinomial estimates are less precise than an ordered model (higher standard errors), but are consistent in terms of not violating crucial model assumptions that would otherwise bias the results. There is an essential "trade off … between recognizing the ordering of the responses and losing the flexibility in specification offered by unordered outcome models."[37]

3.2. Methods chosen

Given the discussion above, it is clear that each model has some pros and cons associated with it. Recall that the decision criteria for a model to be desirable included tractability, precision, and how well it reflects reality. The specific nature of the airborne incident data does not suggest any particular model choice. Though the data does have some sense of ordering to the categories, multinomial models provide some advantages in terms of analysis, especially as the ordering present in the data may be the result of multiple processes.

---

[33] Ibid.

[34] Williams (2006).

[35] Greene (2003), p. 738.

[36] Washington*, et al.* (2011), p. 345.

[37] Ibid., p. 359.

Due to the nature of the data (i.e., severity ratings from Catastrophic to Minimal), it was initially desired to focus on the analysis on the ordered family of models. However, ordered models failed to pass the parallel lines assumption necessary to warrant their use. Partial proportional odds models were also estimated, but time did not allow for a deep analysis of the results.[38] Future research should focus on using PPO models as a way to pass the necessary assumption associated with ordered models. Therefore, the primary model of choice for this report is the multinomial logit model.

[38] Estimation for both ordered and PPO models can be found in Appendices C-H.

4. INPUT DATA

4.1. Air Traffic Safety Action Program Data (ATSAP)

4.1.1. Source

The ATSAP database is maintained by the FAA Air Traffic Organization (ATO) in conjunction with the National Air Traffic Controllers Association (NACTA) and is a voluntary self-reporting system for air traffic controllers for safety and operational concerns. It contains 22,381[39] terminal-area airborne events from May 18, 2007 to January 25, 2013. The self-reported events are processed by a committee of experts and are assigned a severity rating based on event specific information.

4.1.2. Contents

The ATSAP database contains basic information on each airborne event (date, time, facility location), aircraft, parties involved (controllers and pilots), and possible causal factors. This dataset serves as our “base” dataset, where all other datasets will be integrated and merged to.

4.1.3. Concerns

There are inconsistencies in the structure of the dataset from year to year required additional data cleaning. Certain variables (namely causal factors) only appear in some years and do not span the entire dataset. This is due to revisions to the ATSAP form, which now includes more causal factor choices than in previous versions. Because the dataset spans multiple years, the dataset contains answers from multiple versions of the reporting form. It was important to treat these newer variables as “Missing” and not “Yes/No” for the period before these variables were added to the dataset.

Many of the variables in the ATSAP database are free text entry. The use of free text fields creates several problems for analysis. First, data entry can be inconsistent. For example, aircraft model may be entered several different ways despite all referring to the same aircraft (B737, 737, B-737 may all be entries referencing a Boeing 737, for example). Second, free text allows respondents to include multiple values in the field. For example, some causal factors can take on multiple values of which the respondent may select none, some, or all potential values. This provides more information for analysis, but required cleaning to turn the information into a useable format for modeling purposes.

The ATSAP dataset also does not contain any Federally Contracted Towers (FCT), which are by nature all small airports with much fewer operations. This is still important to keep in mind, because the representative sample of facilities is inherently excluding loss of separation incidents that occurred at FCTs. It is unknown whether FCT incidents are more severe or more common, so the scope of this exclusion may be non-trivial. Therefore, all results from this report are applicable only to FAA controlled tower and TRACON facilities.

[39] There were originally 22,704 events reported to the Volpe Center, with 323 non-terminal events that were later dropped by the Volpe Center from the dataset.

4.2. Weather Information

4.2.1. Source

Hourly METAR weather readings at airports are archived by Plymouth State University in New Hampshire.[40] These METAR readings represent a standardized set of information automatically collected by weather stations. Plymouth State University was able to provide weather readings for nearly all of the location-hour pairs in the ATSAP dataset.

4.2.2. Contents

The hourly readings contain information about temperature, humidity, wind conditions, visibility conditions, and information about active weather such as storms. In addition, some readings contain summary amounts of precipitation for the past 6 or 24 hours.

4.2.3. Concerns

Readings of average precipitation over the previous 6 or 24 hours are not reported in every METAR record. Consequently, additional hours of data were required to ensure a precipitation reading is available for every event. However, because the precipitation readings will not be concurrent with the event, there is some discrepancy between the reported precipitation and the conditions during the event. For example, if the next precipitation reading is five hours after the event (so that the previous six hours include the time of the event) but the rain did not start until two hours after the event there would be a precipitation reading, but it is likely unrelated to the outcome of the event. This is largely an interpretation problem rather than a data problem, however.

A similar discrepancy with the weather data is related to the location of the weather readings. Previous experience with this dataset indicated that readings were reliably available for towered airports, as towered airports have a weather station on the property. However, the ATSAP dataset contains reports from both TRACONs and towers. Events that take place at TRACONs can occur anywhere in the airspace of that TRACON and deriving the closest METAR station requires additional assumptions, detailed in the following section.

[40] Website: http://vortex.plymouth.edu/

4.2.4. Integration with the base data set

Extra care was taken when matching the weather data with the appropriate facility. For TRACON events, the event location variable in ATSAP allowed for matching to coincide with the closest towered airport. For the incidents where this was not possible due to ambiguous event location data (around 20% of the total incidents), the primary towered airport was used for that particular airport as the event location reference for that incident. This raises a number of concerns that should not be discounted. Firstly, even though there is a higher probability that the incident occurred in the vicinity of the primary towered airport, there could still be many incidents where this is not the case. Secondly, weather phenomena can vary in certain areas of TRACON space. For example, the TRACON that serves the Oakland and San Francisco International Airports also serves San Jose and Sacramento International Airports. Weather can vary drastically between these two sets of airports, and if the San Francisco airport is incorrectly used as location of the event, then the weather data could very well be completely wrong. Given these potential issues, the assumption will be made that any incorrectly specified location with large varying weather phenomena will be averaged out due to the large number of incidents in the underlying dataset. In other words, the high number of observation in the dataset will help ensure that any noise introduced by data measurement error with be smoothed out. However, as a precautionary measure, tower and TRACON facility incidents will be modelled separately when examining weather data.

The weather data is reported hourly and represent point estimates of the condition at that time. The ATSAP database contains the time of the event down to the minute. Because weather data did not necessarily align with the timing of the incursion event, a way to interpolate the weather at the time of the event was developed. Two methods were developed: one for variables that change continuously (e.g., temperature) and one for variables that change discretely (such as precipitation).

The method for continuous variables relied on linear interpolation. The two weather readings on either side of the incident were used as the basis for the interpolation. The method for variables that changed discretely relied on picking the observation closest to the time of the incident. The weather readings occur roughly hourly (and more frequently in changing weather conditions) so the closest reading is, in general, less than 30 minutes away. This method was used for the variables including the weather code (indicating precipitation, fog, smoke, haze, etc.). The remainder of the variables (temperature, cloud cover, etc.) were all subject to the linear interpolation method. The combination of these two methods provided a set of data that could be matched exactly to the ATSAP database, making the matching trivial after the interpolation steps.

4.3. Facility Characteristics

4.3.1. Source

Facility characteristics are derived from two main sources. The National Flight Data Center (NFDC)[41] provided information on the number of runways at various facilities. The Digital Terminal Resource Book (DTRB)[42] provided information on facility level (a measure of complexity used to adjust controller pay) and the mapping between airports and the TRACONs that serve them.

4.3.2. Contents

Both the NFDC and DTRB contain many variables beyond those used for this analysis. For the purposes of this analysis, only the number of runways, facility level, and the TRACON serving a given airport were collected. For airports, number of runways is a meaningful and easy calculation. For TRACONs, the total number of runways at towered airports served by that TRACON will be used.[43]

4.3.3. Concerns

A concern with this dataset is the exclusion of non-towered airports from the DTRB data. Without non-towered airports, the calculation of total number of runways served by a TRACON is only approximate. If a mapping between non-towered airports and TRACONs could be developed, this calculation could be improved. Further, the facility level calculation collapses many factors down into one number for payroll reasons, rather than analysis reasons. Examining the details of the calculation[44] indicates that many of the factors included correspond to facility complexity. Thus, the facility level appears to be a useful measure for this analysis, but caution should be taken in interpreting the results as the measure is a combination of a variety of factors and labor/negotiating factors affect changes to a facility's complexity level designation.

4.3.4. Integration with the Base Data Set

As the ATSAP reports are filed by controllers, ATSAP reports are only generated by controlled facilities (TRACONs and towered airports). All controlled facilities exist in both the NFDC and DTRB. Therefore, merging facility characteristics onto the base dataset is simple.[45]

4.4. Operations Data

4.4.1. Source

Daily operations data are available from the FAA through the Operations Network (OPSNET) website.[46]

---

[41] Website: https://nfdc.faa.gov

[42] Website: http://terminaltools.faa.gov/DTRB/

[43] See the Concerns section for more information.

[44] http://nso.natca.org/NSO%20Docs/NSO_PDF/2009%20Contract%20Arbitration%20Decisions/Appendix_A-1.pdf

[45] Only 86 events fail to merge with facility characters. This is due to misspelled or otherwise incorrect (gibberish) facility codes in the ATSAP dataset, which cannot be fixed.

4.4.2. Contents

Daily operations are available for both Tower and TRACON facilities, spanning the entire sample period (May 18, 2007 to January 25, 2013). Operation counts per facility are given for both itinerant and overflight IFR and VFR flights for commercial air carriers, air taxis, general aviation (GA), and military traffic.

4.4.3. Concerns

There are no significant concerns with this dataset. Each observation in the ATSAP dataset that has a correct facility identifier is matched with the appropriate operations data for that specific date and time. The main issue to be aware of is matching operations for *combined* tower and TRACON facilities, where the facility ID information is not sufficient in matching the appropriate operations data. That is, tower operations should be matched to the tower portion of the facility while TRACON operations should be matched to the TRACON portion of the facility. It was possible to get a correct operations match per event for nearly all events by discerning what type of position the controller was working when the event happened.

4.4.4. Integration with the Base Data Set

Integration with the ATSAP dataset required accurate facility identifiers and date information. Additional information was required for combined tower and TRACON facilities, where the event had to be identified as either happening at the tower or TRACON in order to appropriately assign operations data. Indicators were generated using ATSAP controller position data to specify which part of the combined facility the event took place. However, there were some controller positions that are ambiguous in terms of which part of the facility they apply to and were not viable for matching purposes.[47]

[46] Website: https://aspm.faa.gov/opsnet/sys/main.asp

[47] Less than 4% of observations were ambiguous.

5. DATA TRANSFORMATION, STATISTICAL PROCEDURES, AND HOW TO READ THE OUTPUT AND GRAPHICS IN THIS REPORT

Data in the ATSAP dataset went through extensive cleaning and data transformation before being modelled. There were also common statistical practices undertaken when modeling. The following were, in general terms, the type of data transformation and statistical procedures Volpe used.

5.1. Categorical Variables

Free text categorical variables in the ATSAP dataset were converted to binary indicators. For example, data on airspace type was originally a single variable in ATSAP that contained the 5 different airspace types (A, B, C, etc.). In order to analysis airspace type, binary indicators for each airspace type were generated.

5.2. Binary Severity Measure

A binary severity measure was generated, splitting the five severity categories into severe and non-severe incidents. Minimal and Minor events (which encompass 87.7% of incidents) were grouped as non-severe, while the remaining three categories, Major, Hazardous, and Catastrophic were group as severe incidents.[48] There appears to be an inflection point between Major and Hazardous that gets ignored when modeling certain variables over this binary severity measure. This causes certain relationships to appear insignificant and raises questions not only about what the appropriate binary severity measure should be, but also the definitions of the individual severity measures.

5.3. Aircraft Type

Extensive cleaning was performed on the aircraft type data. The ATSAP database originally contained descriptions of the aircraft make and model, such as B-737 or Boeing 737. Since there was little consistency in how the data was entered, entries needed to be examined individually. Planes were grouped into 11 categories: single aisle jets, multiple aisle jets, regional jets, corporate jets, multiple-engine props, single-engine props, civilian helicopters, military jets, military props, military helicopters, and experimental aircraft for analysis.

5.4. Data Aggregation

Certain subgroups of variables were aggregated together due to either commonality or a lack of observations. For example, the variable for phase of flight had many categories for arriving (arrival) aircraft. All arrival based variables were combined together to make the model interpretation more straightforward. Certain causal factor variables lacked enough observations to allow for sufficient statistical variation. In order to model these variables, aggregate causal factor groups were generated. Guidance on these categorical and causal factor aggregations was provided by the FAA and can be found in Appendix A.

---

[48] Guidance on the severity groupings was provided by the FAA.

5.5. ATSAP Form Discrepancies

Over the course of the sample period, different versions of the ATSAP reporting form were introduced. Because of this, certain variables are only available for a subset of the sample. This issue was taken into account and any period before certain variables appeared in more recent versions of the ATSAP form were treated as missing. Further, some models contain only events within the time period containing the relevant variables.

5.6. Standard Error Corrections

All logit and multinomial logit models presented here are estimated with clustered standard errors around facility level. This allows for idiosyncratic correlations within individual facility's standard errors to exist, but not across different facilities. In other words, observations across facilities are independent, but not necessarily within groups (clusters) of facilities.

5.7. Relative Risk Ratios and Multinomial Logit Model Output Interpretation

All multinomial logit output is presented in terms of relative risk ratios (RRRs). This was imposed as a means to preserve the coefficient interpretation across both binary and multinomial logits. RRRs are essentially the exponential of the coefficient and can be interpreted as either increasing or decreasing the likelihood of an outcome if the RRR is greater than or less than one, respectively. All models use Minimal severity as the base outcome, so any interpretation of a significant variable must be considered as an increase or decrease in the likelihood of a given severity outcome to the Minimal outcome.

Table 2 presents the coefficient estimates of a multinomial logit model using relative risk ratios. The interpretations of the coefficients are similar to that of Odd Ratios that would appear in a single variable logit: for an incident that occurred in Airspace Type B, the likelihood of it being a Minor incident *increases* relative to the base (Minimal) outcome. Moreover, for an incident that occurred in Airspace Type B, the likelihood of it being a Catastrophic incident *decreases* relative to the base outcome.

Table 2 - Example of Multinomial Logit Model Output, Relative Risk Ratios

	Minor	Major	Hazardous	Catastrophic
Airspace Type B	1.290* (0.161)	1.055 (0.148)	0.911 (0.224)	0.359* (0.146)
Airspace Type C	1.154 (0.105)	1.038 (0.145)	0.611 (0.171)	0.722 (0.261)
Airspace Type D	1.039 (0.106)	0.926 (0.124)	1.206 (0.242)	1.728* (0.381)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 6916				

5.8. Interpretation of Probability Charts and Graphs

Presenting the results from the multinomial logit models are best served in two ways depending on whether or not the variable is categorical in nature. For both set of variables, continuous variables that are not changing are held to their mean value, while categorical values that are not changing are set to zero.

Categorical variables of interest are expressed as the marginal changes in the probability of each severity category when the categorical variable becomes a factor. This is represented graphically, where the x-axis is each severity outcome and the y-axis is the marginal change in probability. The associated percentage change in probability is also shown for selected variables of interest. An example is present in Figure 3. When aircraft or pilot actions are a factor in an incident, the probability of a minimal incident decreases by 0.07, while the probability of a catastrophic incident increases by around 0.01. This has an associated decrease in the percentage change in probability of around 10% for minimal and an increase of close to 100% for Catastrophic.

Presenting both the change in probability and the associated percentage change allows for a more thorough examination of independent variables with respect to changes in severity levels. The example here shows that even though a seemingly large drop in the probability of a minimal incident and a seemingly small increase in the probability of a Catastrophic incident, are in fact characterized by the opposite percentage changes. This is due the relatively rare nature of the more severe incidents, where the probability of an incident being Catastrophic is typically quite low, so even a small increase in the probability will be associated by a large increase in the percentage change.

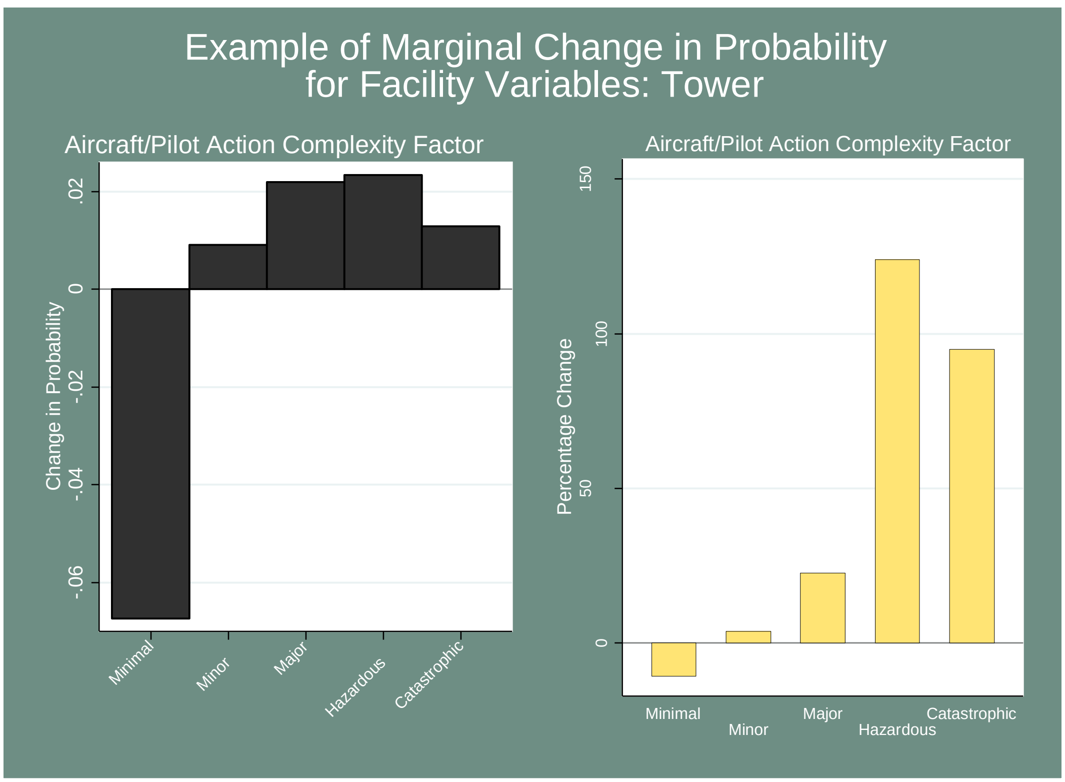

Figure 3 - Example of Marginal Change in Probability Figures

Continuous variables are presented as their impact on the probability for each severity category. The x-axis is the continuous variable of interest and the y-axis is probability area for each severity category at that particular value of the continuous variable. Examples of these probability area charts are provided in Figure 4 and Figure 5.

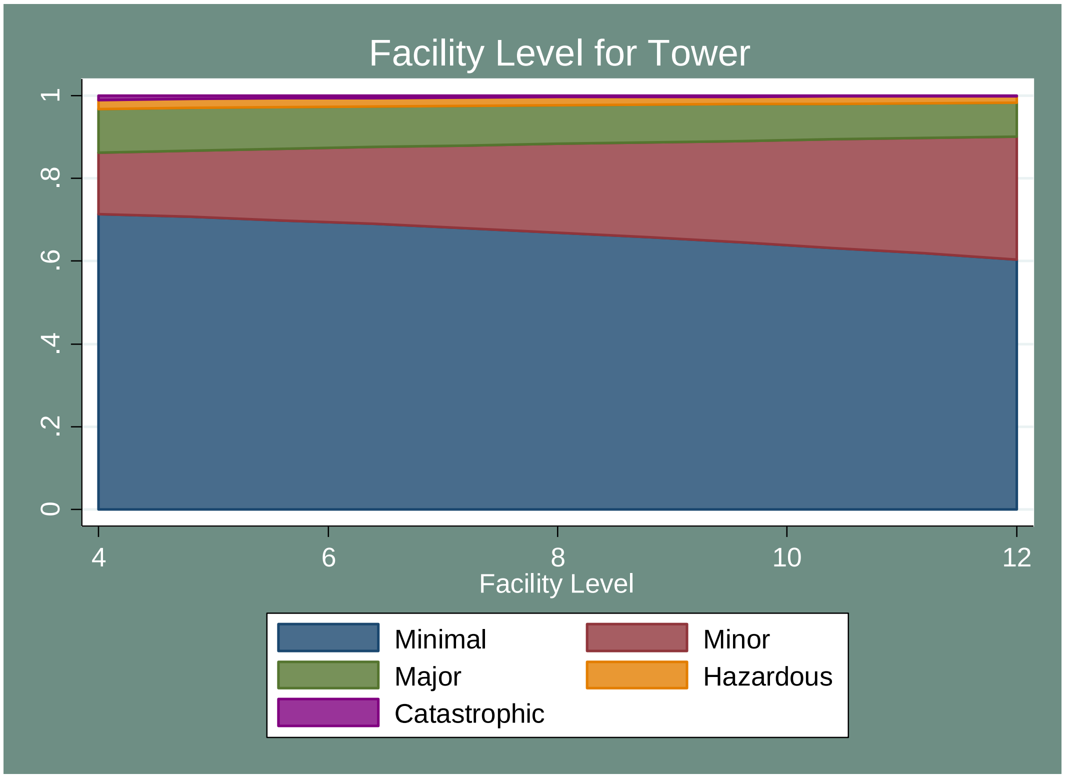

Figure 4 - Example of Probability Area Chart

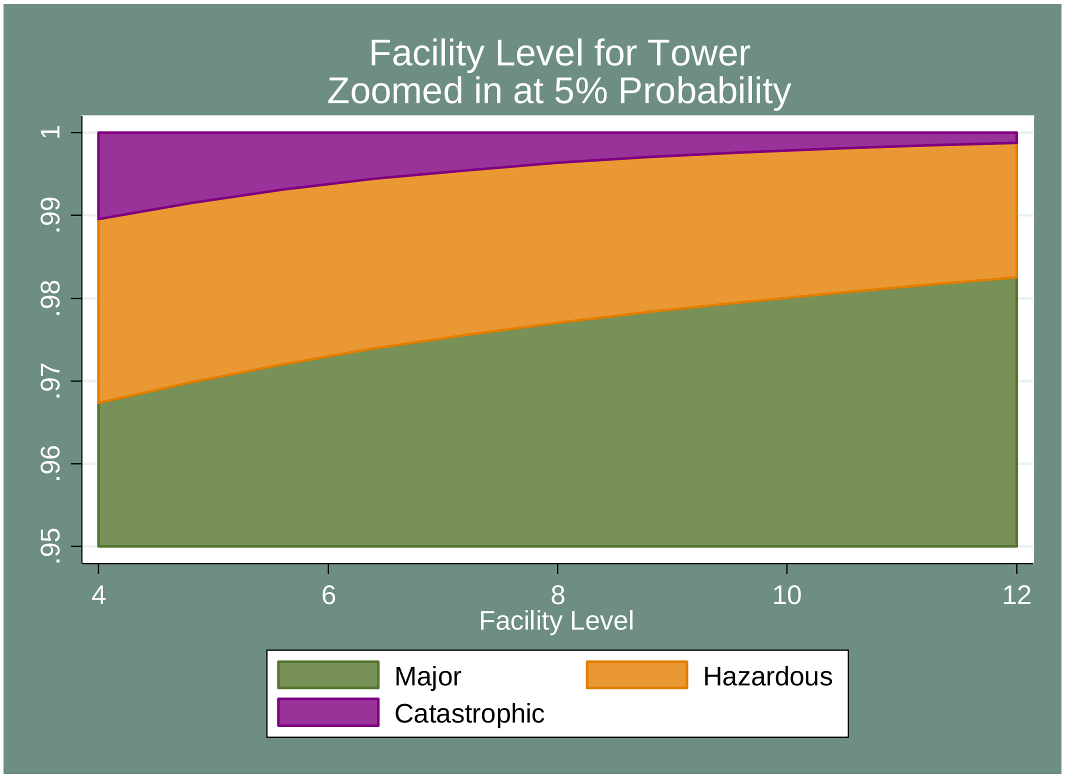

Figure 5 - Example of Probability Area Chart, Zoomed in at the 5% Probability

The probability area charts are simple to interpret; as the independent variable increases (moving from left to right on the x-axis), the associated impact on the probability of each severity category is captured on the y-axis. In this example, as the facility level of a Tower increases, the probability of a Minimal outcome decreases, while Minor increases. In order to examine the relative impact of facility level on the more severe incident levels, a zoomed in figure is also presented for select variables.

5.9. Causal Factors

These variables are some of the most interesting to model, but need to be treated with care since they were entered into the database only when a controller (or analyst) determined that they were a cause of an incident, itself a concept that varies in definition individual-to-individual. The resulting statistical issues are described below. Despite these caveats, causal factors are included in further analysis.

5.9.1. Reporting Bias

There is a potential for systematic bias in the data. For example, suppose that hypothetically, ambient noise is unrelated to incident severity, but there were two incidents where ambient noise was present. One was a Minor incident, in which a plane barely diverged from the correct flight path. The other was a Catastrophic incident in which two planes collided mid-air. Both incidents are reported in the ATSAP database, however, the controller may believe that the first, Minor, incident, was a fluke, and thus expect less effort selecting from the hundreds of available causal factors such as ambient noise. In the second, Catastrophic, incident, the controller may choose to expend more effort reporting the incident, and may list a greater number of causal factors, such as ambient noise. This would lead to a spurious correlation appearing between ambient noise and severity. We do not have a way to separate out a relationship between severity and a propensity to spend time reporting more causal factors in ATSAP from the true relationship between causal factors and severity.

5.9.2. Measurement Error

These variables were entered into the database only when a variable was thought to have been a potential cause of an incident, rather than simply if the variable was present. For example, if there is a noisy environment, it will be only entered into the database if the noise was determined to be a potential cause of an airborne incident. Furthermore, different controllers may have different thresholds of what level of sound constitutes a noisy environment. Thus it is not possible to say that controllers involved in incidents where noise was not marked as a causal factor were not working in a noisy place. This is problematic for modeling purposes, and could create issues in terms of producing robust results in multivariate equations. To reiterate the aforementioned caution: any statistically significant or impactful results concerning causal factors should be taken with care.

5.10. Variable Selection for Modeling

The Volpe Center followed an approach that emphasized examining the complete available dataset in a thorough and well document manner to ensure that any statistically robust relationships with incident severity were documented.

5.10.1. Variable Choices

Given the relatively large nature of the dataset (in terms of number of independent variables) the need to organize and categorize variables was an important step before conducting any analysis. As with previous research, the Volpe team divided the variables into large categories,[49] running a separate model on each, followed by a "best-prediction" model utilizing the results of each of the models for each category. The design and results of each model section are explored in subsequent sections of this report.

The analysis strove to provide a comprehensive look at ATSAP data, including a broad array of variables. When multiple variables measured highly similar concepts, not all were included in the models in order to avoid problems with multicollinearity.

The choice of variables for inclusion in models was driven by:

- Intuitive, logical framework
- Data quality of individual variables
- Degree of correlation to incident severity level
- Interaction with other variables

The specific variables used in each model are described below, in sections 5.10.2 through 5.10.7.

Causal Factors, a Special Case

Because ATSAP contains several hundred causal factors, contained in their own three level hierarchies, additional procedures were necessary to maximize the utility of these variables.

Primarily, many of the causal factors occur quite infrequently due to the high level of specificity about the factors. These small sample sizes make it unlikely that a relationship would be detected, even if present, and may overly differentiate between factors that are highly similar to each other. Consequently, as detailed in Appendix A, causal factors are considered independently only when sample sizes and simple descriptive statistics indicate it may be useful to do so. Otherwise, they are aggregated into Sub-Sub-Categories largely (but not entirely) in line with the hierarchy inherent to the ATSAP reporting form. In the tables below, causal sub-sub-categories factors are listed in the variable column in *italics*, and the table in Appendix A contains the breakdown of specific factors within that sub-sub-category.

---

[49] Aircraft, facilities, controllers, communication, airspace/pilots, and weather

5.10.2. Aircraft information

These variables contain descriptors of the type of aircraft involved in an incident. Analyzing these variables allowed the relationship between incident severity and the phase of flight, aircraft size, or the number of aircraft involved to be determined.

Table 3 - Aircraft Information Variables

Aircraft Information Sub-Category	Variables
Aircraft Type	Aircraft Type
Flight Plan	Flight Plan
Number of Aircraft Involved	Number of Aircraft Involved
Phase of Flight	Control Status
Phase of Flight	Phase of Flight
Special Events	*Emergency Situation*
Special Events	*Military Action*
Special Events	*Special Event*
Special Events	*Traffic Management Initiative*

5.10.3. Airport (Facility) Characteristics

These variables describe characteristics of the airports where incidents occurred. Analyzing these variables allowed the relationship between incident severity and airport attributes to be determined. The airport characteristics variables in the ATSAP database primarily provide information on the complexity of each airport's operations.

Table 4 - Facility Characteristics Variables

Facility Characteristics Sub-Category	Variables
Facility Complexity	ATC Level
Facility Complexity	Daily Operations
Facility Complexity	Number of Runways
Facility Complexity	Traffic Complexity Rating
Facility Complexity	Traffic Volume Rating
Organizational and Complexity Factors	*Abnormal Configuration*
Organizational and Complexity Factors	Aircraft Performance or Pilot Action Complexity Factor
Organizational and Complexity Factors	Airspace Procedure Complexity Factor
Organizational and Complexity Factors	Communication Complexity Factor
Organizational and Complexity Factors	Coordination Complexity Factor
Organizational and Complexity Factors	*Facility Influences*
Organizational and Complexity Factors	*Organizational Influences*
Organizational and Complexity Factors	*Policy/Procedure Influences*
Organizational and Complexity Factors	*Staffing Configuration*
Organizational and Complexity Factors	*Supervisory Influences*
Organizational and Complexity Factors	Traffic Complexity Factor
Organizational and Complexity Factors	Traffic Volume Complexity Factor

5.10.4. Controller Variables

These variables describe attributes and environment of the air traffic controllers assigned to a flight where an incident occurred. Particular interest was in the relationship between controller experience and incident severity, and the relationship between a controller's position and severity.

Table 5 - Controller Variables

Controller Sub-Category	Variables
Controller Actions	Controller Actions
Controller Experience	Years of Experience
Controller Experience	Years at Facility
Controller Experience	*Specific Training Issue*
Controller Experience	*Currency/Proficiency Level*
Time-of-Incident Descriptors	Control Status
Time-of-Incident Descriptors	Control Position
Time-of-Incident Descriptors	*Information Exchange*
Unsafe Acts	*Decision Error*
Unsafe Acts	*Perceptual Error*
Unsafe Acts	*Skill-Based Error*
Unsafe Acts	*Violation*
Controller Influences	*Controller Influences*
Controller or Equipment Capacity	*Controller or Equipment Capacity*
Work Area Influences	*Work Area Influences*
Equipment Influences	*Equipment/Software Design Issue*
Equipment Influences	*Equipment Malfunction/Area/Coverage*

5.10.5. ATC/Pilot Communication/Clearance

These variables describe communication issues between ATC and pilots. Analyzing these variables allowed the relationship between incident severity and ATC/pilot communications to be determined.

Table 6 - Communication Variables

Communication Sub-Category	Variables
Communication Problems	*Loss of Communication*
Communication Problems	*Phraseology*
Communication Problems	*Readback Problem*
Communication Problems	*Aircraft Acknowledgement Problem*
Communication Problems	*Clearance Problem*
Flight Data, Display Problems, Aircraft Observation	*Computer Entry Problem*
Flight Data, Display Problems, Aircraft Observation	*Flight Plan/PDC Processing Problem*
Flight Data, Display Problems, Aircraft Observation	*Radar Misidentification Problem*
Flight Data, Display Problems, Aircraft Observation	*Displayed Data Problem*

5.10.6. Airspace and Pilot Actions

These variables describe the airspace the aircraft is operating and pilot actions at the time that an incident occurred. Analyzing these variables will enable the relationship between incident severity and certain types of airspace characteristics and pilot behavior to be determined.

Table 7 - Airspace and Pilot Actions Variables

Airspace and Pilot Action Sub-Category	Variables
Airspace Characteristics	*New Airspace*
Airspace Characteristics	*Old Airspace*
Airspace Characteristics	*Special Use Airspace*
Airspace Classification	Airspace Type (altitude)
Aircraft Equipment	*Aircraft Equipment Issues*
Aircraft Performance or Pilot Response	Compression on Final
Aircraft Performance or Pilot Response	Untimely Aircraft Descent/Climb
Aircraft Performance or Pilot Response	Untimely Aircraft Turn
Aircraft Performance or Pilot Response	Untimely Roll
Aircraft Performance or Pilot Response	Untimely Runway Exit
Aircraft Performance or Pilot Response	Untimely Speed Adjustment
Expectation Bias	Expectation bias (pilot)
Non-Conformance with a Clearance	Altitude
Non-Conformance with a Clearance	Altitude Crossing
Non-Conformance with a Clearance	Course
Non-Conformance with a Clearance	Speed
Pilot Reaction	Evasive Action
Pilot Reaction	Go Around
Pilot Reaction	TCAS-RA
Pilot Reaction	Other
Pilot Reaction	Unknown
Procedure Type	RNAV Procedure Type
Procedure Type	Conventional Procedure Type
Procedure Type	Directive/Publication/Regulation problem type

5.10.7. Weather Variables

These variables describe weather characteristics at the time that an incident occurred. Analyzing these variables will allow the relationship between incident severity and certain types of weather characteristics to be determined.

Table 8 - Weather Variables

Weather Characteristics Sub-Category	Variables
Dew	Dew Point Temperature
Dew	Temperature
Dew	Temperature-Dew Point Difference
Other Weather	Sea Level Pressure Deviation
Other Weather	Weather Complexity Factor
Other Weather	Weather Phenomena
Other Weather	Wind Speed
Precipitation	Precipitation Last 6 Hours
Visibility	Cloud Ceiling
Visibility	Cloud Coverage
Visibility	Visibility

6. DATA EXPLORATION AND MODELING

6.1. Aircraft Information

Variables in the aircraft category contain descriptors of the type of aircraft involved in an incident. Volpe determined that a number of these variables are correlated with incident severity. Most strikingly, incidents involving experimental aircraft are associated with increased likelihood of catastrophic severity. In tower facilities, planes flying under visual flight rules also see increased severity.

Aircraft variables are grouped into sub-categories, which are discussed in the following sections. The sub-categories include aircraft type, control status, flight plan, the number of aircraft involved, phase of flight, and special events. Data in each sub-category is analyzed separately; this information is then brought together in the full aircraft model.

6.1.1. Aircraft Type

Experimental planes are more likely to be involved in catastrophic incidents. Single engine props are and corporate jets are also associated with increased severity in tower facilities.

Data on aircraft type was derived from the ATSAP fields for aircraft make and model. Volpe aggregated these into 11 categories[50] for analysis. An additional category of “ground” was used for any ground vehicles at an airport. Volpe expected that results from this model would be tied to pilot skill level and the type of route generally flown by a specific aircraft type, rather than any inherent safety differences between the aircrafts. Table 9 shows the result of a binary logit analysis of severity and aircraft type.

---

[50] Civilian helicopter, corporate jet, experimental aircraft, ground vehicle, military helicopter, military jet, military prop, multiple aisle jet, multiple engine prop, regional jet, single aisle jet, single engine prop

Tower

Table 9 - Binary Logit Results for Aircraft Type, Tower and TRACON

Variable	Odds Ratio	Standard Error
Civilian Helicopter	1.567*	0.354
Corporate Jet	1.684***	0.247
Experimental	7.646***	2.599
Ground	0.609	0.365
Military Helicopter	1.262	0.620
Military Jet	2.116*	0.638
Military Prop	1.881	0.813
Multiple Aisle Jet	1.031	0.168
Multiple Engine Prop	1.312*	0.151
Regional Jet	1.588***	0.212
Single Engine Prop	2.088***	0.214
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 N = 6,932. Base: Single Aisle Jet		

TRACON

Variable	Odds Ratio	Standard Error
Civilian Helicopter	0.802	0.291
Corporate Jet	1.278*	0.147
Experimental	1.705	0.841
Military Helicopter	0.758	0.335
Military Jet	0.850	0.151
Military Prop	1.195	0.344
Multiple Aisle Jet	0.996	0.200
Multiple Engine Prop	0.792	0.095
Regional Jet	1.022	0.118
Single Engine Prop	0.835	0.147
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 N = 13,634. Base: Single Aisle Jet		

The most striking finding is that incidents involving experimental planes had a much higher frequency of catastrophic severity levels than an even distribution would predict. Since experimental aircraft are often flown by recreational pilots from small airports with fewer controllers, these human factors may be the cause of the increased severity levels of incidents.

In general, particularly in Tower facilities, smaller planes such as single engine props, corporate jets, and regional jets were also had higher than incident severity levels than single aisle jets. One possibility is that small planes are typically flown by less experienced pilots into smaller regional airports. If large planes are frequently commercial planes flying into large airports, they may have the most experienced pilots being controlled by the most experienced controllers.

6.1.2. Control Status

Visual Approaches are associated with increased severity levels.

Control Status refers to how an aircraft is being controlled. Instrument Approach was the most common status, and is thus used as the baseline for Volpe's analysis. The most interesting results from this model are that incidents with visual approaches are 2.5 times more likely to be severe than the mean incident. A similar result was found in Section 6.1.3, where planes flying under visual flight rules were also shown to be more than twice as likely to be involved in severe incidents. It may be that planes on visual approaches are more likely to be flown by inexperienced pilots,[51] or it may be that instrument approaches have significant safety advantages. Where there is likely some truth to both explanations (among others), determining the specific causality and the associated relative weight of each would require follow-up analysis.

A binary logit model, as shown in Table 10, additionally showed that On Vector status is associated with increased severity in both Tower and TRACON Facilities. "No" control status is associated with increased severity in Tower facilities. SID (Standard Instrument Departure) /STAR (Standard Terminal Arrival Route) and Radar Advisories show increased severity in TRACON facilities.

[51] While pilots of all types fly visual approaches, pilots who do not (or do not yet) possess an instrument rating must necessarily fly visual approaches.

Table 10 - Control Status and Severity

Tower

Variable	Odds Ratio	Standard Error
None	1.542***	0.163
On SID/STAR	1.120	0.147
On Vector	1.873***	0.241
NORDO	1.719	1.320
On Route	1.182	0.286
Radar Advisories	1.473	0.298
Visual Approach	2.356***	0.281
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 N = 6,932; Base: Instrument Approach		

TRACON

Variable	Odds Ratio	Standard Error
None	1.316	0.211
On SID/STAR	1.755***	0.270
On Vector	1.829***	0.171
NORDO	3.178	2.525
On Route	0.942	0.158
Radar Advisories	1.600**	0.247
Visual Approach	1.884***	0.229
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 N = 13,634; Base: Instrument Approach		

6.1.3. Flight Plan

Flight plan variables designate the technology (and related procedures) that a pilot and controller are using to direct traffic. Summary statistics show that flight plan variables have a more consistent relationship with severity than controller status, although our final model does use variables from both groups.

VFR (Visual Flight Rules) incidents are likely to be more severe than IFR (Instrument Flight Rules) incidents.

The flight plan an aircraft is flying under had a noticeable relationship with severity. Flights with no flight plan were associated with a higher frequency of Catastrophic incidents – although the relationship lost statistical significance when other variables were introduced, as shown in Table 15 on page 46. Since flight plans are used by air traffic control to track flights, it may be that flights with no flight plan are more likely to have a miscommunication with air traffic control. IFR incidents were less likely to be severe than VFR incidents. This is striking, since IFR flight plans are required whenever visibility is poor; planes are only allowed to fly with VFR plans when visibility is good.[52] 79% of incidents are categorized as having an IFR flight plan, making up the majority of the incidents recorded (see Table 11), demonstrating that while incidents involving IFR are more common, incidents with VFR are more serious. Also, since pilots need a higher level of certification to fly IFR, it is possible that these more experienced pilots are less likely to be involved in severe incidents.

Regression analysis confirmed that incidents with IFR flight plans are less likely to be Hazardous or Catastrophic in the tower space, and are also less likely to be Catastrophic in the TRACON space. Incidents categorized as VFR are more likely to be fall into the Catastrophic category. Incidents with no flight plan were also substantially more likely to be Catastrophic.[53] Table 11 illustrates these results further, showing that IFR incidents are less likely to be Hazardous or Catastrophic than an equal distribution would predict.

Table 11 - Flight Plan, Actual and Expected Frequency[54]

	Minimal	Minor	Major	Hazardous	Catastrophic
IFR Actual	10214	5030	1851	242	60
IFR Expected	*10391*	*4784*	*1778*	*287*	*158*
VFR Actual	1213	371	153	67	78
VFR Expected	*1124*	*518*	*192*	*31*	*17*
None Actual	254	64	22	9	25
None Expected	*228*	*100*	*37*	*6*	*3*

6.1.4. Number of Aircraft

The more planes involved, the higher the typical severity of an incident.

The vast majority of incidents documented had one or two aircraft involved. A frequency distribution of the number of aircraft in an incident is shown in Table 12. Higher severity is associated with a larger number of planes in an incident; this association is likely related to how severity is categorized, since the probability of an accident increases when more planes are in close proximity.

[52] FAA VFR Minimums, accessed 3/12/2014:http://www.faasafety.gov/files/gslac/courses/content/25/185/VFR%20Weather%20Minimums.pdf

[53] Regression results are shown in the Appendix C.

[54] Based on Chi-square calculations

Since the data series has a long, dispersed, tail, a multinomial analysis of individual severity classes did not yield useful results. Very few incidents had more than two planes involved, and those instances were widely dispersed across severity categories. Thus, in this case, a binary analysis that categorizes incidents as either severe or not severe yielded stronger results. In both tower and TRACON spaces, a larger number of aircraft was associated with severe incidents (Table 12).

Table 12 - Number of Aircraft, Actual and Expected Frequency

Number of Aircraft	1	2	3	Chi2 score	P Value
Tower Actual	329	476	34	113	0
Tower Expected	*468*	*343*	*28*		
TRACON Actual	560	1072	55	249	0
TRACON Expected	*859*	*790*	*41*		

6.1.5. Phase of Flight

Incidents with VFC traffic patterns have higher severity levels than other phases of flight.

Phase of flight describes what an airplane is doing – for example if it is climbing, level, or descending. For the most part, the relationship between phase of flight and severity is fairly subtle, with the exception of a few of the less frequent variables. Incidents were most commonly recorded in the climbing, departure, descending, and level flight phases. Figure 6 shows that Arrival is the most common phase of flight for ATSAP incidents to occur in.

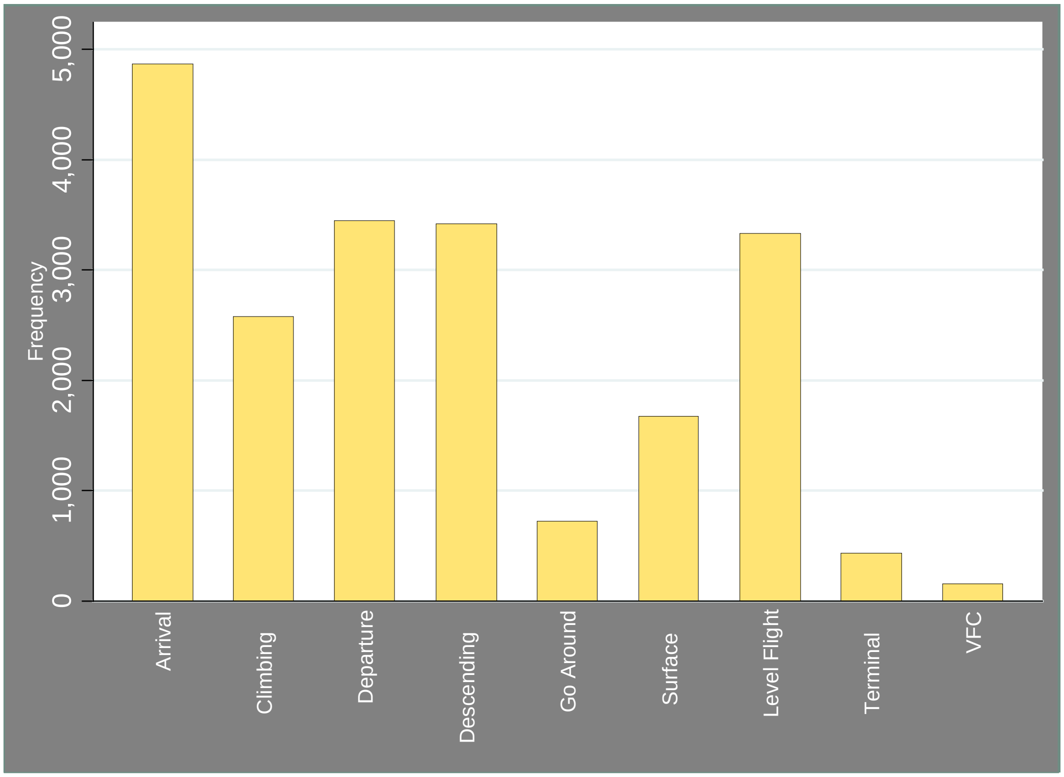

Figure 6 - Airborne Incidents by Phase of Flight

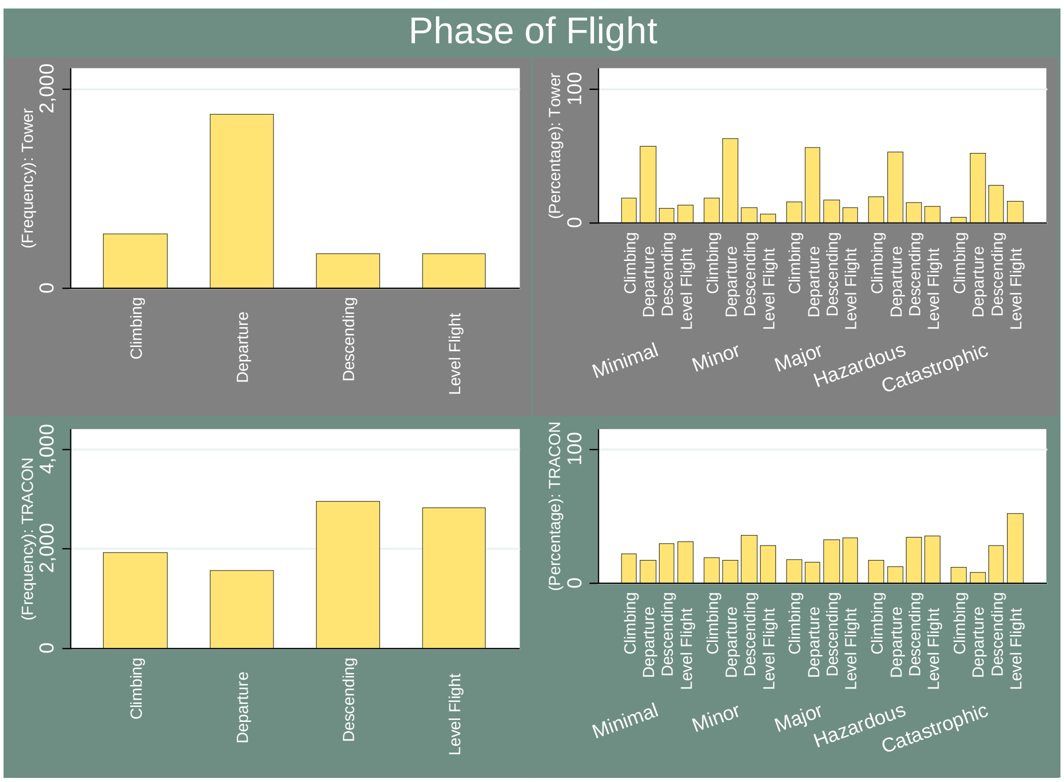

Figure 7 - Incident Severity by Phase of Flight (Major Categories)

Logit regressions,[55] performed separately for Tower and TRACON facilities, show that while some of the variables are statistically significant, differences are slight in most cases. One notable finding is that incidents with an aircraft on the surface are significantly more likely to be Catastrophic in both the tower and TRACON space. This may be because such incidents, especially in a TRACON facility, are likely a crash. For example, one such incident is described as a loss of engine power. In Tower facilities, VFC traffic patterns have an increased severity distribution that becomes more pronounced at each step (Minor to Major, Major to Hazardous, etc.).

[55] Regression results are in Appendix C

Table 13- Phase of Flight Binary Analysis[56]

Tower

Variable	Odds Ratio	Standard Error
Arrival	1.580***	0.192
Climbing	1.091	0.177
Descending	2.059***	0.336
Go Around/Missing Approach	1.740***	0.252
Level Flight	1.346	0.297
Incidents with an Aircraft on the Surface	1.289*	0.161
Terminal Enroute Transition	1.739	0.554
VFC Traffic Pattern	2.734***	0.691
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 N = 6,932; Base: Departure		

TRACON

Variable	Odds Ratio	Standard Error
Arrival	1.053	0.110
Climbing	1.101	0.107
Descending	1.352**	0.128
Go Around/Missing Approach	1.191	0.272
Level Flight	1.501***	0.151
Incidents with an Aircraft on the Surface	1.653	0.628
Terminal Enroute Transition	0.742	0.156
VFC Traffic Pattern	1.683	0.880
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 N = 13,634; Base: Departure		

[56] Full regression results are shown in Appendix D

6.1.6. Special Events

Emergency situations and special events have a higher frequency of severe incidents. Data on TFM is inconclusive.

Causal factor variables for special events were aggregated into three categories so that each category would have a sufficient number of incidents to fit into a regression analysis. The aggregated categories are emergency situations,[57] special events,[58] and TFM (traffic management initiatives). Table 14 shows the results of a binary logit model run separately on each aggregated category.

Not surprisingly, incidents during emergency situations had substantially higher severity, especially Hazardous and Catastrophic occurrences. Since this category by definition consists of unplanned, sudden events where there may be a loss of control such as in an emergency landing, this correlation is not surprising, as shown in Table 14.

Table 14- Binary Logit Results during Emergency Situations

Tower

Variable	Odds Ratio	Standard Error
Emergency Situation	5.645***	1.040
Special Event	2.241***	0.462
TFM Initiative	0.379	0.225
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 **N = 3,692**		

TRACON

Variable	Odds Ratio	Standard Error
Emergency Situation	2.041**	0.490
Special Event	2.390***	0.335
TFM Initiative	1.877*	0.558
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 **N = 7,575**		

[57] Emergency situations include controller declared emergency, emergency landing, expedited handling, operator declared emergency, pilot declared emergency, other emergency

[58] Special events include airshow, flight check operations, large event, openskies photo, VIP, and skydiving

Figure 8 indicates that every sub-category of emergency situation had a much higher frequency of Catastrophic incidents than the ATSAP database average. Due to the small number of observations in each emergency category (for example, some of the emergency categories have too few observations for a model to converge) the emergency variables were aggregated together for modeling.

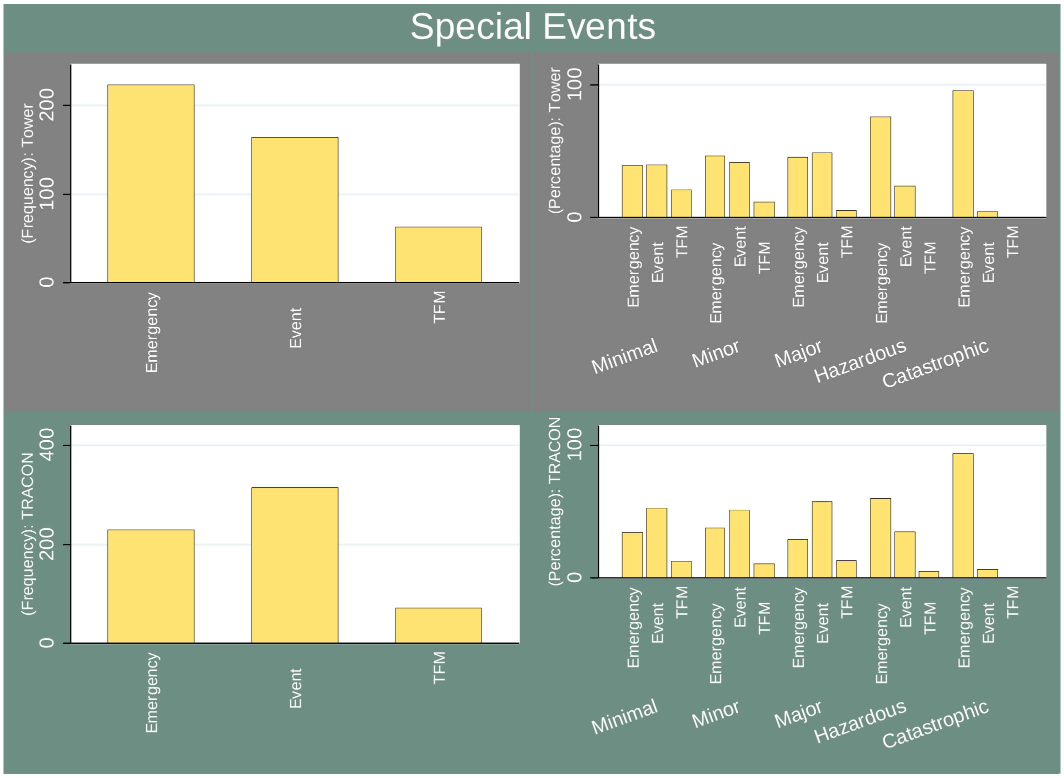

Figure 8 - Special Events

Special events, which include VIP flights, skydiving, and airshows, had a higher occurrence of Major incidents in both tower and TRACON spaces, and additionally had a higher frequency of Hazardous incidents in the case of the TRACON space. There was not statistically significant relationship with Catastrophic incidents.

An individual analysis of individual special event variables yielded little of significance, because of the small number of observations in each category. The only notable result was that openskies photos (when a plane is flying to be photographed for publicity) are associated with higher instances of Major and Hazardous results.

Data on TFMs were inconclusive. There was a positive relationship between TFM's and severity level in TRACON facilities, but not in Tower facilities. The lack of a correlation between TFMs and increased severity in tower facilities may speak to the effectiveness of TFMs. Disaggregated individual types of TFM events did not have a significant relationship with severity, due to the small number of observations for each type of TFM.

6.1.7. Aircraft Model

The next several sections put together the prior data explanations into a model that incorporates all of the relevant aircraft variables. Since the Special Event variables only started being recorded in 2011, while the other variables are available from 2008, binary logit models were run in two sets (with and without the special event variables for 2008-2013 and 2011-2013). The small number of observations for special events caused the multinomial logit models to fail to converge when special events were included; thus, the multinomial models were run without the special event variables. Multinomial model results without the special event variables are shown in this section; binary models as well as multinomial models with special events included can be viewed in the appendix.

6.1.7.1. Tower Multinomial Logit Model

Experimental Planes are tied to Catastrophic incidents.

Incidents involving experimental planes have a seven-times higher risk of a severe (major, hazardous, or catastrophic) incident than single-aisle jets (the base case), and have a 21 times higher risk of catastrophic incidents. Corporate jets and single engine props were also associated with increased severity, although not to as great an extent. It may be that small planes are more likely to have less experienced pilots, and be flown into smaller airports with less experienced controllers.

Similarly, on vector and visual approach control statuses were also associated with increased severity levels. On vector approaches had a strikingly high number of major incidents, while visual approaches were tied to a high number of catastrophic incidents. Since visual approaches are often used by less experienced pilots, this may be tied to the level of a pilot's skill.

The special event version of the binary models showed that emergency situations and special events both retained a statistically significant relationship with increased severity even when other variables are held constant, while Traffic Management initiatives were not significantly tied to incident severity. This makes sense, since traffic management initiatives reduce the number of planes that fly into an airport during bad weather or other special circumstances.

Table 15 - Tower Aircraft Multinomial Model (no special events)

	Minor	Major	Hazardous	Catastrophic
Civilian Helicopter	1.1 (.23)	1.1 (.31)	1.7 (.76)	3* (1.6)
Corporate Jet	.94 (.12)	1.5* (.26)	1.1 (.4)	1.1 (.64)
Experimental Plane	.9 (.47)	4** (1.9)	3.8 (3)	21*** (13)
Ground Vehicle	.73 (.27)	.43 (.44)	3.9e-07*** (1.6e-07)	2.8 (2.4)
Military Helicopter	.68 (.28)	.61 (.46)	2.2 (1.6)	3.1 (3.5)
Military Jet	1.1 (.28)	1.4 (.51)	3.3* (1.7)	2.8 (2.2)
Military Prop	1.1 (.39)	2 (.93)	1.6 (1.7)	8.3e-07*** (4.0e-07)
Multiple Aisle Jet	1.4* (.23)	.93 (.21)	1.4 (.61)	.77 (.8)
Multiple Engine Prop	.92 (.1)	1 (.17)	1.4 (.35)	1.7 (.7)
Regional Jet	1.1 (.11)	1.4* (.22)	1.3 (.44)	.38 (.39)
Single Engine Prop	1 (.1)	1.4* (.2)	2.2*** (.53)	3.6*** (1.3)
Control Status: Instrument Approach	1.6** (.24)	1.2 (.29)	1.6 (.69)	8*** (3.6)
Control Status: None	.98 (.1)	1 (.17)	1.4 (.46)	2* (.64)
Control Status: NORDO	.63 (.5)	1.3 (1.4)	2.1e-07*** (9.7e-08)	6.3 (7.3)
Control Status: On Route	.68 (.16)	1.2 (.38)	.79 (.53)	1.8e-06*** (7.4e-07)
Control Status: On SID/STAR	1.2 (.15)	1.4 (.28)	1.2 (.6)	.61 (.65)
Control Status: On Vector	1.1 (.15)	2*** (.39)	1.6 (.68)	2.1 (1.6)
Control Status: Radar Advisories	.54** (.11)	.78 (.23)	.45 (.24)	2.1 (1)
Control Status: Visual Approach	1.1 (.16)	1.9*** (.34)	2 (.82)	3.6** (1.6)
Flight Plan: IFR	1.3 (.25)	2.9** (1)	3.1 (2.4)	.36 (.22)
Flight Plan: None	1.6 (.37)	2.7* (1.3)	3.8 (3.4)	1.7 (1.2)
Flight Plan: Unknown	1.2 (.3)	3.3** (1.4)	4.7 (3.9)	1.3 (.59)
Flight Plan: VFR	1.8** (.35)	3.4** (1.3)	6.1* (4.7)	1.2 (.74)
Number of Aircraft	1.7*** (.12)	1.8*** (.16)	1.8*** (.18)	.11*** (.052)
Phase of Flight: Arrival	1.4* (.22)	1.7 (.53)	1.5 (.73)	1.6 (1)
Phase of Flight: Climbing	1.5* (.28)	1.3 (.43)	1.6 (.82)	.72 (.88)
Phase of Flight: Departure	1.7** (.26)	1.5 (.45)	1.3 (.62)	2.3 (1.3)
Phase of Flight: Descending	1.4 (.29)	2.3* (.8)	1.5 (.83)	3.6 (2.5)
Phase of Flight: Go Around/Missing Approach	1.5* (.28)	1.9* (.62)	1.5 (.86)	2.3 (1.7)
Phase of Flight: Level Flight	.91 (.16)	1.9 (.65)	1.3 (.81)	1.4 (1.1)
Phase of Flight: Surface	1.3 (.19)	1.2 (.34)	1.2 (.52)	6.7*** (3.8)
Phase of Flight: Terminal Enroute Transition	.6 (.26)	2.1 (.94)	2.4 (1.6)	8.0e-07*** (5.5e-07)
Phase of Flight: VFC Traffic Pattern	1.6 (.46)	2.2 (.99)	2.4 (1.4)	5.3* (3.9)
Total Operations	1* (.00075)	1 (.00092)	1 (.0014)	.99 (.0028)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 6,874				

6.1.7.2. TRACON Multinomial Logit Model

Corporate and Regional Jets have increased rates of Major incidents in TRACON facilities. Experimental planes have high rates of catastrophic incidents.

SID/STAR and Visual Approaches also have higher than expected rates of Major incidents.

The TRACON aircraft models confirmed the high relationship between experimental planes and increased severity levels that were seen in other models. Since this result has appeared in several models with different specifications, it ought to be considered carefully.

Multinomial logit models showed that the increased severity seen for corporate and regional jets in TRACON facilities was strongest for Major incidents. This is a slight difference from Tower facilities, where corporate and regional jets had the most elevated risk ratios in the Hazardous category.

SID/STAR and Visual Approaches also had elevated risk ratios for Major incidents. This parallels the results found in Tower facilities.

The special events versions of the binary models showed that, similarly to Tower facilities, special events and emergency situations had a significant correlation with increased severity levels, while Traffic Management initiatives were not significantly tied to incident severity. The parallel results in both the Tower and TRACON spaces provide an additional measure of statistical validity for this result.

Table 16 - TRACON Aircraft Multinomial Model (no special events)

	Minor	Major	Hazardous	Catastrophic
Civilian Helicopter	.99 (.28)	.57 (.3)	.55 (.52)	4.4* (2.7)
Corporate Jet	1 (.066)	1.3* (.15)	.74 (.19)	.28 (.33)
Experimental Plane	.47 (.28)	1.1 (.68)	2.1 (2.2)	22*** (17)
Military Helicopter	1.2 (.44)	.76 (.36)	.81 (.84)	2.2e-08*** (1.2e-08)
Military Jet	1.2 (.17)	1 (.22)	.44 (.32)	2.9e-08*** (1.4e-08)
Military Prop	1 (.21)	1.4 (.4)	.62 (.67)	6.4e-08*** (3.8e-08)
Multiple Aisle Jet	1.6*** (.19)	.95 (.22)	1.8 (.57)	7.0e-08*** (3.1e-08)
Multiple Engine Prop	1 (.083)	.81 (.12)	.57 (.2)	2.5 (1.2)
Regional Jet	.9 (.088)	.93 (.12)	.98 (.3)	5.8e-08*** (2.1e-08)
Single Engine Prop	1.2*** (.084)	.8 (.16)	1.1 (.32)	6*** (2.8)
Control Status: Instrument Approach	2.4*** (.35)	1.8** (.34)	1.1 (.5)	2.9 (2.5)
Control Status: None	1.8** (.33)	1.3 (.39)	2.8* (1.5)	2.4 (2)
Control Status: NORDO	1.3 (.9)	3.1 (3)	13* (14)	3.9e-08*** (4.0e-08)
Control Status: On Route	1.1 (.18)	1.1 (.29)	.93 (.46)	1.8 (1.7)
Control Status: On SID/STAR	1.3 (.22)	1.9** (.41)	1.2 (.58)	6.3e-08*** (5.5e-08)
Control Status: On Vector	2.4*** (.38)	2.6*** (.44)	1.3 (.47)	1.3 (1.3)
Control Status: Radar Advisories	1.6* (.35)	1.6 (.45)	2 (1.2)	2.1 (1.9)
Control Status: Visual Approach	2.1*** (.34)	2.5*** (.5)	2.1 (.89)	2.4 (2.1)
Flight Plan: IFR	1.7*** (.19)	1.8 (.67)	1.5 (.65)	2 (1.8)
Flight Plan: None	.58 (.2)	2 (1)	1.4 (1.7)	8.5 (9.7)
Flight Plan: Unknown	.74 (.27)	1.3 (.81)	3.7 (2.7)	9.6* (10)
Flight Plan: VFR	.87 (.15)	1.7 (.78)	1.7 (1.2)	5.9* (5.1)
Number of Aircraft	1.7*** (.089)	2.1*** (.15)	2*** (.23)	1.6e-08*** (6.9e-09)
Phase of Flight: Arrival	1.3 (.21)	.84 (.3)	.87 (.36)	3.9 (4.1)
Phase of Flight: Climbing	.82 (.17)	.83 (.29)	.65 (.31)	.63 (.71)
Phase of Flight: Departure	.89 (.18)	.84 (.32)	.56 (.25)	.72 (.85)
Phase of Flight: Descending	1 (.21)	.99 (.32)	.9 (.42)	1.7 (1.6)
Phase of Flight: Go Around/Missing Approach	1.3 (.3)	.88 (.36)	1.8 (.96)	2.3 (3.3)
Phase of Flight: Level Flight	.89 (.17)	1.3 (.41)	1 (.39)	1.2 (1.3)
Phase of Flight: Surface	1.3 (.54)	1.8e-12*** (8.0e-13)	1.9 (1.8)	51*** (58)
Phase of Flight: Terminal Enroute Transition	.54* (.14)	.69 (.26)	.38 (.28)	1 (1.3)
Phase of Flight: VFC Traffic Pattern	.72 (.35)	.44 (.33)	1.8e-09*** (1.2e-09)	8.8 (11)
Total Operations	1*** (.00033)	1*** (.00042)	1** (.00052)	1 (.0011)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 **N = 13,832**				

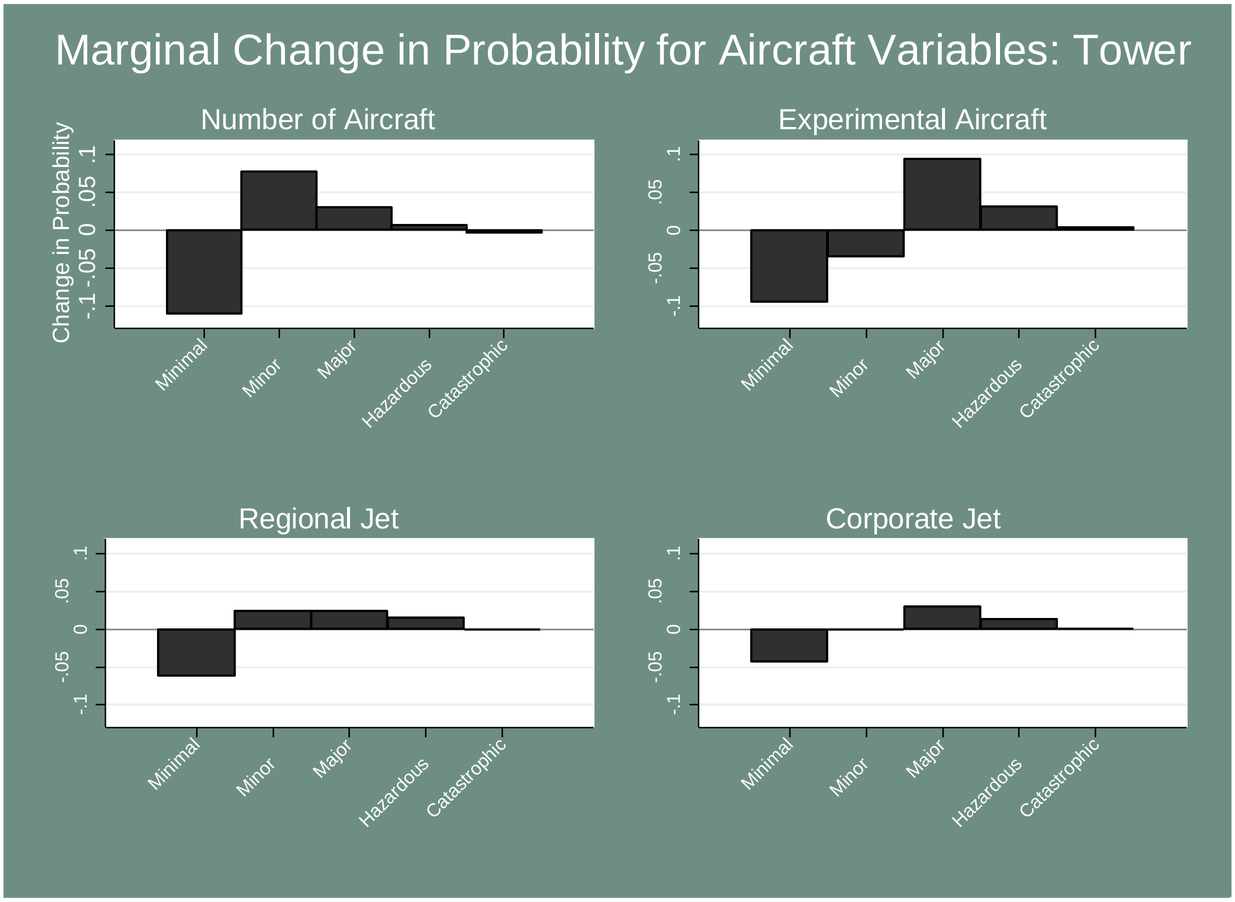

Figure 9 - Experimental Planes have Higher Odds of Severe

Figure 9 shows the marginal change in probability as the incident moves from one severity level to the next. The graph illustrates one of the more dramatic results from this model – incidents involving experimental aircraft have a higher than expected severity distribution – especially in Tower facilities. Unsurprisingly, each additional aircraft involved in an incident also increases the likelihood of a more severe incident.

6.2. Facility Characteristics Model

These variables describe characteristics of the facilities where incidents occurred. Data for this section is a combination of ATSAP and facility specific data detailed in Section 4. Certain facility characteristics will remain static across incidents shifting the focus to be on the variation between these variables across facilities.[59] All variables examined in this section are split between tower and TRACON facilities, due to the significant differences between the two facility types. In other words, studying facility characteristics combined across tower and TRACON will obfuscate nuanced difference between the two, and could lead to incorrect conclusions about certain variables.[60] It should also be noted that due to high collinearity between sub-categories in this section, the final facility based model will focus only on a subset of these variables.

6.2.1. Facility Complexity

6.2.1.1. ATC Level

ATC levels are a sliding scale based on the complexity of that particular facility's airspace. Therefore, higher ATC levels are a loose proxy for the overall complexity of that facility. Figure 10 presents the distribution of operations over ATC levels, while Figure 11 presents the distribution ATC levels across severity levels. Similar to the runway count variable, this variable is conditional on an incident having occurred at a facility, not the ATC levels at all facilities.

---

[59] Variables that remain static across time for each facility include the number of runways and ATC pay-level.

[60] An example of this is runway count. TRACON facilities control many more runways, on average, than tower facilities. Therefore, it would be misleading to combine the two facilities types and attempt to interpret results involving runway count and severity because of this inherent difference in runway management.

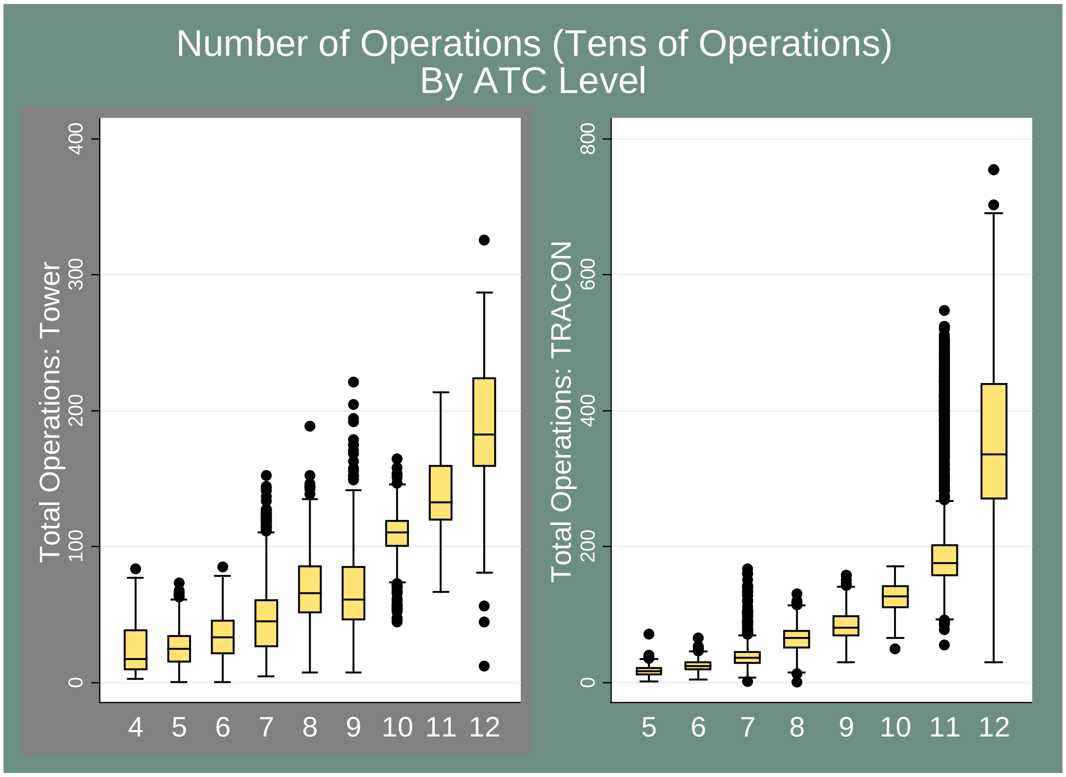

Figure 10 - Distribution of Operations by ATC Levels

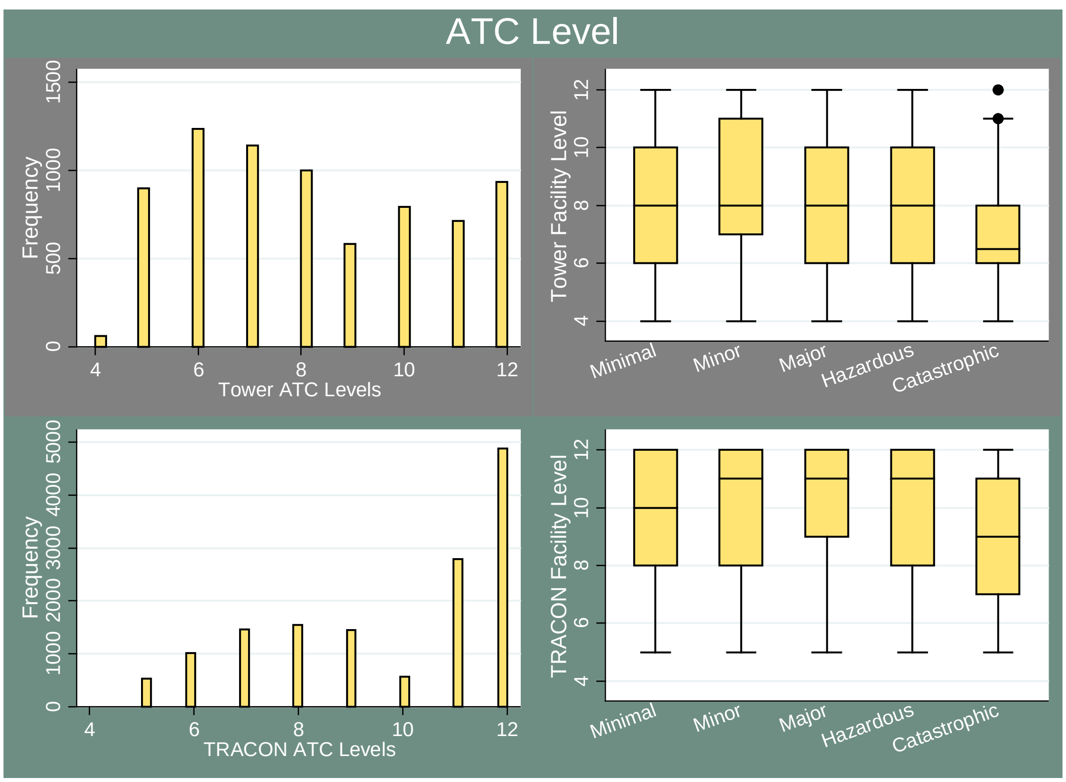

Figure 11 - Distribution of ATC Levels

As expected, the ATC level rises as operations increase, suggesting higher ATC levels are associated with more complex (busier) terminal airspace. Results from a Kruskal-Wallis test on the ATC level indicate that the severity levels are jointly different from each other for both facility types. The results from the pairwise test for tower and TRACON facilities are presented in Table 17 and Table 18 below. In both cases, the relationship between Catastrophic outcomes and the other severity levels are all significantly different from each other (excluding Minimal and Catastrophic for TRACON facilities).

Table 17 - Kruskal-Wallis Test of ATC Level for Tower Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major					
Hazardous		X			
Catastrophic	X	X	X	X	

Table 18 -Kruskal-Wallis Test of ATC Level for TRACON Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major	X				
Hazardous	X				
Catastrophic		X	X	X	

Table 19 presents the results the single variable logit of ATC levels on severe/non-severe incidents by facility type. The logit results show that there is a statistically significant relationship between the ATC levels and severity for both facility types. Again, the direction of the odds ratio differ between the facility type, with higher ATC levels reducing the probability of a severe event for towers and increasing the probability for TRACONs.

Table 19 - Logit Estimate of the ATC Level by Facility Type

Variable	Odds Ratio	Standard Error	Obs
ATC Level - Tower	0.962***	0.020	6,899
ATC Level - TRACON	1.116***	0.028	13,742
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.1.2. Daily Operations

Daily operations were taken from OPSNET and provide an accurate measure for the volume of traffic at facility on any given date. Unlike ATC levels and runway numbers, operations are expected to vary on a day-to-day, if not hour-to-hour basis for each facility. Ideally, hourly data would be preferred but due to data quality and accessibility these data were not available for this study. Figure 12 presents the distribution of daily operations over severity for both facility types.

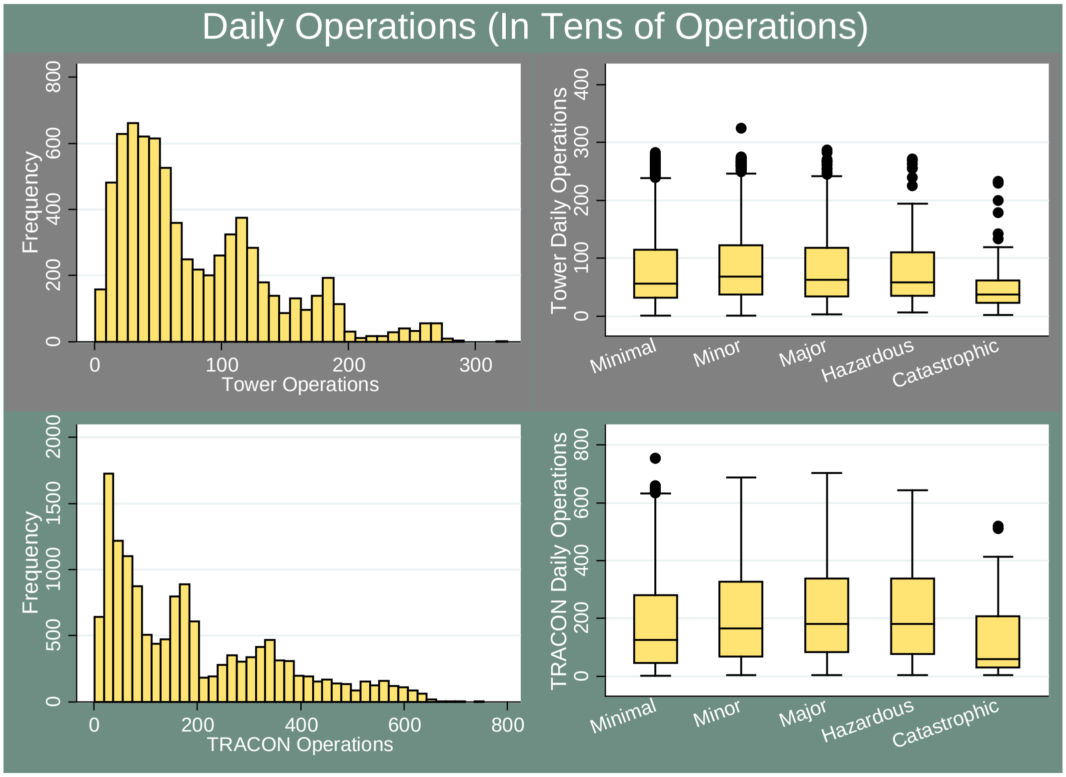

Figure 12 - Distribution of Daily Operations

The results from the Kruskal-Wallis test indicate that daily operations jointly differ across severity levels for both facility types. A Pairwise comparison is presented for both facilities in Table 20 and Table 21. Again, Catastrophic outcomes are significantly different than the other severity levels in both facility types.

Table 20 - Kruskal-Wallis Test of Daily Operations for Tower Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major					
Hazardous					
Catastrophic	X	X	X	X	

Table 21 - Kruskal-Wallis Test of Daily Operations for TRACON Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major	X	X			
Hazardous	X				
Catastrophic	X	X	X	X	

Results from a single variable logit on daily operations can be found in Table 22. In both cases daily operations has a statistically significant relationship with severity, but odds ratios go in opposite directions. For tower incidents, an increase in 1,000 daily operations would *reduce* the odds of a severe event by about 20%, while it would increase the odds of a severe event by about 10% for TRACON events. The disparity between the different facility types could be due in large part to the simplicity of the single variable logit estimation holding all other relevant factors constant. Therefore, these results need to be taken with caution, and it is expected after controlling for other variables the relationship between daily operations and severity will change (see Section 6.2.3).

Table 22 - Logit Estimate of Daily Operations (in Tens of Operations) by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Daily Operations - Tower	0.998	0.001	6,874
Daily Operations - TRACON	1.001***	0.000	13,832
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.1.3. Number of Runways

This variable indicates the total number of runways at the airport where the incident occurred. Note that this is not the number of runways in operation at the time of the incident due to lack of data availability. Figure 13 presents the distribution of the number of runways for tower and TRACON events.

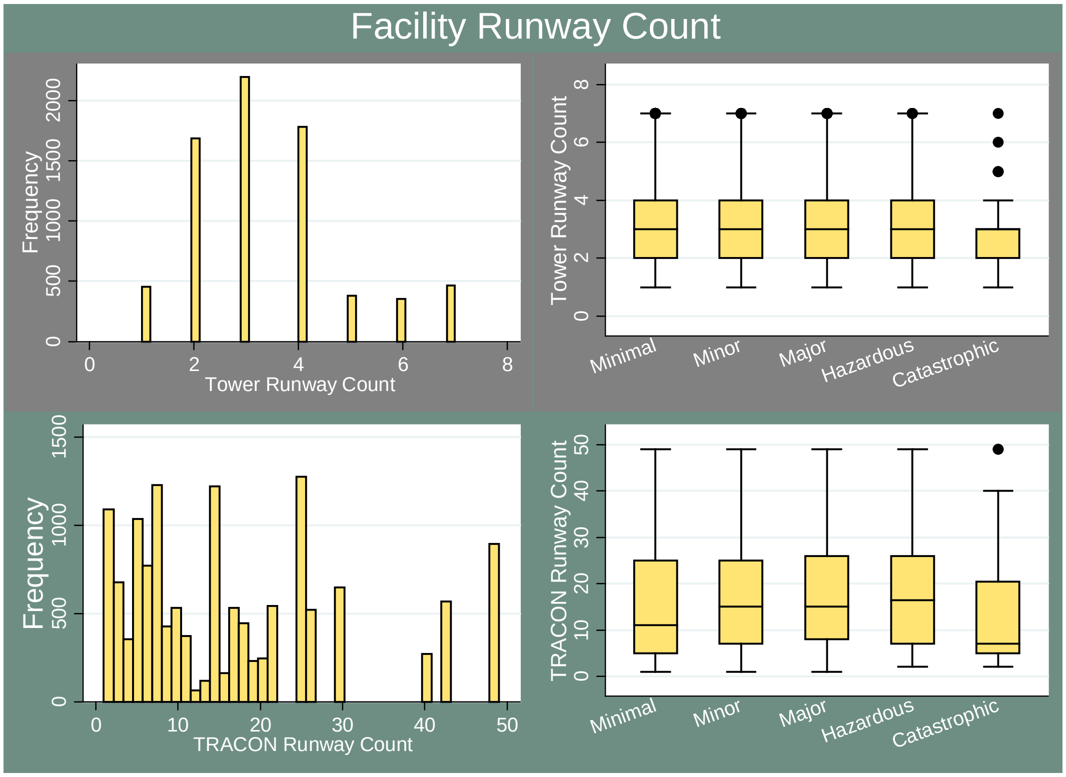

Figure 13 - Distribution of Number of Runways

Again, the distribution shown here are not the total number of runways at all airports; it is only the total count of runways at airports where an incident occurred, and airports with multiple incidents in the dataset are counted multiple times. The results from a Kruskal-Wallis test on the number of runways indicate that the five severity levels are jointly different from each other for both tower and TRACON facilities.[61] A pairwise comparison test for both facilities types is presented in Table 23 and Table 24 below.

Table 23 - Kruskal-Wallis Test of Runway Count for Tower Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major					
Hazardous					
Catastrophic	X	X	X		

[61] Results from the Kruskal-Wallis test can be found in Appendix E.

Table 24 - Kruskal-Wallis Test of Runway Count for TRACON Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major	X				
Hazardous	X				
Catastrophic	X	X	X	X	

Table 25 presents the results from a single variable logit of runway count on severe/non-severe incidents. Both logit estimations indicate a relationship between the number of runways and severity, however, the direction differ between the facility types. These results suggest that for tower incidents, more runways are associated with less severe incidents, while the opposite is true for TRACON incidents.

Table 25 - Logit Estimate of the Number of Runways by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Runway Count - Tower	0.921**	0.031	6,864
Runway Count - TRACON	1.016***	0.004	13,742
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.1.4. Traffic Complexity Rating

The relative distribution of severity over both traffic rating variables shows a clear difference between Catastrophic and the other four severity measures.

Complimentary to the traffic volume rating variable, traffic complexity rating is a non-mandatory field in ATSAP that represents the reporter's perceived traffic complexity during the time of the incident. Traffic complexity does not necessarily mean heavy traffic; however, a high correlation between the two variables would suggest there is a causal link between the two.[62] Careful consideration will be given when deciding whether to include one or the other or possibly both in the final model.

[62] Estimated correlation values of 0.843 and 0.849 for Tower and TRACON incidents, respectively.

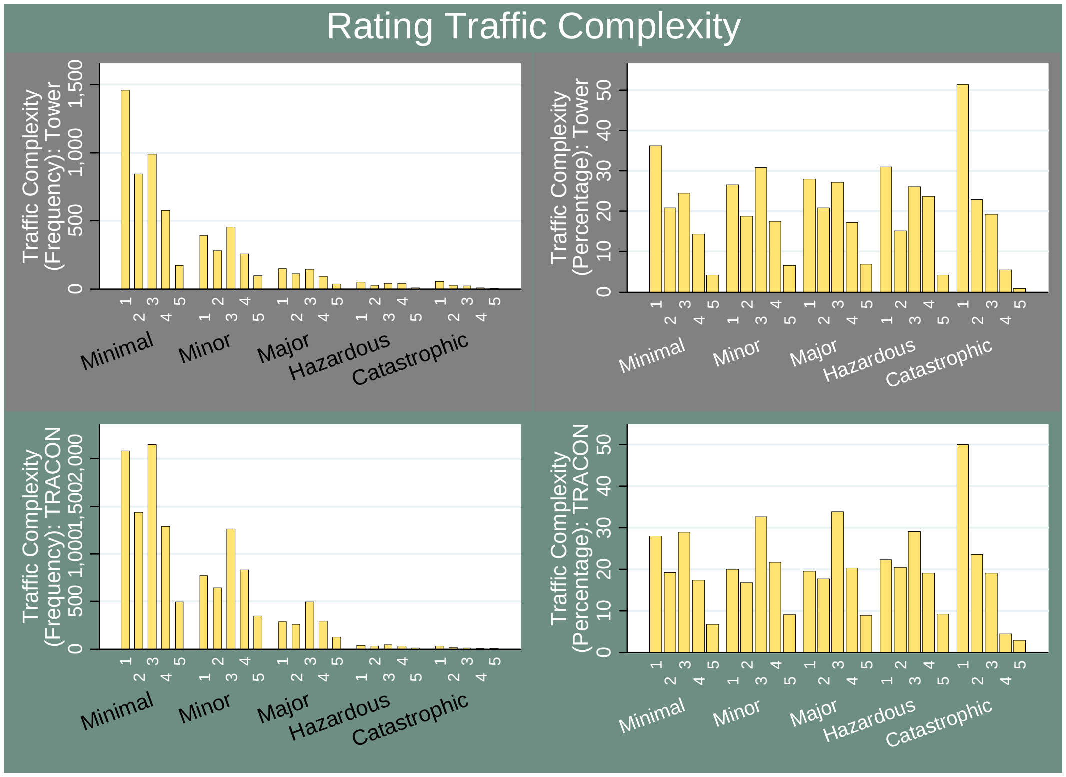

Figure 14 - Rating Traffic Complexity

An interesting takeaway from the distribution presented above is that it appears Catastrophic incidents are largely made up of low traffic complexity ratings. A Kruskal-Wallis test reveals that traffic complexity ratings jointly differ from each across severity levels for both facility types and a pairwise comparison is presented in Table 26 and Table 27 for both facility types. This test confirms the findings from the distributions above that Catastrophic incidents appear to be different in nature relative to the other severity outcomes.

Table 26 - Kruskal-Wallis Test of Traffic Complexity for Tower Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major	X				
Hazardous					
Catastrophic	X	X	X	X	

Table 27 - Kruskal-Wallis Test of Traffic Complexity for TRACON Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major	X				
Hazardous					
Catastrophic	X	X	X	X	

Results from the single variable logit of traffic complexity ratings on severe/non-severe incidents are presented in Table 28. There is no statistically significant relationship between traffic complexity ratings and severity for tower incidents, while there is a statistically significant relationship for TRACON events, with an odds ratio suggesting close to an 8% increase in the odds of a severe incident occurring for an incremental change in the complexity rating.

Table 28 - Logit Estimate of Traffic Complexity Ratings by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Traffic Complexity Rating - Tower	1.049	0.033	6,323
Traffic Complexity Rating - TRACON	1.079*	0.033	13,009
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.1.5. Traffic Volume Rating

This variable is a non-mandatory field in ATSAP that the reporter may fill out representing how high, on a scale from 1-5, they perceived traffic volume to be during the time of the incident. This variable serves as additional measure of the overall traffic volume and theoretically provides information at the time of the incident as opposed to the whole day, which is the case with the daily operations count data.

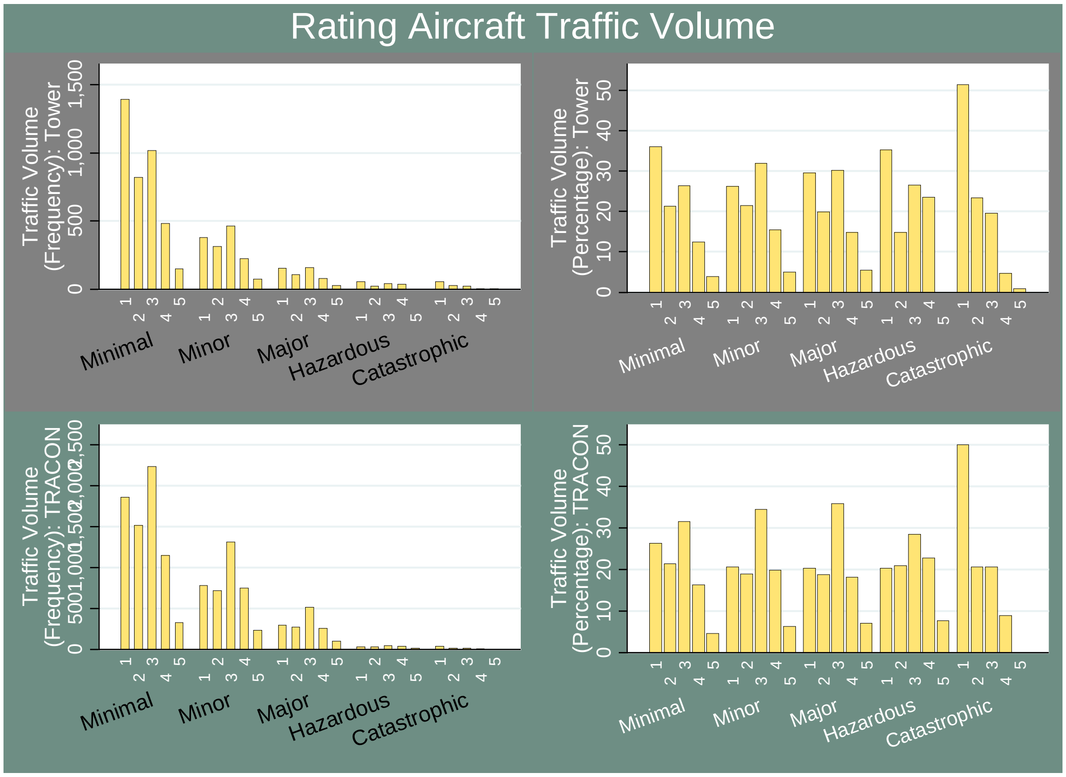

Figure 15 - Rating Aircraft Traffic Volume

Similar to the Rating Traffic Complexity, there are a relatively high percentage of incidents with reported low rated traffic volumes falling into the Catastrophic category in Figure 15. The results from a Kruskal-Wallis test indicate that traffic volume ratings jointly differ across severity levels for both facility types. A pairwise comparison is presented in Table 29 and Table 30 for both facility types. This test also confirms that Catastrophic incidents appear to be different across all other severity outcomes.

Table 29 - Kruskal-Wallis Test of Traffic Volume for Tower Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major	X				
Hazardous					
Catastrophic	X	X	X	X	

Table 30 - Kruskal-Wallis Test of Traffic Volume for TRACON Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major	X				
Hazardous					
Catastrophic	X	X	X	X	

Results from the single variable logit of traffic volume ratings on severe/non-severe incidents are presented in Table 31. There appears to be no statistical relationship in the likelihood of a severe event given an increase in the traffic volume ratings for tower incidents. The logit results for TRACON incidents do show a statistical relationship between traffic volume ratings and an increase in the likelihood of a severe loss of separation incident.

Table 31 - Logit Estimate of Traffic Volume Ratings by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Traffic Volume Rating - Tower	1.018	0.036	6,103
Traffic Volume Rating - TRACON	1.082**	0.032	12,560
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2. Organizational and Airspace Complexity Factors

This section contains causal factor variables connected to facility characteristics. These variables are complimentary to variables discussed in the previous section. Causal factors were first examined at an individual, disaggregated level to determine if the variables had an appropriate amount of statistical variation to be left as-is. Not all causal factors are examined here due to either lack of relevance or a severe lack of variation even after aggregation, and a complete list of causal factors can be found in Appendix D. As a final note, not all figures are presented in this section for sake of brevity, but all figures can be found in Appendix F.

6.2.2.1. Abnormal Configuration

This variable is the aggregated group of variables that refer to idea that the operational configuration of the control environment contributed to a loss of separation incident. Results from a single variable logit of control environment configuration on severity are presented in Table 32. While the odds ratios for the different facilities are in opposite directions, neither result can be determined statistically different from zero.

Table 32 - Logit Estimate of Configuration by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Configuration Contributed to the Event - Tower	0.700	0.135	6,932
Configuration Contributed to the Event - TRACON	1.170	0.133	13,859
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2.2. Aircraft Performance or Pilot Action Complexity Factor

Aircraft/Pilot action complexity factors increase the likelihood of a severe incident for Tower facilities.

The aircraft performance or pilot variable is a non-mandatory field in ATSAP where the reporter may indicate true or false if the complexity of the aircraft performance or pilot action was a significant factor during the incident. Figure 16 presents the distribution of the aircraft performance or pilot action complexity factor variable over severity. There is a large difference between the two facility types in terms of percentage true/false, with a much higher percentage of "true" cases for higher level of severity in Tower events.

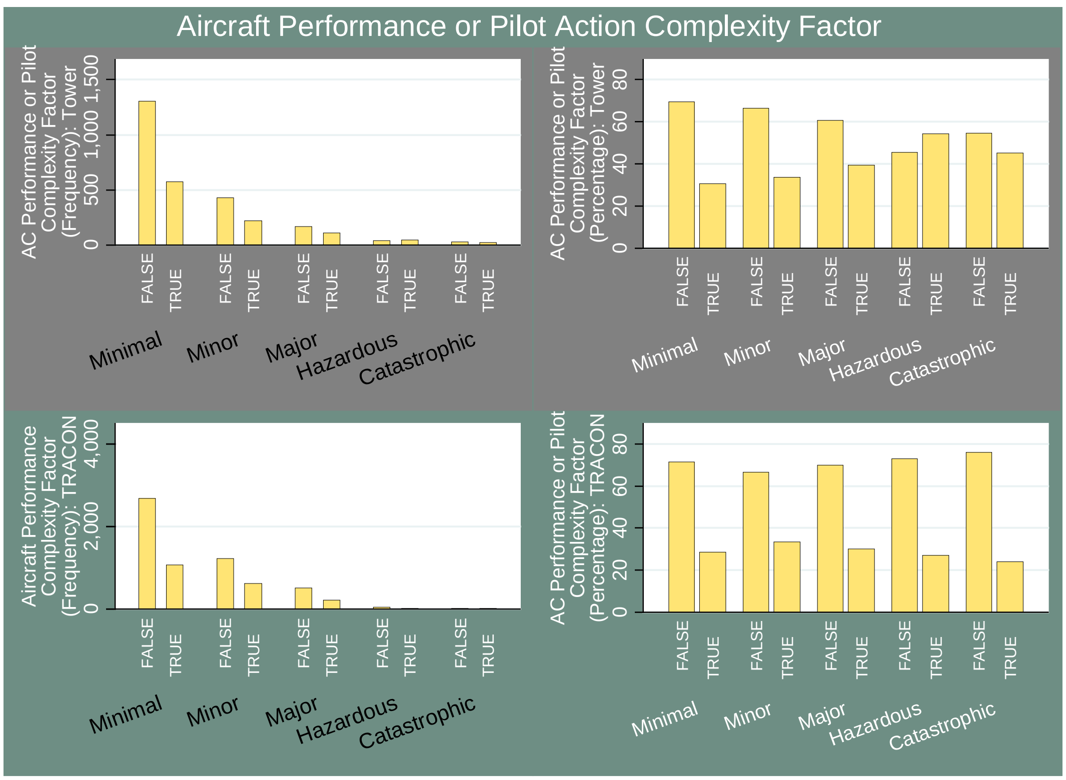

Figure 16 - Aircraft Performance or Pilot Action Complexity Factor

Table 33 reports the single variable logit of this complexity factor on severity. Tower incidents have a statistically significant relationship with severe and non-severe events, with an odds ratio suggesting a nearly 66% increase in the likelihood of a severe event. TRACON events, however, are not statistically significant.

Table 33 - Logit Estimate of Aircraft Performance Complexity Factor by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Aircraft/Pilot Action Complexity Factor - Tower	1.657***	0.213	2,931
Aircraft/Pilot Action Complexity Factor - TRACON	0.978	0.091	6,427
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2.3. Airspace and Procedure Complexity Factor

The airspace and procedure complexity factor variable is a non-mandatory field in ATSAP where the reporter may indicate true or false if the airspace and associated procedure was complex to the point of being a factor during the incident.

Table 34 presents the single variable logit of airspace and procedure complexity on severity. While there is no statistical relationship between airspace and procedure on severity for either type of facility, the odds ratio for both facility types implies a lower likelihood of a severe event if the complexity of the airspace and related procedure was a factor. Lacking statistical significance, no conclusions can be drawn on the relationship between this variable and severity without additional research.

Table 34 - Logit Estimate of Airspace Procedure Complexity Factor by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Airspace & Procedure Complexity Factor - Tower	0.726	0.119	2,931
Airspace & Procedure Complexity Factor - TRACON	0.864	0.114	6,322
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2.4. Communication Complexity Factor

The communication complexity factor variable is a non-mandatory field in ATSAP where the reporter may indicate true or false if communication complexity was influenced an incident. Figure 17 presents the distribution of this variable.

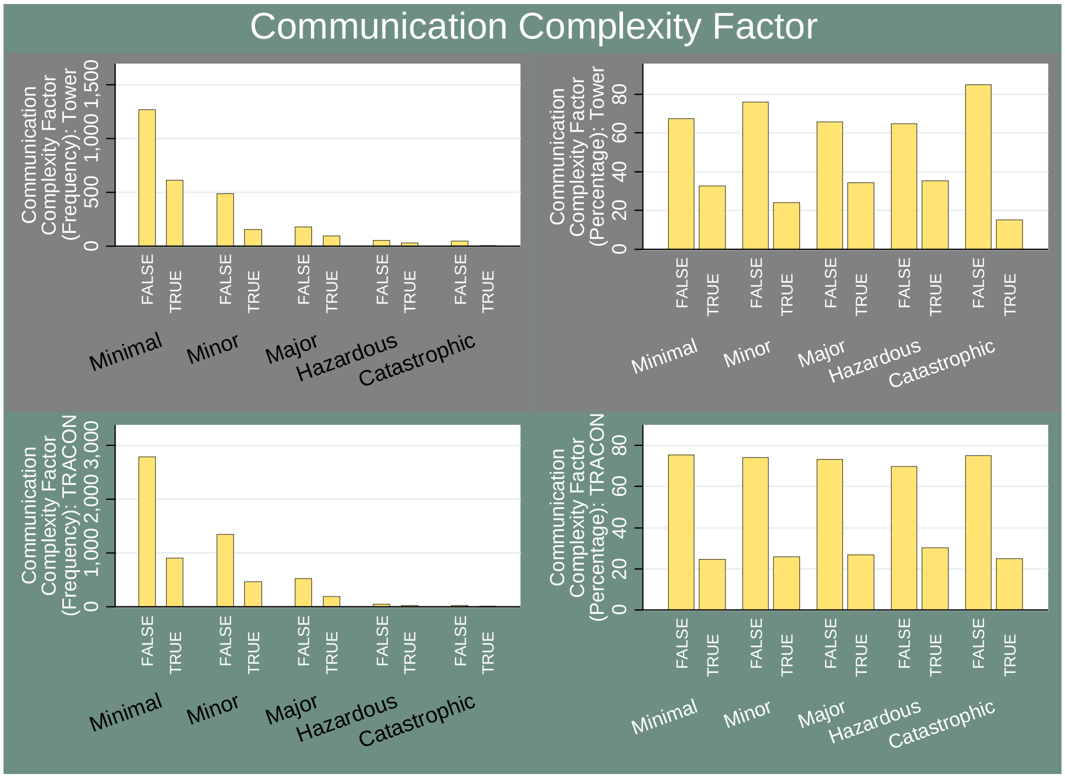

Figure 17 - Communication Complexity Factor

The results from a single variable logit of communication complexity factors on severe/non-severe events can be found in Table 35. Unlike coordination complexity factors, the odds ratios of communication complexity factors would suggest a slight increase in the odds of a severe incident occurring, however, the high p-values deter any statistical significance inference for both facility types.

Table 35 - Logit Estimate of Communication Complexity Factor by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Communication Complexity Factor - Tower	1.078	0.122	2,931
Communication Complexity Factor - TRACON	1.111	0.144	6,322
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2.5. Coordination Complexity Factor

Further research should examine the relationship between communication and coordination causal factors and their interaction with severity.

Coordination complexity factor variable is a non-mandatory field in ATSAP where the reporter may indicate true or false if coordination complexity was a factor during an incident. Theoretically speaking, this variable is to some degree related with the previous variable due to the fact that the any degree of coordination of air traffic would typically require communication. Although low correlation values (0.25 and 0.23 for tower and TRACON incidents, respectively) would indicate otherwise, it could still be worth investigating whether there is any interaction between the two variables in a final model.

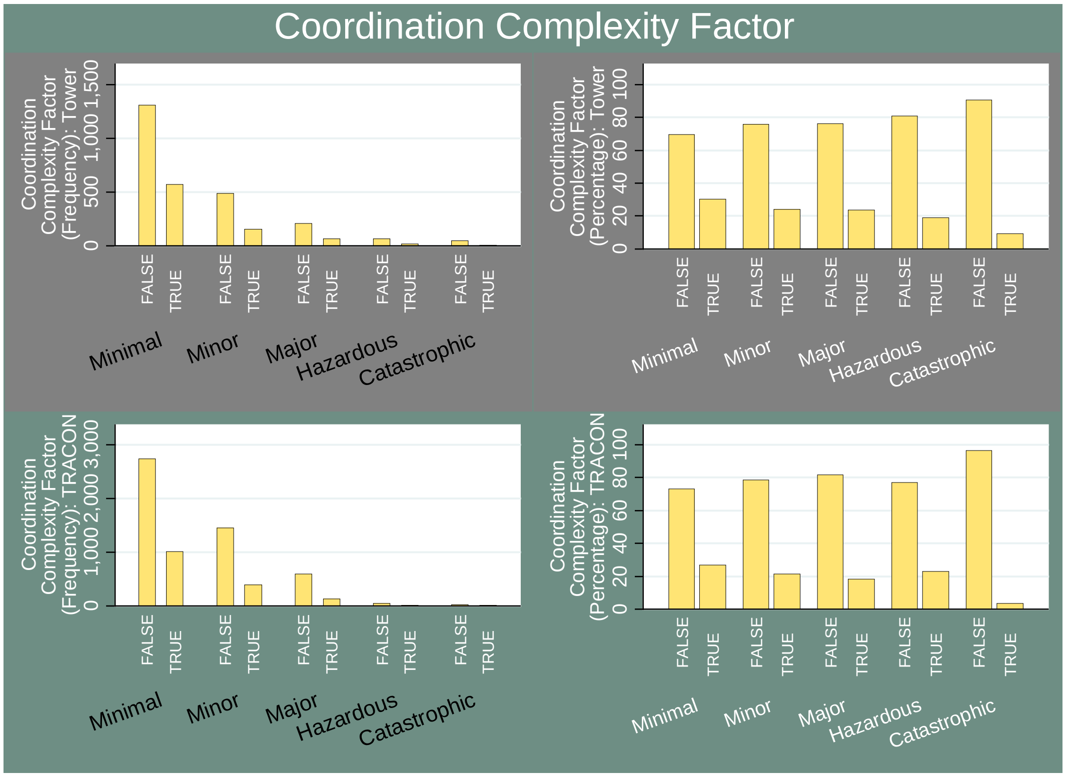

Figure 18 - Coordination Complexity Factor

Results in Table 36 from a single variable logit of coordination complexity on severe/non-severe incidents show a clear statistical relationship for both facility types. The odds ratios are nearly identical and both indicate a near 35% reduction in the probability of severe incident occurring if coordination complexity was marked as true. This may seem counterintuitive; however, it is possible that there is a relationship between coordination complexity and highly complex airspace that requires only the most experience or skilled controllers to manage, thus dampening the overall severity level if a loss of separation incident should occur. The final facility model will provide more insight for coordination complexity factors once all facility characteristics are controlled for.

Table 36 - Logit Estimate of Coordination Complexity Factor by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Coordination Complexity Factor - Tower	0.656***	0.098	2,931
Coordination Complexity Factor - TRACON	0.658***	0.078	6,427
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2.6. Facility Influences

The facility influences causal factor variable is an aggregation of variables that describe the culture, procedures, and processes that impact an individual facility's operational effectiveness negatively. This aggregated level variable reflects issues with controller pairing and teamwork, information flow issues where communication of information from leadership breaks down, and staffing shortages.

Table 38 presents the logit estimation output from a single variable logit of facility influences on severity. There is no statistical significance for both facility types.

Table 37 - Logit Estimate of Facility Influences by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Facility Influences - Tower	0.996	0.258	3,693
Facility Influences- TRACON	0.999	0.193	7,711
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2.7. Organizational Influences

The organizational influences variable is an aggregation of all the causal factors that fall within the organizational influences section. The organizational influences causal factor section refers to the cultural, procedural, and processes that impacted the organization's operational effectiveness, which also includes interfacing with other facilities/organizations. In theory, a disorderly organizational structure at a facility could have real, negative impacts on safety, translating to possibly more severe loss of separation incidents.

Results from a single variable logit of organizational influences on severe/non-severe events can be found in Table 38. These results indicate that organizational influences do not influence whether the loss of separation incident is severe or not for both types of facilities given the high p-value.

Table 38 - Logit Estimate of Organizational Influences by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Organizational Influences - Tower	1.092	0.308	3,692
Organizational Influences - TRACON	0.970	0.195	7,710
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2.8. Policy/Procedure Influences

The policy/procedure influences variable is an aggregation of causal factors that relate to policy or procedures that affected an individual group’s operational environment negatively. This includes such factors as inadequate, outdated, or lack of policy and/or procedures set in place.

The estimation output from a single variable logit of policy/procedure influences on severity is found in Table 39. Following the general trend for causal factors, there is not statistical relationship between policy/procedure influences and severity for either facility type.

Table 39 - Logit Estimate of Coordination Complexity Factor by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Policy/Procedure Influences - Tower	0.694	0.171	3,693
Policy/Procedure Influences- TRACON	1.068	0.178	7,710
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2.9. Staffing Configuration

The staffing configuration causal factor refers to staffing situation where sectors/positions are combined or decombined in such a way that it influences a loss of separation incident. Table 40 presents the single variable logit estimation output of staffing configuration on severity for each facility type. There appears to be no statistical relationship between staffing configuration and severe/non-severe incidents for either facility type.

Table 40 - Logit Estimate of Coordination Complexity Factor by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Staffing Configuration - Tower	0.852	0.140	6,932
Staffing Configuration Influences- TRACON	1.016	0.13	13,859
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2.10. Supervisory Influences

Supervisory influences are an aggregated group of causal factors relating to the leadership's effect on the individual or operation that contributed to a loss of separation outcome. This variable serves an indicator of the failure of leadership and management, including such things as intentional rule and regulation breaking or failure to hold realistic expectations of controllers' capabilities.

Table 41 presents estimation output from a single variable logit of supervisory influences on severity. As with the many other causal factors so far, there is no statistical relationship between severity and supervisory influences for both facility types.

Table 41 - Logit Estimate of Supervisory Influences by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Supervisory Influences - Tower	0.967	0.171	3,693
Supervisory Influences- TRACON	1.041	0.128	7,710
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2.11. Traffic Complexity Factor

Traffic complexity factor is a non-mandatory field in ATSAP where the reporter may indicate true or false if the traffic complexity was a factor during the incident. This variable could be viewed as a simplified version of the traffic rating variable and Table 42 and Table 43 highlight these variables relationship by facility type. In both cases the p-value from the Pearson Chi-Squared test indicates that there is a relationship between these two variables. This warrants careful consideration when including these variables in final mode to avoid any issues with multicollinearity. Figure 19 presents the distribution of the traffic complexity factor variable over severity.

Table 42 - Traffic Rating and Traffic Complexity Factor for Tower Facilities

	Traffic Complexity Rating					
Traffic Complexity Factor	1	2	3	4	5	Total
FALSE	908	612	722	254	31	2,527
TRUE	6	14	86	246	114	466
Total	914	626	808	500	145	2,993
P-value: 0.00						

Table 43 - Traffic Rating and Traffic Complexity Factor for TRACON Facilities

	Traffic Complexity Rating					
Traffic Complexity Factor	1	2	3	4	5	Total
FALSE	1,466	1,121	1,855	638	111	5,221
TRUE	6	18	202	636	406	1,277
Total	1,472	1,139	2,096	1,274	517	6,498
P-value: 0.00						

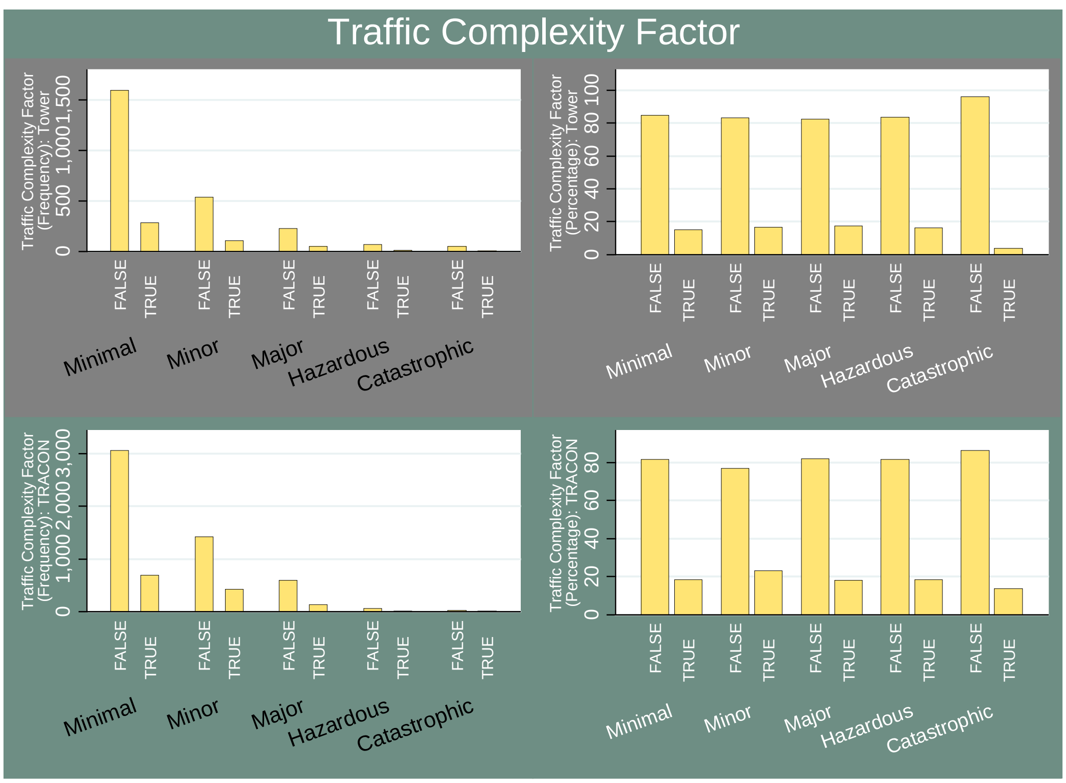

Figure 19 - Traffic Complexity Factor

The single variable logit estimation of traffic complexity factor on severity presented in Table 44 signals a lack of a statistically significant relationship between severe and non-severe events with the traffic complexity factor variable for both facility types. These results provide evidence in favor of using the traffic complexity rating variable over the traffic complexity factor variable in any further modeling.

Table 44 - Logit Estimate of Traffic Complexity Factor by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Traffic Complexity Factor - Tower	0.999	0.157	2,931
Traffic Complexity Factor - TRACON	0.881	0.116	6,427
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.2.12. Traffic Volume Complexity Factor

The traffic volume complexity factor variable is a non-mandatory field in ATSAP where the reporter may indicate true or false if the traffic volume was complex to the point of influencing an incident. This variable also has the distinction of being a simplified version of the traffic volume rating variable and Table 45 and Table 46 show the tabulation between the two variables by facility type. While there are a few 'true' traffic volume complexity records for lower rated traffic volume (ratings 1 and 2 in particular), the vast majority of the 'true' cases appear in ratings of 3 and higher. Additionally, the significant Chi-Squared test indicates that there is a relationship between the two variables for both facility types. Figure 20 reports the distribution of traffic volume complexity across severity.

Table 45 - Traffic Volume Rating and Traffic Volume Complexity Factor for Tower Facilities

	Traffic Volume Rating					
Traffic Volume Complexity	1	2	3	4	5	Total
FALSE	1,152	772	860	205	34	3,023
TRUE	11	30	116	273	115	545
Total	1,163	902	976	478	149	3,568
P-value: 0.00						

Table 46 - Traffic Volume Rating and Traffic Volume Complexity Factor for TRACON Facilities

	Traffic Volume Rating					
Traffic Volume Complexity	1	2	3	4	5	Total
FALSE	1,845	1,618	2,182	649	57	5,958
TRUE	15	43	229	632	342	1,261
Total	1,860	1,661	2,411	1,281	399	7,612
P-value: 0.00						

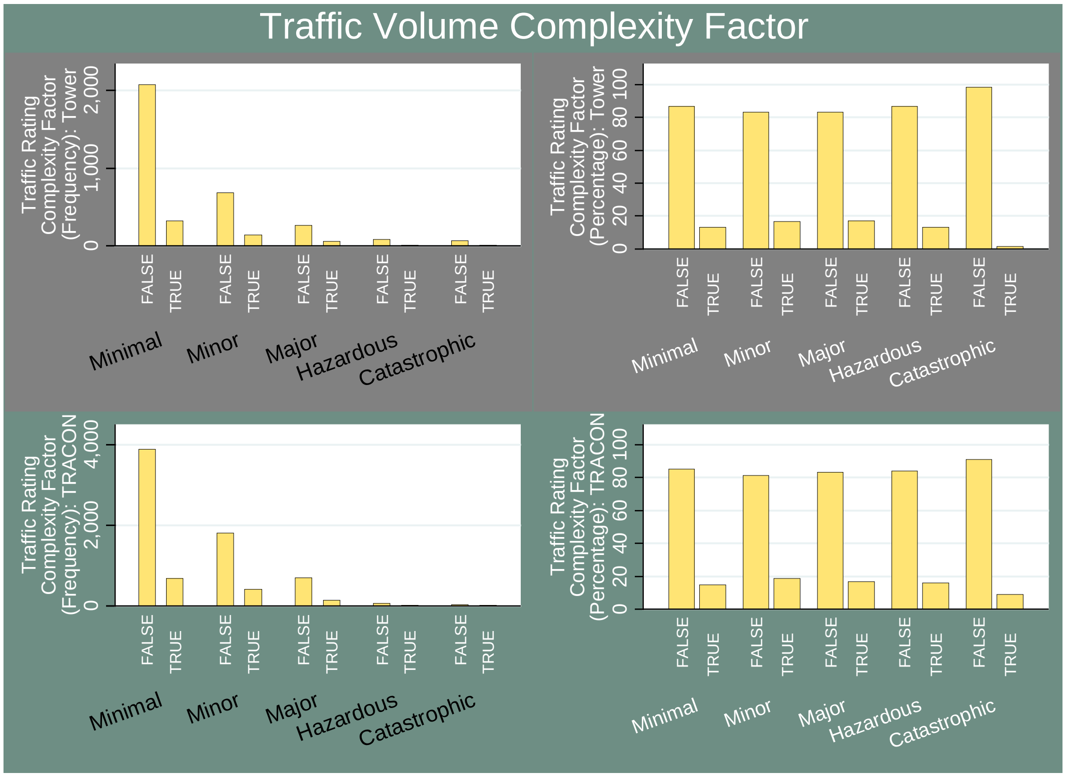

Figure 20 - Traffic Volume Complexity Factor

Table 48 presents the single variable logit results of traffic volume complexity on severity. For both facility types there is no statistical relationship between the traffic volume complexity factor and severe/non-severe events. These results shift the argument in favor of using traffic volume ratings over the traffic volume complexity factor variable in order to maintain accurate and interpretable results.

Table 47 - Logit Estimate of Traffic Volume Complexity Factor by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Traffic Volume Complexity Factor - Tower	.999	0.162	3,696
Traffic Volume Complexity Factor - TRACON	1.028	0.093	7,749
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.2.3. Facility Model

This section contains fully specified multinomial logit models containing facility characteristic variables examined in the previous section. Certain variables were excluded during this modeling exercise due to redundancy with other variables. Complexity factor variables for traffic volume rating and traffic complexity were not included due to both the statistically insignificant relationship with severity and the greater level of detail provided by the rating system variables. The variable traffic complexity rating was selected over traffic volume rating for two reason: data is already available on traffic volume in the form of daily operations, and traffic complexity rating, while highly correlated with traffic volume rating, is not exactly the same and understanding the relationship between traffic complexity and traffic volume at different severity levels are best served by analyzing only operations and traffic complexity.

All models include an interaction variable between coordination and communication complexity given the conceivable theoretical relationship between coordination and communication. Models also include year indicators. Since certain variables only span a portion of the time series, a second model was estimated with variables that span the entire dataset and are presented in Appendix D.5.

6.2.3.1. Tower Multinomial Logit Model

When aircraft/pilot action complexity is a factor, the likelihood of a more severe incident increases relative to the Minimal severity outcome. Communication complexity factor and traffic complexity rating both decrease the likely of a Catastrophic incident.

Table 48 presents the multinomial logit results of facility characteristics for tower incidents. With the lowest severity measure (minimal) being the base case, a coefficient value above 1 can be interpreted as increasing the probability of being in that severity classification with respect to the minimal severity outcome.

Table 48 - Multinomial Logit of Facility Characteristics for Towers

	Minor	Major	Hazardous	Catastrophic
Aircraft/Pilot Action Complexity Factor	1.2 (.11)	1.4* (.22)	2.5*** (.59)	2.2* (.72)
Airspace Procedure Complexity Factor	.85 (.12)	.93 (.17)	.66 (.23)	.2* (.13)
ATC Level Grade	1.1* (.055)	.99 (.07)	.98 (.12)	.78 (.14)
Communication Complexity Factor	.63** (.096)	.82 (.14)	1.1 (.34)	.23** (.13)
Configuration Contributed to the Event	1.2 (.33)	.42 (.2)	1.5 (.93)	1.1e-07*** (1.1e-07)
Coordination Complexity Factor	.76 (.11)	.43*** (.11)	.53 (.24)	.11* (.11)
Coordination X Communication	1.3 (.3)	2.7** (.88)	1.1 (.64)	15* (20)
Daily Operations	1 (.0014)	1 (.0024)	1 (.0046)	1 (.0075)
Facility Influences	.57 (.17)	1.1 (.38)	.86 (.55)	.25 (.3)
Organizational Influences	.81 (.25)	.77 (.33)	2.6 (1.4)	3 (2.3)
Policy/Procedure Influences	.66* (.13)	.67 (.2)	.4 (.2)	.56 (.38)
Staffing Configuration	1.9 (1)	4.3 (3.3)	1.6e-07*** (1.0e-07)	70*** (79)
Supervisory Influences	1.1 (.2)	1.1 (.26)	1.7 (.63)	.75 (.57)
Traffic Complexity Rating	1.1* (.045)	1.1* (.066)	1.2 (.13)	.58*** (.091)
Runway Count	.97 (.067)	.98 (.087)	.97 (.14)	.83 (.17)
Year 2011 Indicator	2.3* (.78)	1 (.38)	3.8 (3.9)	1.5 (.95)
Year 2012 Indicator	1.1 (.36)	.88 (.3)	2.3 (2.4)	.87 (.5)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 **N = 2,920**				

Firstly, it is interesting to note that daily operations and runway count are statistically insignificant in all severity levels and ATC level is only significant (and higher than one) in minor severity level cases. With that said, interpreting more than the sign of these coefficients in a multinomial model framework is unadvisable without first estimating their associated probability. Moreover, all coefficients must be interpreted in terms of changes in probability within their severity category, as opposed to across categories. Therefore, categorical variables of interest are expressed as the marginal and percentage changes in probability of severity categories in Figure 21 and Figure 22, while continuous variables are presented as their impact on the probability for each severity category in Figure 23 through Figure 28. For both set of figures, continuous variables that are not changing are held to their mean value, while categorical (causal factor) values that are not changing are set to zero.

When aircraft or pilot action complexity is a factor, the change in probability for a Minimal severity outcome is lowered while the Major, Hazardous, and Catastrophic outcomes are all increased. Moreover, the percentage change in probability for Hazardous and Catastrophic outcomes increases by over 100% each. This suggests that aircraft or pilot actions complexity factors may increase the likelihood of a more severe incident. The other three variables all appear to reduce the likelihood a severe incident, most notably for Catastrophic outcomes. When either communication complexity or coordination complexity are a factor, the percentage change in the likelihood of a Catastrophic incident decreases by nearly 100%. The percentage change for abnormal operational configuration was not possible to estimate due to zero Catastrophic observations for this variable.

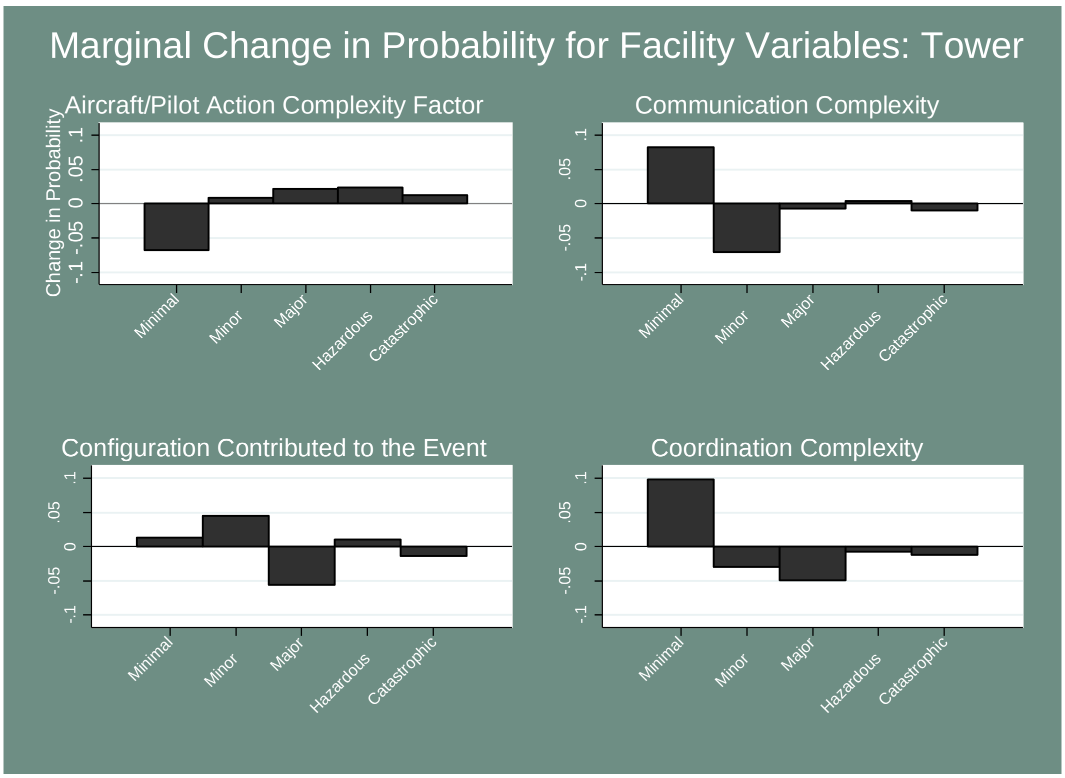

Figure 21 - Marginal Change in Probability for Facility Categorical Variables: Tower

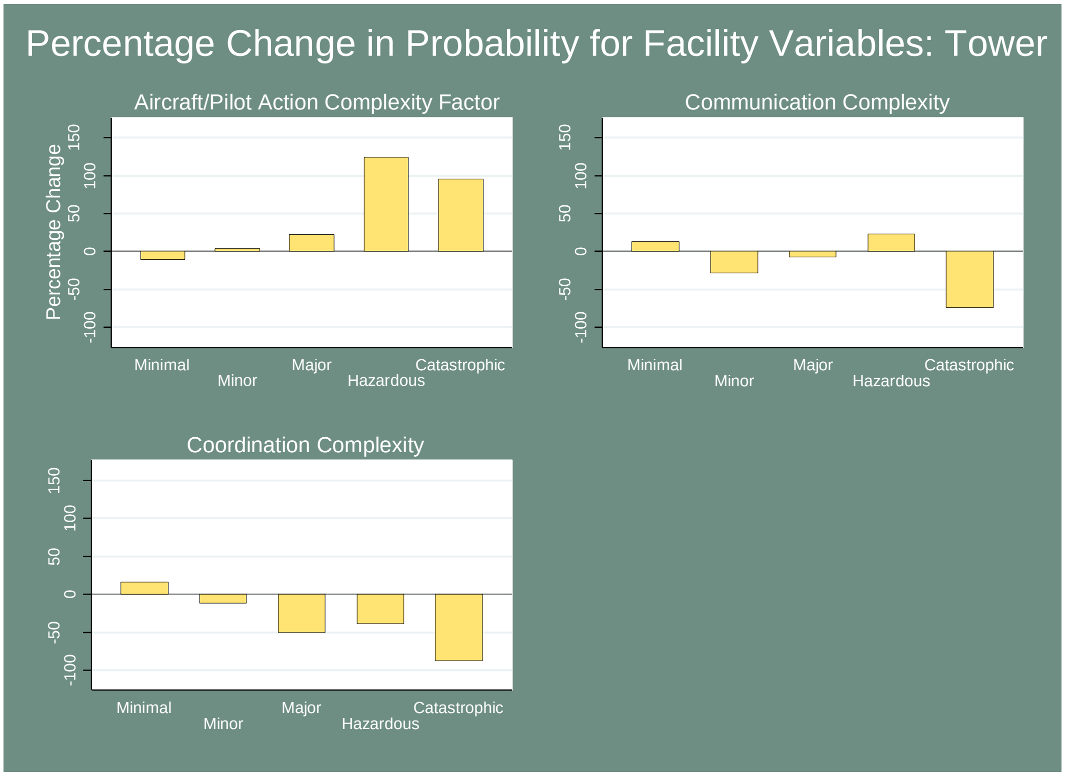

Figure 22 - Percentage Change in Probability for Facility Categorical Variables: Tower

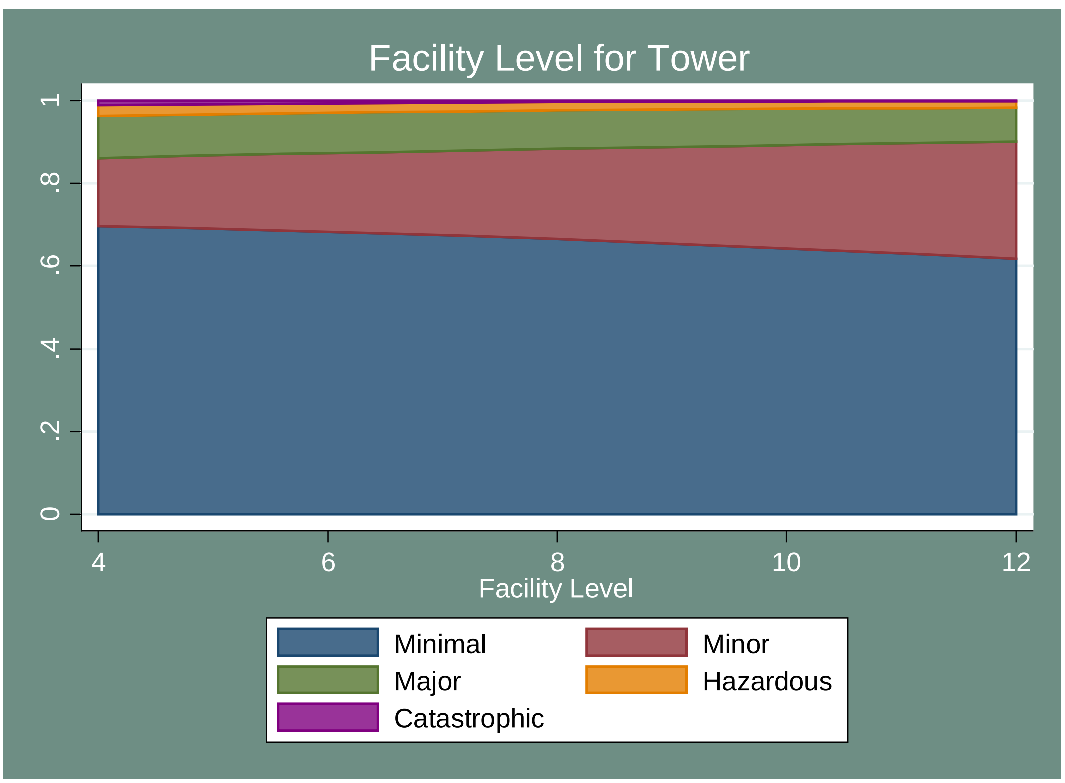

Figure 23 - Impact on Probability of Severity Categories of Facility Level for Tower Incidents

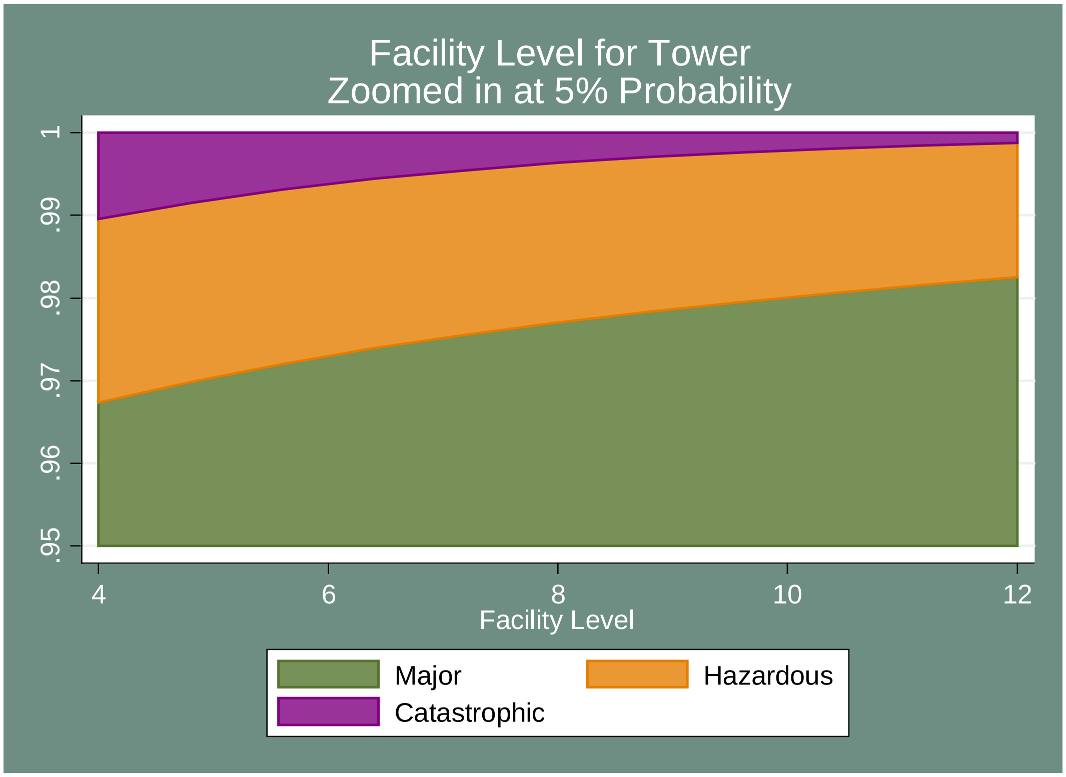

Figure 24 - Impact on Probability of Severity Categories of Facility Level for Tower Incidents, Zoomed In

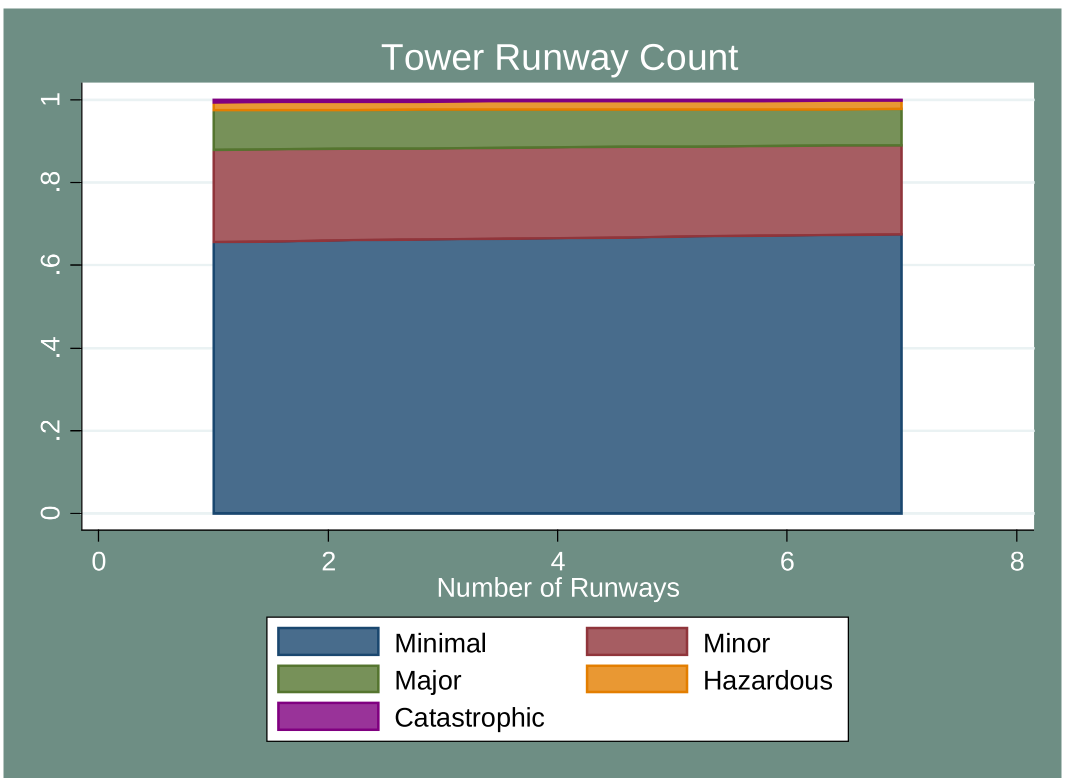

Figure 25 - Impact on Probability of Severity Categories of Runway Count for Tower Incident

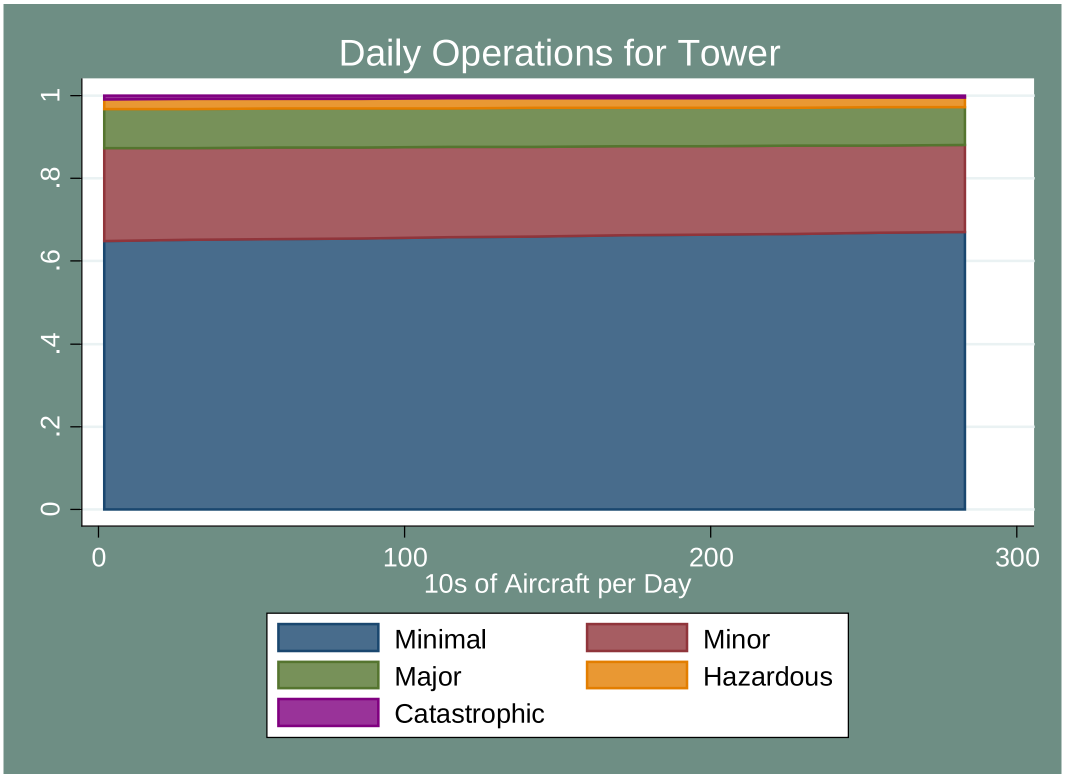

Figure 26 - Impact on Probability of Severity Categories of Daily Operations for Tower Incidents

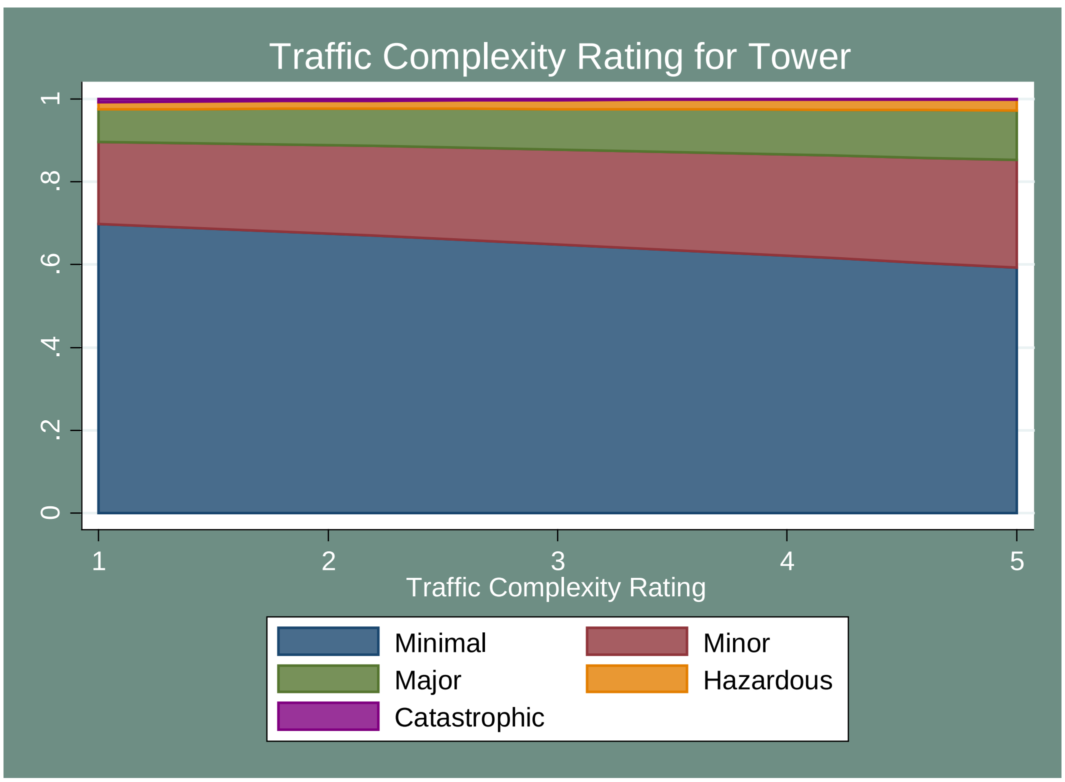

Figure 27 - Impact on Probability of Severity Categories of Traffic Complexity Rating for Tower Incidents

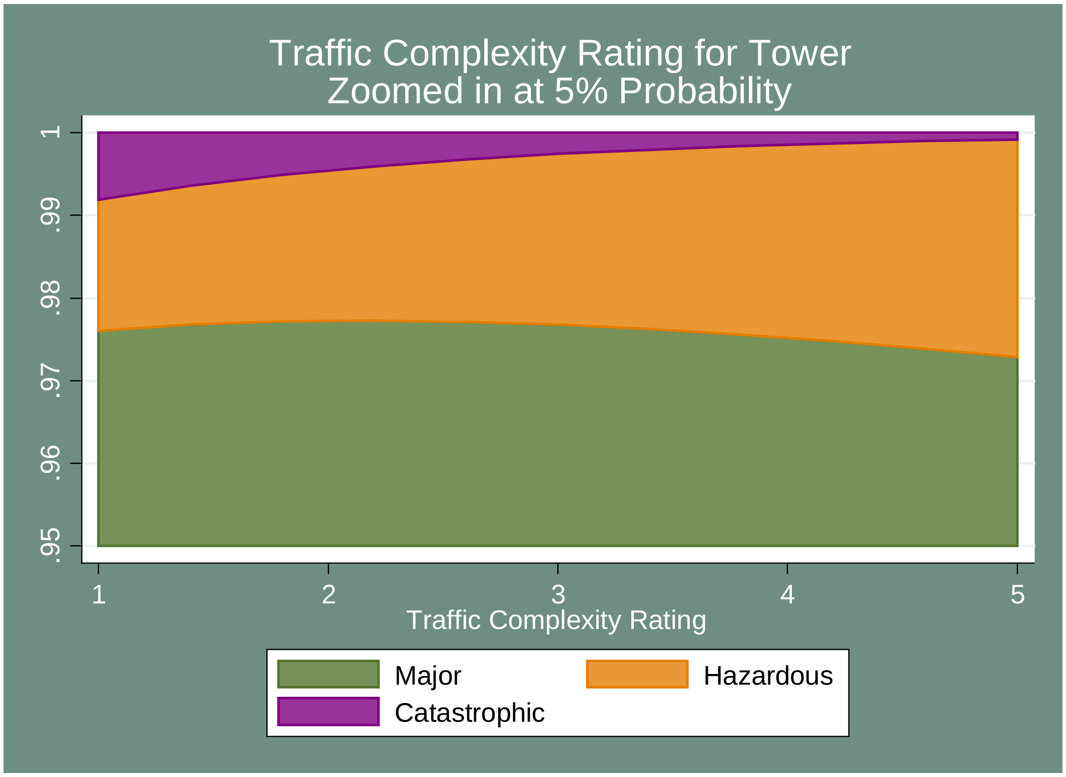

Figure 28 - Impact on Probability of Severity Categories of Traffic Complexity Rating for Tower Incidents, Zoomed

The variables daily operations and runway count are statistically insignificant for all severity categories. This is clearly evident in the figures presented above due to the relatively unchanged probabilities across the independent variable values. For Facility pay levels you see a significant increase in the probability of a minor event as the facility pay level increases. The traffic complexity rating variable indicates an increasing shift in the probabilities of minor, major and hazardous variables as the traffic complexity rating increases.

6.2.3.2. TRACON Multinomial Logit Model

Increases in operations have a decreasing effect on severity. When facility influences are a factor, the likelihood of a Catastrophic incident increases.

Results from the multinomial logit of facility characteristics for TRACON facilities are presented in Table 49. Figure 29 and Figure 30 present the change in probability for the categorical variables.

Table 49 - Multinomial Logit of Facility Characteristics for TRACONS

	Minor	Major	Hazardous	Catastrophic
Aircraft/Pilot Action Complexity Factor	1.2** (.082)	1 (.1)	.84 (.25)	.7 (.27)
Airspace Procedure Complexity Factor	.69*** (.061)	.78 (.12)	1 (.35)	.48 (.34)
ATC Level Grade	1.1* (.047)	1.2* (.097)	1.2 (.13)	1.1 (.15)
Communication Complexity Factor	1 (.089)	1.2 (.19)	1.5 (.44)	1.5 (.69)
Configuration Contributed to the Event	1.2 (.19)	1.7** (.35)	2.5 (1.8)	1.6e-06*** (6.4e-07)
Coordination Complexity Factor	.72*** (.067)	.53*** (.084)	.63 (.3)	2.4e-07*** (7.3e-08)
Coordination X Communication	1.1 (.19)	1.2 (.27)	1.3 (.82)	882007*** (1006215)
Daily Operations	1*** (.00087)	1* (.0013)	1 (.0019)	.99 (.0037)
Facility Influences	1.1 (.22)	1.2 (.27)	.3 (.34)	11*** (7.2)
Organizational Influences	.77 (.11)	.96 (.28)	.25 (.21)	8.9e-07*** (6.6e-07)
Policy/Procedure Influences	1 (.17)	1 (.18)	2 (.96)	5.9e-07*** (3.6e-07)
Staffing Configuration	.93 (.26)	.98 (.31)	.83 (1)	5.6e-06*** (3.8e-06)
Supervisory Influences	.99 (.12)	.91 (.18)	.9 (.44)	.53 (.47)
Traffic Complexity Rating	1.2*** (.033)	1.1*** (.036)	.99 (.07)	.67* (.12)
Runway Count	1.1*** (.012)	1** (.016)	1 (.025)	1.1* (.039)
Year 2011 Indicator	4.5*** (1.1)	2.1* (.78)	5.3 (5.6)	6642125*** (3718419)
Year 2012 Indicator	1.9** (.42)	1.1 (.36)	2.5 (2.6)	5897773*** (3264859)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 6,417				

The first interesting take away from the marginal change in probabilities figures for TRACON incidents is the lack of variation in the probabilities for more severe incidents (Hazardous and Catastrophic categories). This is most likely a statistical artifact for certain variables that lack observations in the Catastrophic category, pushing the change in probability for that category to zero. However, for categorical variables that are not lacking severe incident observations, the lack in variation could be due to the larger impact the continuous variables (daily operations, runway count, etc.) have on the TRACON facility model compared with the Tower model. In other words, controlling for these continuous variables dampens the overall effect of the categorical variables within severity categories.

The configuration contributed to the event variable appears to have the biggest impact on increasing the likelihood of a Major incident, while the coordination complexity factor variable reduced the likelihood of being in a severe category. The facility influences causal factor variable had the largest statistically relevant impact on Catastrophic incidents, with the percentage change increasing by close to 150%.

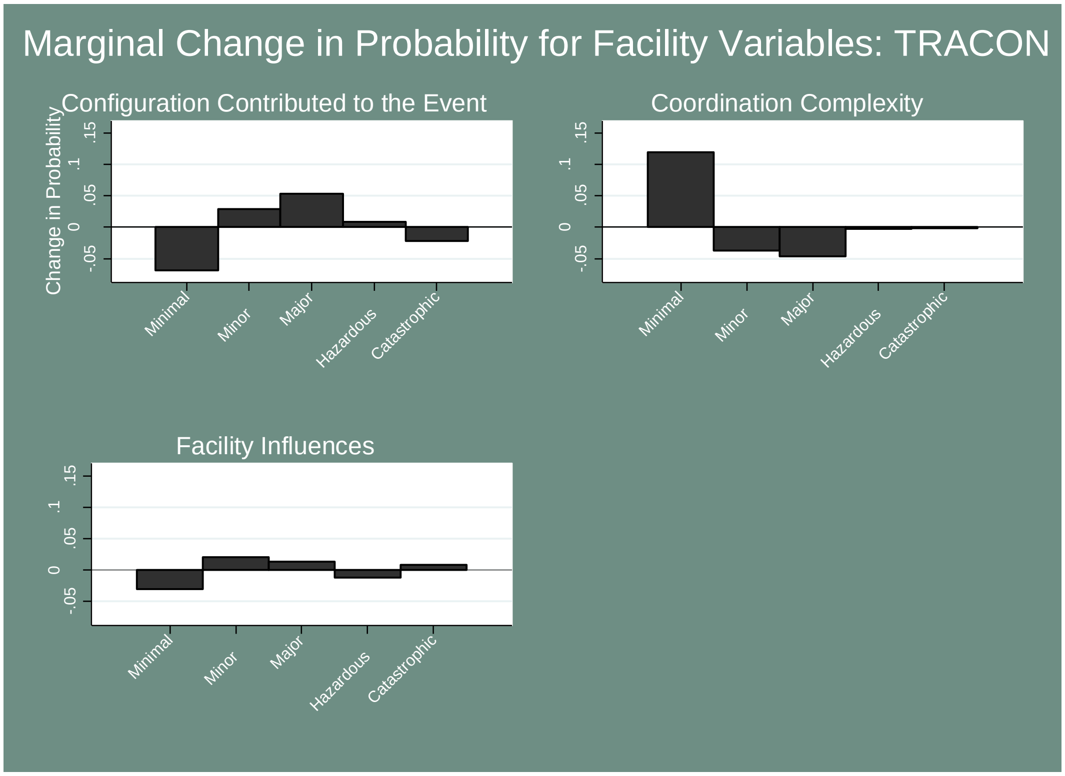

Figure 29 - Marginal Change in Probability for Facility Categorical Variables: TRACON

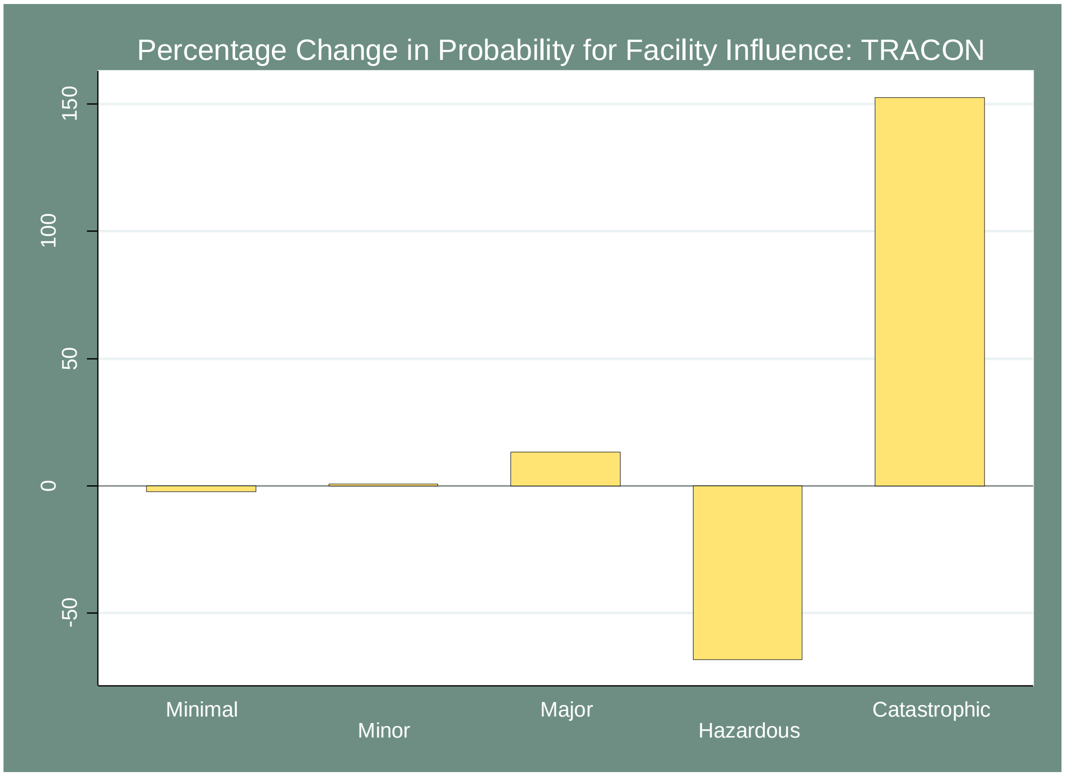

Figure 30 - Percentage Change in Probability for Facility Categorical Variables: TRACON

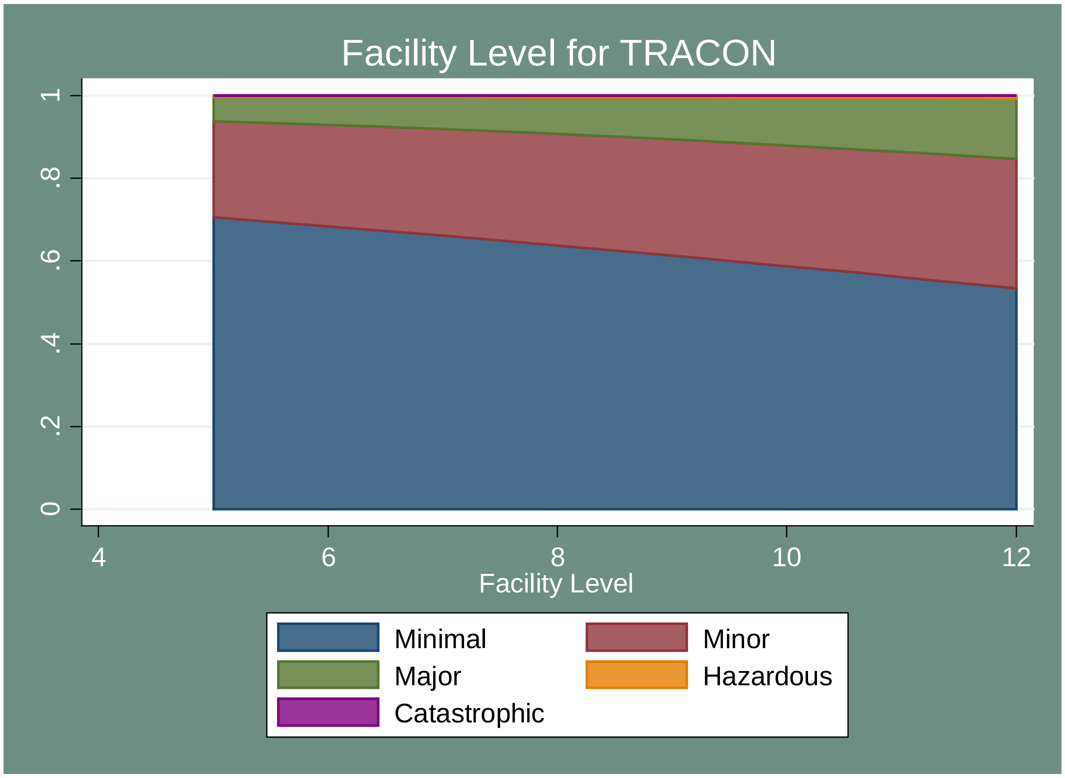

Figure 31 - Impact on Probability of Severity Categories of Facility Level for TRACON Incidents

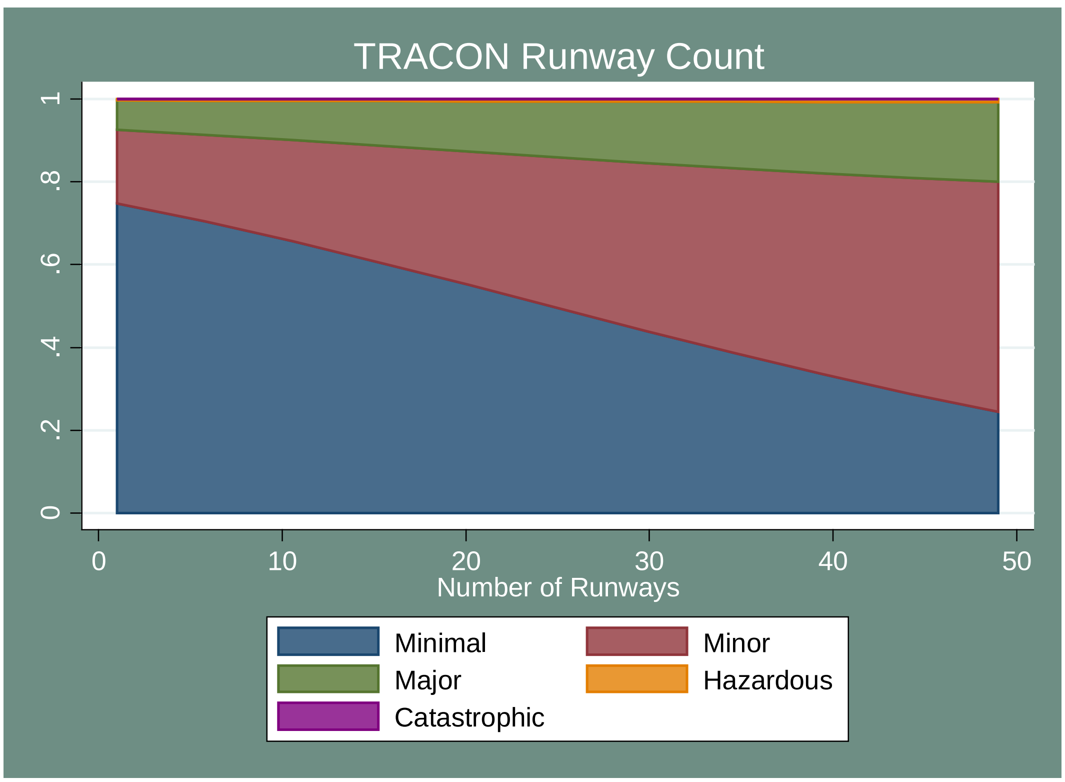

Figure 32 - Impact on Probability of Severity Categories of Runway Count for TRACON Incidents

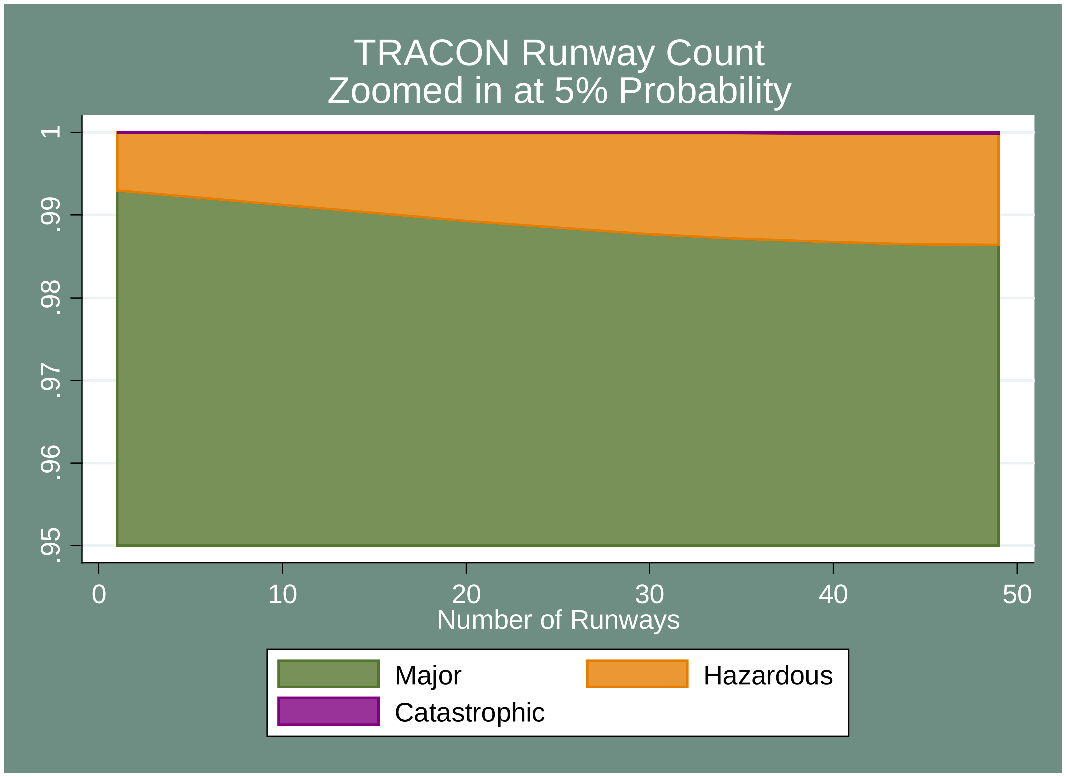

Figure 33 - Impact on Probability of Severity Categories of Runway Count for TRACON Incidents, Zoomed In

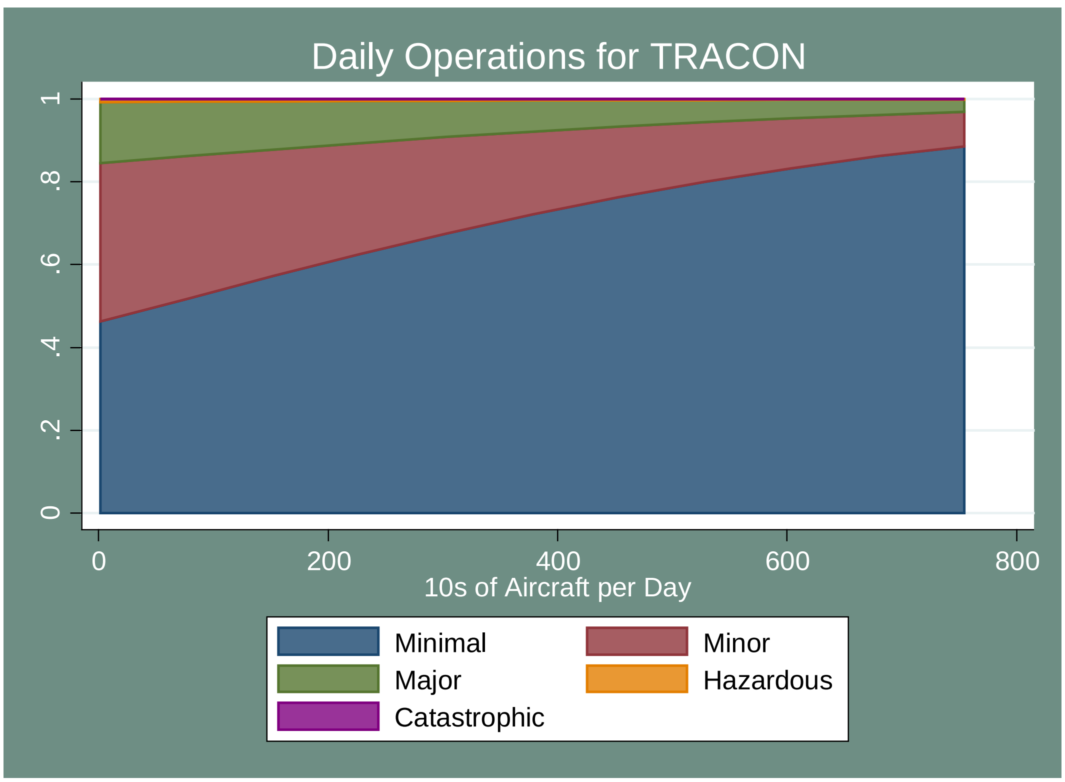

Figure 34 - Impact on Probability of Severity Categories of Daily Operations for TRACON Incidents

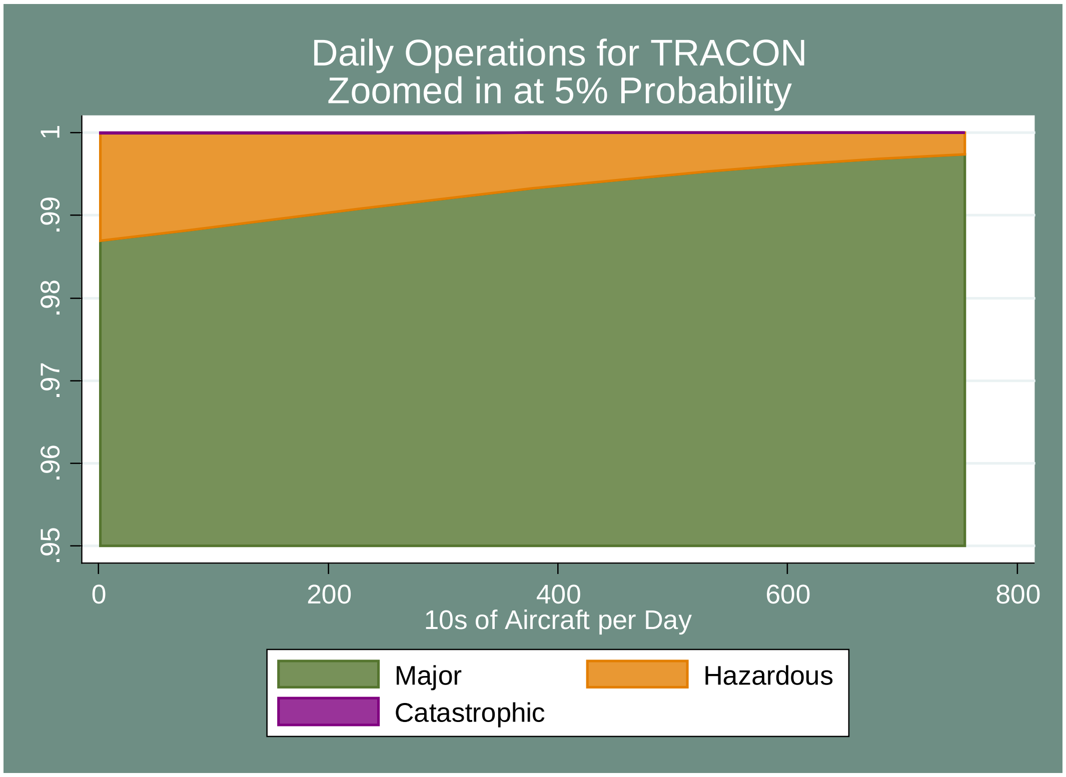

Figure 35 - Impact on Probability of Severity Categories of Daily Operations for TRACON Incidents, Zoomed

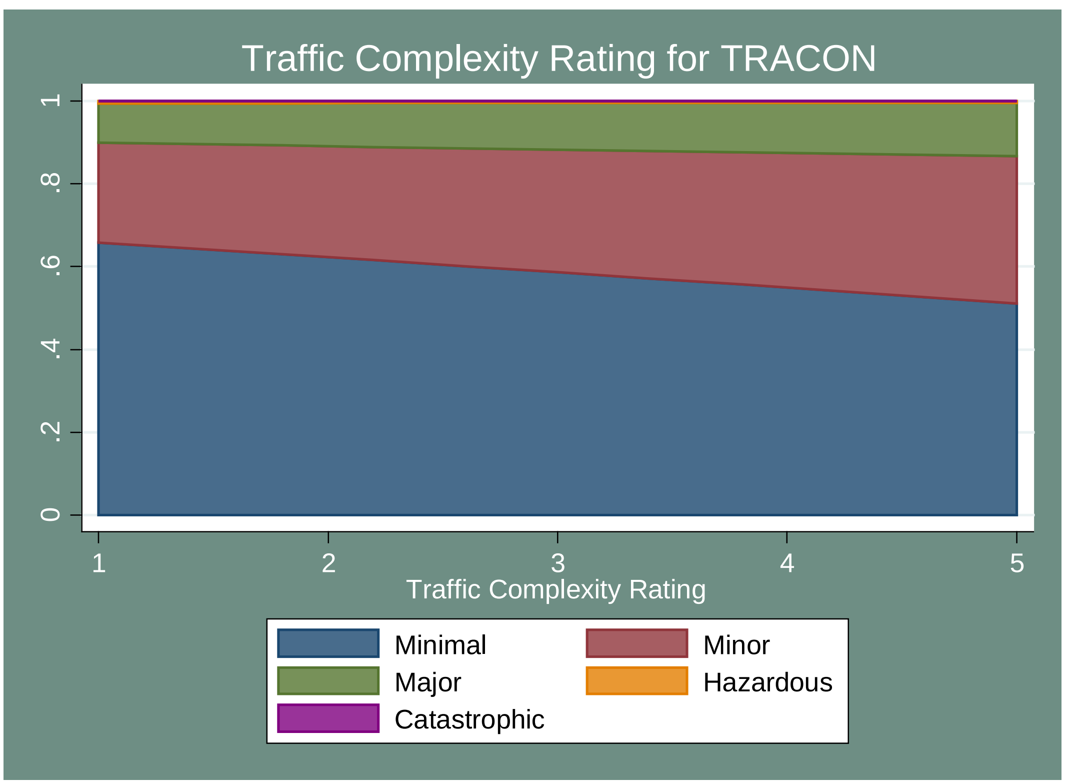

Figure 36 - Impact on Probability of Severity Categories of Traffic Complexity Rating for TRACON Incidents

Figure 31 through Figure 36 present the impact on probability of severity categories for the four different count variables of interest. With the exception of daily operations, an increase in the facility level, runway count, or traffic complexity ratings all increase the probability of a Minor or Major event. The daily operations variable has the opposite effect, where lower operations are associated with higher severity categories. This result seems counterintuitive, but controlling for traffic complexity should allow the relationship between traffic volume and severity to be more apparent. Since these variables are going in opposite directions as they increase, there is a mitigating effect between volume (daily operations) and complexity (runway count and facility level) at work in the TRACON airspace.

6.3. Controller Variables

Variables in the controller category primarily contain descriptors of the type of environment a controller is working in as well as controller actions that may be related to an airborne incident. Volpe is particularly interested to see if controller experience and/or position is tied to incident severity.

Some of the more interesting findings from this analysis are that navigation equipment failures are correlated with increased severity. Some modeling approaches found a small relationship between controller experience and lower severity, while others found no relationship. Regardless, the effect is too small to have any policy implications.

Controller variables are grouped into sub-categories, which are discussed in the following sections. The sub-categories include Approach Type, Experience, Capacity, Controller Actions, Controller Influences, Equipment Influences, Information Exchange, Training Issue, Unsafe Acts, and Work Area Influences. Data in each sub-category is analyzed separately; this information is then brought together in the full controller model. When variables, such as years of controller experience and years at a facility, measure similar concepts and are highly correlated; only one is used in the final model to avoid multicollinearity.

6.3.1. Controller Experience

In the binary model, there is a small but statistically significant reduction in Severe incidents with more experienced controllers in TRACON facilities, but this result do not hold in other model specifications. Thus, it must be viewed with caution.

Controller experience variables provide information on how long a controller has worked as a controller and at a particular facility. Variables on training issues and proficiency levels are also included. Volpe is particularly interested in seeing the relationship that controller experience has with severity.

After looking at summary statistics for both how long a controller has worked at as a controller and how long a controller has worked at a particular facility, Volpe decided to focus on how long a controller has been at a facility because it had a slightly stronger relationship with severity. Since controllers change facilities infrequently, the two variables are very similar.

The negligible relationship between controller experience and severity parallels the finding in the runway incursions report that any relationship between severity and experience is minor.

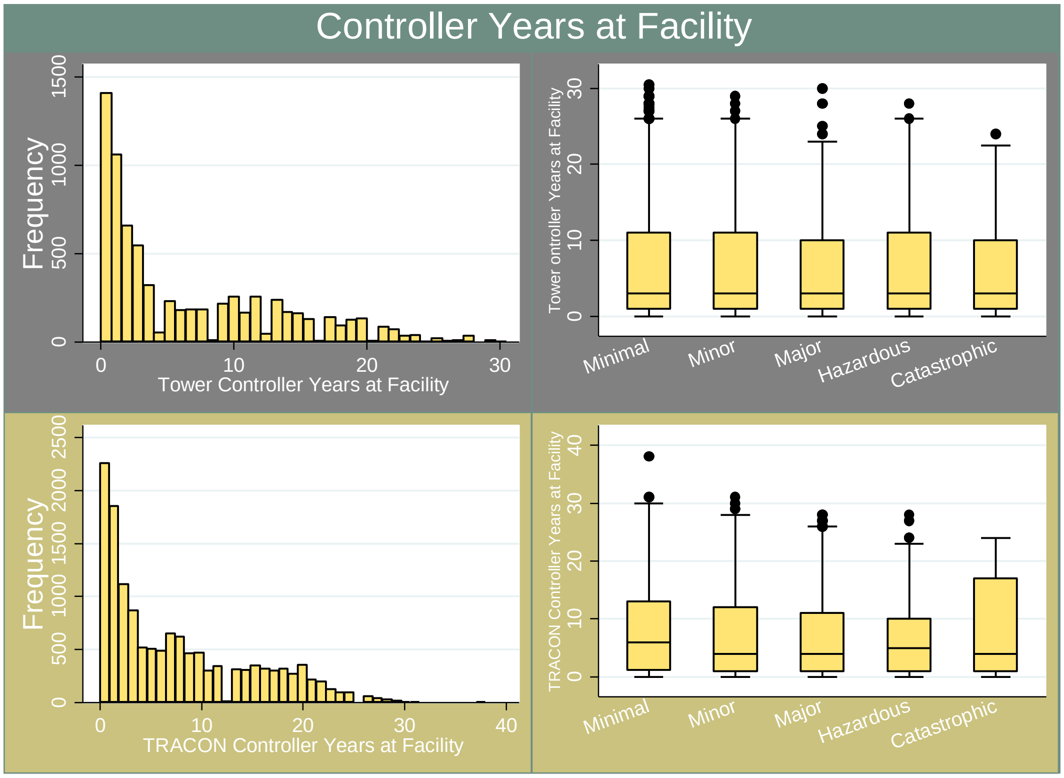

Figure 37 - Controller Experience and Severity

A binary analysis of the number of years a controller has been at a particular facility showed that there was a small, significant correlation between decreased severity and increased controller experience in TRACON facilities when only controller experience was analyzed. When other variables were held constant, as in the full controller model on page 88, there was no statistically significant relationship. There was also no statistically significant relationship in Tower facilities. The small magnitude of the differential, as well as the lack of robust results, makes this not likely to be useful for policy formulations.

Table 50 - Binary Controller Experience Model

Variable	Odds Ratio	Standard Error	Obs
Years at Facility - Tower	0.994	0.006	6,852
Years at Facility - TRACON	0.985*	0.006	13,431
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.3.2. Control Position

The Control Position variables refer to the type of control an aircraft was under when an incident occurred. Initially, 14 position types were analyzed, but Volpe found that many of the smaller position types had only a few observations and caused problems with multicollinearity. Thus, final modeling excludes Approach Control Coordinator, Approach Control Monitor, Approach Control Assistant, Approach Control Route Clearance Delivery, Traffic Management Coordinators / Supervisory Traffic Management Coordinators, Operations Manager, and Flight Data. This is because firstly, these variables have very few observations, and secondly, there did not appear to be other categories that were similar enough to merit combining these variables.

Binary logit models show that incidents with Ground Controller Position were unlikely to be severe. The data regarding ground controllers is likely skewed, however, because the ATSAP database only involves airborne incidents, with which ground controllers are unlikely to be involved. In TRACON facilities, incidents under Satellite Control had higher than expected severity levels, and incidents under FLM Control had lower than expected severity levels.

Table 51 - Control Position and Severity

Tower

Variable	Odds Ratio	Standard Error
Assistant	0.54*	0.15
Cab Coordinator	0.74	0.23
Clearance Delivery	0.56**	0.12
Flight Data	1.21	0.25
FLM/CIC	1.08	0.16
Gate Hold Metering	0.00***	0
Ground	0.59***	0.07
OM	0.00***	0
Other	1.06	0.25
Radar	1.03	0.32
TMC	0.57	0.55
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 N=6,914; Base = Local		

TRACON

Variable	Odds Ratio	Standard Error
Arrival	0.86	0.08
Departure	0.94	0.08
FLM	0.65*	0.12
Handoff	0.73	0.13
Satellite	1.42*	0.2
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 N = 13,535; Base = Final		

6.3.3. Capacity

The small sample size yielded no statistically significant results between Capacity and Severity Level.

The controller and equipment capacity variables denote if the controller entering an incident into ATSAP believed that either the controller or available equipment did not have the capacity to handle a given situation. There were only a small number of observations, and the sample size is too small to yield robust results.

6.3.4. Controller Actions

The small sample size prevented analysis of these variables from yielding robust results.

Controller actions describe when a controller did something, such as did not follow a procedure or made a poor choice that contributed to an incident. There are four categories:

- Plan poorly executed
- Information misinterpreted
- Policy or procedure not followed
- Inadequate plan

Since each individual category only had a small number of observations, they were aggregated together for analysis purposes. Furthermore, these variable need to be treated with care since they were entered into the database only when a controller determined that they were a cause of an incident. The resulting statistical issues of Reporting Bias and Measurement Error are described in Section 5.9.

The small number of observations means that there are few statistically significant results. A binary logit model showed no significant relationship between a controller action problem and incident severity distribution.

6.3.5. Controller Influences

Controller Influences are tied to decreased severity, but the possibility of Reporting Bias means that the results need to be treated with skepticism.

Controller influences describe the way in which a controller approached handling an incident, and are tied to a controller's job performance. Categories include complacency/boredom, reliance on automation, inconsistent with experience, lack of planning, and personality conflict. Lack of planning is named as a causal factor in the largest number of instances.

A binary analysis that looked at controller influences in isolation found a relationship between controller influences being marked in ATSAP and low severity incidents. Since these variables are subjective, they are highly sensitive to reporting bias, as described in the Volpe Methods section of this report; thus model results are difficult to interpret.

6.3.6. Equipment Influences

The small sample size prevented analysis of these variables from yielding robust results.

Equipment Influences identify problems with equipment that controllers report are associated with airborne incidents. There were only a small number of observations from these variables; thus modeling efforts did not yield statistically significant results.

6.3.7. Information Exchange

Information Exchanges were tied to lower than expected severity.

This variable category refers to if an Information Exchange was considered as having contributed to an incident. Volpe's analysis showed that incidents where an information exchange was listed as a contributing factor were less severe than expected; the results were statistically significant at the 1% level in TRACON facilities. A possible explanation is that if information is misinterpreted or there is another similar issue, the mistake typically corrected before an incident becomes severe.

Table 52 - Information Exchange and Severity

Variable	Odds Ratio	Standard Error	Obs
Info Exchange - Tower	0.768	0.135	3,692
Info Exchange - TRACON	0.721**	0.089	7,575
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.3.8. Training Issue

There is no statistically significant relationship between severity and Training Issues.

Training Issues indicate if either someone was being trained when an incident occurred or if a deficiency in a controller's training contributed to an incident. There is no statistically significant relationship between Training Issues and severity.

6.3.9. Unsafe Acts

Incidents with Unsafe Actions have an increased likelihood of Minimal severity in both Tower and TRACON facilities.

Unsafe Acts refer to specific errors made by controllers. They are categorized into decision, perceptual, skill-based, and violation errors. Perception and Skill-Based Errors had the largest number of incidents associated with them – although the overall frequency is low. These variables were aggregated together for analysis purposes, since there are only a small number of observations.

Multinomial models with other controller variables held constant showed a small but statistically significant relationship between incidents of Minimal severity and Unsafe Acts in both Tower and TRACON facilities. It is important to note that these variables are likely subjected to the issues of reporting bias and measurement error that are described in section 6.3.1, so any analysis of model results must keep this in mind.

6.3.10. Work Area Influences

Work Area Influences do not have a statistically significant relationship with severity.

Work area influence variables are causal factors that describe elements in a controller's work environment that are thought to contribute to an incident. Ambient noise is the most frequently reported variable, but even for this, there are only a small number of entries, making analysis difficult. Thus, these variables are aggregated together for analysis. Work Area Influences do not have a statistically significant relationship with severity.

6.3.11. Controller Model

The next sections put together the prior data explanations into a model that incorporates all of the relevant Controller variables. Since the Causal Factor variables only started being recorded in 2011, while the other variables are available from 2008, the models were run in two sets (with and without the Causal Factor variables). Multinomial models that include the causal factors (2011 onwards) are shown in this section; results from the models with the full date range, as well as binary models, are available in the appendix.

6.3.11.1. Tower Multinomial Logit Model

Controller Experience has no relationship with severity in Tower Facilities.

In Tower facilities, the number of years a controller had worked at a facility was not correlated with incident severity. This is a surprising finding, since Volpe’s expectation had been that more experienced controllers would have better safety records, but it parallels findings found in the Runway Incursions report.

Assistant Control Status was associated with a higher than expected frequency of catastrophic severity levels, while conversely, ground controller positions were associated with minimal severity incidents. Since the ATSAP database is focused on airborne incidents, the dataset involving ground controller positions is likely a skewed sample.

The causal factor model version showed that unsafe actions were associated with Minor severity incidents. Information exchange, training issues, and work area incidents had no reported catastrophic incidents; but the relatively small number of reported observations means that this is likely related to the low overall frequency of catastrophic incidents.

Table 53 - Tower Multinomial Controller Model (2011 - 2013)

	Minor	Major	Hazardous	Catastrophic
Other	.65 (.25)	1.3 (.51)	.5 (.38)	2.5 (2.2)
Cab Coordinator	1.1 (.4)	.54 (.33)	1.3e-07*** (3.9e-08)	1.1 (.85)
Flight Data	1.3 (.28)	.85 (.31)	1.5 (1.1)	.84 (.62)
FLM/CIC	1.5** (.21)	.89 (.22)	1 (.41)	1.1 (.45)
Clearance Delivery	.41*** (.094)	.38* (.15)	.097 (.12)	.5 (.38)
Ground	.57*** (.079)	.49*** (.1)	.35** (.14)	.72 (.28)
Assistant	1.8** (.37)	.59 (.3)	.53 (.55)	5.8** (3.9)
Radar	.41* (.16)	1.7 (.79)	.58 (.63)	8.5e-08*** (5.3e-08)
TMC	3.1 (1.9)	2.9 (3.5)	3.7e-07*** (2.9e-07)	5.1e-07*** (4.1e-07)
OM	2.2e-08*** (2.3e-08)	2.3e-08*** (2.5e-08)	1.1e-08*** (1.4e-08)	369883*** (615506)
Gate Hold Metering	6.3e-08*** (3.6e-08)	1.3e-07*** (8.0e-08)	2.4e-07*** (1.8e-07)	3.3e-07*** (3.9e-07)
Years at Facility	1 (.0068)	.99 (.0091)	1 (.018)	.98 (.019)
Capacity	.63 (.25)	.75 (.37)	1.5 (1.2)	2.8 (2.9)
Controller Actions	1.4** (.16)	1.5** (.2)	.91 (.21)	2.0e-07*** (3.9e-08)
Controller Influences	.8 (.11)	1.1 (.21)	.79 (.24)	.52 (.38)
Equipment Design Problem	1 (.7)	.69 (.71)	2.6 (3.1)	2.5e-06*** (1.9e-06)
Equipment Malfunction	.44 (.19)	1.2 (.49)	1.1 (.92)	.84 (.96)
Information Exchange	.68** (.096)	.8 (.16)	.59 (.21)	1.9e-07*** (4.0e-08)
Training Issue	1.3 (.19)	1.5* (.29)	1.3 (.43)	2.9e-07*** (7.1e-08)
Unsafe Actions	1.7*** (.2)	.95 (.17)	1.4 (.43)	.24 (.24)
Work Area Influences	.79 (.16)	.97 (.28)	1.3 (.52)	2.7e-07*** (9.2e-08)
Total Operations	1*** (.00068)	1 (.0011)	1 (.0016)	.98*** (.0047)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N=3,680				

6.3.11.2. TRACON Multinomial Logit Model

Controller Experience has no significant relationship with severity in TRACON facilities, when the variables introduced in 2011 are controlled for. Satellite Control Status is associated with a higher incidence of Catastrophic severity levels.

In the TRACON model for 2011 onwards, controller experience had no statistically significant association with severity levels; this was mirrored in the Tower model. This indicates that controllers with more experience are neither more nor less likely to be involved in severe incidents. Keep in mind that this analysis only looks at the severity of incidents that have occurred; it does not look at the probability that an incident will occur in the first place. It is possible that there is a relationship between controller experience and the likelihood of an incident occurring.

There was a small but statistically significant relationship between controller experience and a lower incidence of minor and major severity levels in the model specification that included the full time series but not the variables introduced in 2011; this model is shown in Appendix E.5. Since this result does not appear consistently with different model specifications, it must be viewed with caution.

A multinomial logit model showed that incidents with satellite control positions have a noticeably higher frequency of catastrophic severity ranks, with an odds ratio of 3.6. This result merits further research to determine possible reasons for this relationship.

The addition of the causal factor variables into the model showed that in TRACON facilities, information exchanges were associated with low severity incidents. It may be that errors involving information exchanges are typically resolved rapidly before incidents become severe.

Table 54 - TRACON Multinomial Controller Model (2011-2013)

	Minor	Major	Hazardous	Catastrophic
Arrival	.94 (.095)	.86 (.12)	.65 (.15)	.98 (.28)
Departure	.69*** (.041)	.86 (.11)	.85 (.21)	1.1 (.36)
Final	1.3*** (.1)	1 (.15)	1.1 (.35)	.95 (.34)
FLM	.81 (.15)	.73 (.2)	5.7e-08*** (1.3e-08)	4.2* (2.4)
Handoff	.94 (.15)	.7 (.15)	.99 (.54)	1 (.85)
Satellite	.9 (.079)	1.5* (.24)	1.3 (.32)	3.6*** (1.1)
Years at Facility	.99 (.0054)	.99 (.0097)	.99 (.018)	1 (.019)
Capacity_All	1.5 (.4)	.58 (.28)	3 (2.3)	4.5e-07*** (1.7e-07)
Controller Actions	1.3*** (.081)	1.1 (.12)	.85 (.22)	.32 (.2)
Control Influences	.88 (.083)	.68** (.094)	1.3 (.46)	.4 (.4)
Equipment Design Problem	1.2 (.33)	1.1 (.52)	2.4 (2.5)	5.8e-07*** (2.1e-07)
Equipment Malfunction	.82 (.19)	.59 (.2)	.81 (.84)	2.7e-07*** (9.0e-08)
Information Exchange	.73*** (.062)	.74* (.097)	.51 (.22)	3.0e-07*** (7.1e-08)
Training Issue	1.1 (.097)	1 (.14)	.58 (.32)	5.8e-07*** (1.3e-07)
Unsafe Actions	2.1*** (.2)	1.4* (.21)	1.8 (.54)	5.2e-07*** (1.3e-07)
Work Area Influences	.83 (.12)	1.2 (.25)	.33 (.33)	4.3e-07*** (1.2e-07)
Total Operations	1*** (.00044)	1*** (.00062)	1* (.0007)	1 (.001)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 7,566				

6.4. ATC/Pilot Communications/Clearance

Variables in the ATC/Pilot Communications/Clearance category describe communication issues between ATC and pilots. Analyzing these will allow us to determine if there is a relationship between incident severity and ATC/pilot communications. Where necessary, similar variables are aggregated together when there are not enough observations to yield robust results with an independent analysis of the variables.

6.4.1. Loss of Communication

These variables describe why a pilot lost communication with a controller. Looked at individually, most have too few observations to yield robust statistical results; the number of observations ranges from 4 (Stuck Mike) to 382 (Other Common Issue). Especially given that the most frequent occurrence is “other”, Volpe decided to aggregate these into a single Loss of Communication Variable. *Table 55 shows that a loss of communication incident is associated with a higher than expected severity distribution.*

Table 55 - Loss of Communication Actual and Expected

	Actual	Expected
Minimal	580	603
Minor	227	269
Major	148	99
Hazardous	17	16
Catastrophic	24	9

6.4.2. Readback Problem

Readback problems refer to specific instances where a pilot or controller does not repeat a direction or position to confirm that communicated information has been understood correctly. There are 8 separate types of readback problems in the database, ranging from 10 observations (speed) to 673 observations (clearance). These variables are combined because the only variable with a large enough number of observations to be analyzed singly does not have a statistically significant relationship with severity. The multinomial logit analysis, shown in Table 63 below, shows that readback problems are associated with a slightly lower frequency of catastrophic incidents.

6.4.3. Acknowledgement Problem

Acknowledgement problems refer to when a call sign is not used to identify an aircraft, an acknowledgement is not received, or when the wrong aircraft is acknowledged. While these variables were combined for analysis*, it is worth noting that the “acknowledgement not received” category had a high frequency of catastrophic incidents (6 catastrophic out of 286 incidents).* A multinomial logit analysis, displayed in Table 64, showed that acknowledgement problems had a higher than expected frequency of Major incidents.

6.4.4. Clearance Problem

The 16 Clearance Problem variables refer to the permission given a pilot to go to a specific place or perform a maneuver. The number of observations range from 61 to 2,553 (altitude). The larger variables were analyzed independently; Table 56 shows that routing-clearance problems are associated with lower severity.

Table 56 - Clearance Problem: Routing

Variable	Odds Ratio	Standard Error	P-Value	95% CI LB	95% CI UB	Obs
Routing	0.648	0.083	0.001	.5034043	.8336606	735

An aggregated Clearance Problem variable was created so that the variables with few observations would be included in the regressions. On average, as shown in Table 57, incidents with a clearance problem have a lower than expected severity distribution.

Table 57 - Actual and Expected Severity of Clearance Problems

	Actual	*Expected*
Minimal	3127	*3409*
Minor	1857	*1512*
Major	612	*612*
Hazardous	77	*93*
Catastrophic	8	*55*

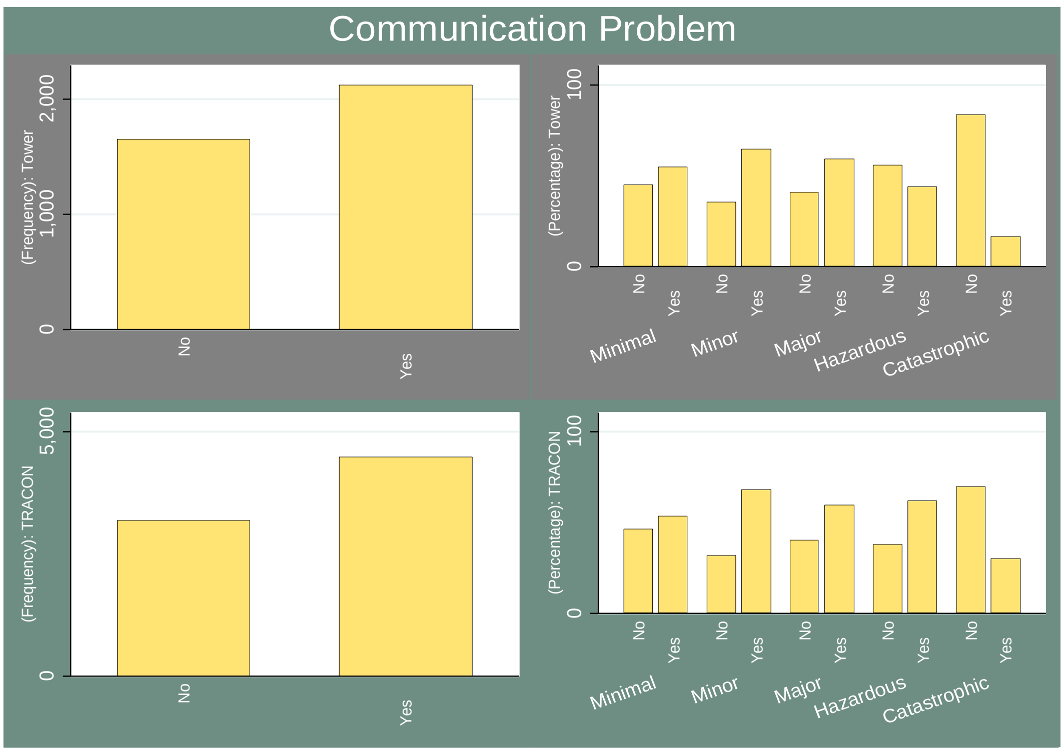

Figure 38 - Communication Problems

Looking at all of the variables associated with communication problems together, Clearance problems have the largest number of associated incidents in the ATSAP database, followed by readback problems. This can be seen in Table 58. This makes sense, since a problem with a clearance is likely to directly lead to an incident since it means that a plane is going where it has not been directed to go.

Table 58 - Communication Problem Frequency

	ATSAP Incidents
Clearance	5761
Readback/Hearback	1658
Loss of Communication	1016
Aircraft Acknowledgement	441
Phraseology	428

6.4.5. Computer Entry Problem

Computer Entry problem variables refer to typo-style mistakes in how data are entered into a system. There are a very small number of incidents in each category; the number of observations ranges from 10 (premature termination of data) to 41 (update entry). No "yes" observations are Hazardous or Catastrophic; only one is major. Even when aggregated together, the robustness of any results from Computer Entry problems must be questioned. In this case, the Tower and TRACON facilities are examined together so that the model can draw on more data. *Logit results in Table 59 show that Computer Entry problems are associated with Low Severity incidents.*

Table 59 - Binary Logit Computer Entry Problem

Variable	Odds Ratio	Standard Error	P-Value	95% CI LB	95% CI UB	Obs
Computer Entry Problem	0.069	0.065	0.01	0.011	0.441	41

6.4.6. Displayed Data Problem

Displayed data problems refer to incidents where the controller is not able to correctly view needed information about the plane(s) being controlled. There are 10 different categories that are aggregated together for analysis, since each individual category has only a small number of "yes" observations (ranging from 0 to 96). Tower and TRACON spaces were analyzed together. While a binary model (Table 60) showed a lower severity distribution with Displayed Data problems, the results were not statistically significant. Multinomial results were also insignificant.

Table 60 - Binary Logit Displayed Data Problem

Variable	Odds Ratio	Standard Error	P-Value	95% CI LB	95% CI UB	Obs
Displayed Data Problem	0.703	0.154	0.11	0.458	1.080	276

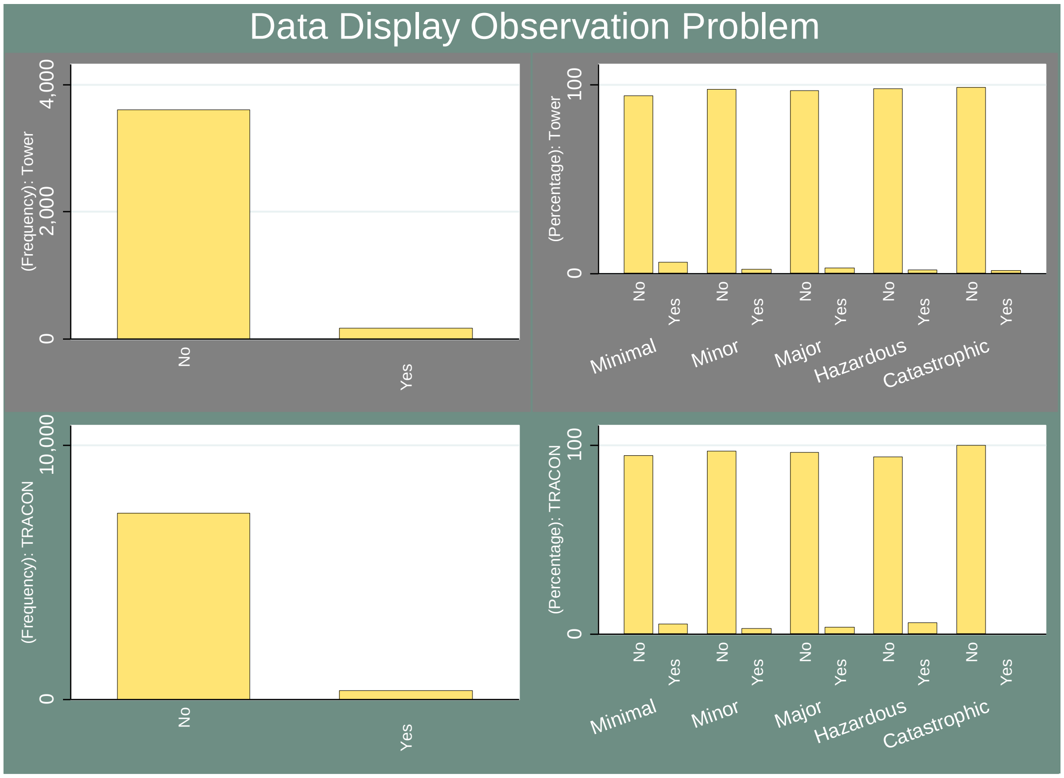

Figure 39 - Data, Display, and Observation Problems

In Tower facilities, Flight Plan Processing Problems have the largest number of associated incidents. Data Display Problems have the largest number of associated incidents in TRACON facilities. While it appears from the graph that Data Display problems are associated with catastrophic events; that is simply because there was only one catastrophic event with any type of data, display or observation problem in a Tower facility – a function of the small number of observations for this category.

6.4.7. Flight Plan/PDC Processing Problem

There are five variables describing different types of flight plan processing problems; all have very few observations, ranging from 15 (premature removal) to 78 (interpretation). The observations covering Flight Plan Processing are primarily low severity; only 3 Major incidents are present; there are no hazardous or catastrophic incidents. The small number of observations led Volpe to aggregate this variable, as well as analyze Tower and TRACON facilities together. As expected, *Logit results show that Flight Plan Processing problems are associated with Low Severity incidents.*

Table 61 - Binary Logit Flight Plan Processing Problem

Variable	Odds Ratio	Standard Error	P-Value	95% CI LB	95% CI UB	Obs
Flight Plan Processing Problem	0.107	0.058	0.00	0.037	0.310	183

6.4.8. Phraseology

Phraseology variables are used to denote when a pilot and controller use non-standard wording to communicate with each other. These variables looked at individually also have small numbers of observations (ranging from 71 to 219); thus they are also combined. Looked at both individually and as a group, phraseology issues do not have a statistically significant relationship with severity.

6.4.9. Radar Misidentification Problem

Radar Misidentification variables refer to when mistakes are made with radar readings due to:

- Overlapping datablocks
- Position and target correlation

There are only a small number of observations, so the two variables are combined, and Tower and TRACON facilities are analyzed together. None of the 105 incidents reported are Catastrophic, although 4 are Hazardous. The relatively high frequency of Major and Hazardous instances caused a binary logit analysis to show that incidents with a radar misidentification are substantially more likely to be Severe.

Table 62 - Binary Logit Radar Misidentification

Variable	Odds Ratio	Standard Error	P-Value	95% CI LB	95% CI UB	Obs
Radar Misidentification Problem	2.237	0.488	0.00	1.458	3.430	105

6.4.10. ATC/Pilot Communications/Clearance Model

Combining all of the ATC/Pilot Communications and Clearance variables into one multinomial model (run for Tower, TRACON, and combined Tower/TRACON) provided results that echoed the results found above when smaller groups of these variables were analyzed: the small number of observations available for these variables means that in many cases, even when a result is technically statistically significant, the standard error is too large to draw any meaningful conclusion.

Binary and multinomial models were both run in two groups: a subset of variables than encompasses the full ATSAP date range and all relevant variables from 2011-2013. Only two aggregated variable groups (loss of communication and readback problem) covered the full 2008-2013 ATSAP date range, while the remainder are available from 2011 on. Full model results are available in the appendix.

Most notably, in both the TRACON models, loss of communication is associated with a higher than expected frequency of catastrophic events. In both Tower and TRACON facilities, computer entry problems were associated with low severity incidents; it may be that these mistakes are typically resolved quickly before they become severe.

6.4.10.1. Tower ATC/Pilot Communications/Clearance Model

Table 63- Tower ATC/Pilot Communication/Clearance Multinomial Model (full date range)

	Minor	Major	Hazardous	Catastrophic
Loss of Communication	.73* (.11)	1.3 (.24)	.85 (.38)	1.8 (.6)
Readback Problem	.85 (.1)	.91 (.15)	1.1 (.29)	.21* (.15)
Total Operations	1** (.00092)	1 (.00097)	1 (.0015)	.99*** (.0031)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 6,836				

Table 64 - Tower ATC/Pilot Communication/Clearance Multinomial Model, (2011—2013)

	Minor	Major	Hazardous	Catastrophic
Loss of Commun.	.72 (.13)	1.2 (.27)	.6 (.31)	.78 (.39)
Readback	.76 (.12)	.82 (.17)	.77 (.3)	.3 (.22)
Acknowledgment	1.3 (.29)	1.8* (.47)	1.7 (.99)	2.3 (1.6)
Clearance Problem	1.7*** (.15)	1.1 (.15)	.58* (.12)	.084*** (.044)
Computer Entry Problem	.15 (.16)	6.0e-07*** (1.7e-07)	6.4e-07*** (2.2e-07)	8.7e-07*** (3.7e-07)
Data Display Problem	.97 (.28)	.98 (.45)	1.1 (.69)	.94 (.98)
Flight Plan/PDC Processing Problem	.099*** (.06)	3.0e-07*** (6.5e-08)	3.5e-07*** (8.0e-08)	1.0e-06*** (3.1e-07)
Phraseology	.76 (.19)	1.2 (.34)	1.4 (.59)	.25 (.26)
Radar Misidentification	.67 (.53)	3.4 (2.1)	2.8 (2.7)	4.2e-07*** (3.1e-07)
Total Operations	1*** (.00065)	1 (.0012)	1 (.0016)	.98*** (.0041)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 3,692				

6.4.11. TRACON ATC/Pilot Communications/Clearance Model

Table 65 - TRACON ATC/Pilot Communication/Clearance Multinomial Model (full date range)

	Minor	Major	Hazardous	Catastrophic
Loss of Communication	.73* (.11)	1.3 (.24)	.85 (.38)	1.8 (.6)
Readback Problem	.85 (.1)	.91 (.15)	1.1 (.29)	.21* (.15)
Total Operations	1** (.00092)	1 (.00097)	1 (.0015)	.99*** (.0031)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 6,836				

Table 66 - TRACON ATC/Pilot Communication/Clearance Multinomial Model (2011-2013)

	Minor	Major	Hazardous	Catastrophic
Loss of Commun.	.84 (.11)	1.7*** (.24)	1.2 (.48)	5.1*** (1.6)
Readback	1.1 (.17)	1.2 (.24)	.49 (.24)	.22 (.24)
Acknowledgment	1.4* (.22)	2*** (.35)	1.4 (.92)	1.5 (1.2)
Clearance Problem	1.9*** (.093)	1.2 (.22)	1.4 (.31)	.068*** (.049)
Computer Entry Problem	.86 (.31)	.27 (.23)	1.6e-06*** (5.4e-07)	4.1e-06*** (2.0e-06)
Data Display Problem	.73 (.14)	.59 (.18)	.68 (.58)	2.0e-06*** (5.7e-07)
Flight Plan/PDC Processing Problem	.57 (.17)	.33* (.15)	1.0e-06*** (3.0e-07)	2.4e-06*** (7.4e-07)
Phraseology	1.2 (.18)	.92 (.24)	2.6 (1.4)	.98 (1)
Radar Misidentification	.92 (.33)	2.9*** (.89)	5.1* (3.3)	3.5e-06*** (1.1e-06)
Total Operations	1*** (.0004)	1*** (.00056)	1** (.00071)	1 (.001)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 7,575				

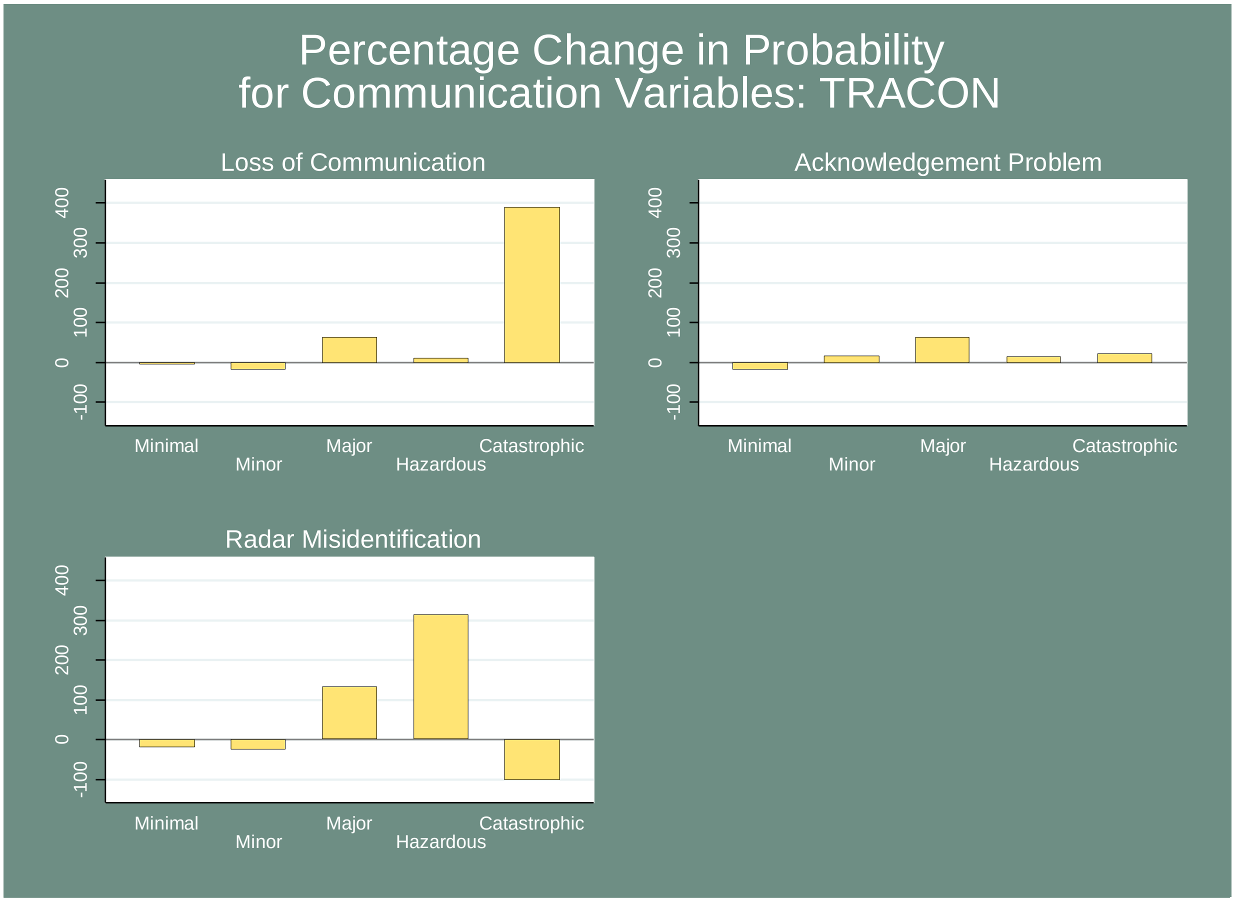

Figure 40 - Percentage Change in Probability for Communication Variables: TRACON

Figure 40 shows that loss of communication is significantly associated with major, hazardous, and catastrophic severity levels in TRACON facilities. The graph shows the percentage change in probability of an incident being classified at each incremental level of severity, moving from one level to the next. Acknowledgement problems in TRACON facilities are also associated with an increase in each consecutive level of severity, although not as dramatically. Radar misidentifications in TRACON facilities are associated with a higher risk of major and hazardous incidents.

6.5. Airspace and Pilot Action Variables

These variables describe the airspace the plane was operating in and pilot characteristics at the time that an incident occurred. These variables are exclusively causal factor and categorical in nature. The main concern surrounding the bulk of these variables is the subjectively of the pilot causal factors, since controllers initially fill out the reports. In more severe loss of separation incidents, controllers may be more apt to shift some of the culpability to pilots and mark down causal factors such as “pilot expectation bias.” It is important that any findings in this section be taken with this in mind with regards to pilots’ actions.

6.5.1. Airspace Design Issues

The airspace design and use variables are causal factors variables in ATSAP. Reporters have the option of indicating whether the airspace design (new, poor, or special use) played a significant factor in the loss of separation incident. Figure 41 and Figure 42 present the distribution of these variables by facility type.

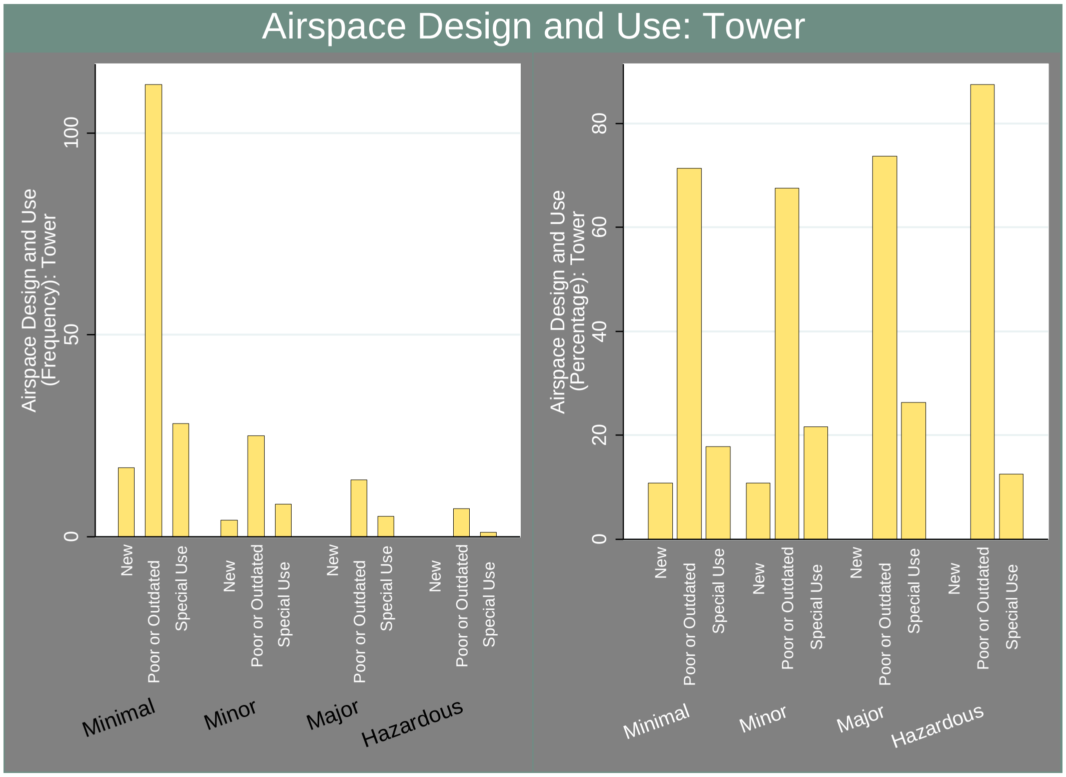

Figure 41 - Distribution of Airspace Design by Facility

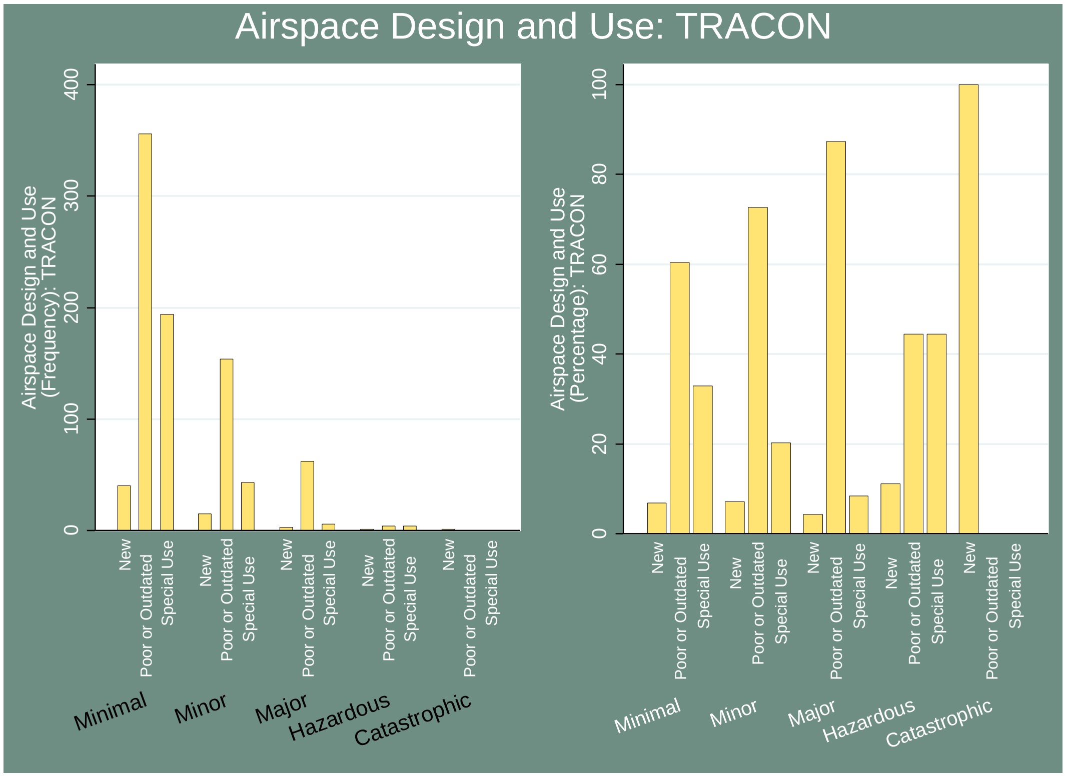

Figure 42 - Distribution of Airspace Design by Facility

Poor or outdated airspace design was the most frequent airspace design issue noted and tends to overwhelm the distributions for both tower and TRACON facilities in Figure 41 and Figure 42. Table 67 present the estimation results from a logit of the airspace design issues over severe/non-severe incidents for tower and TRACON facilities. For tower facility incidents, new airspace design issues was indicated for only non-severe incidents, forcing this variable to be omitted in the binary logit. The other airspace design issues appear to not have any statistical relationship with severity.

The TRACON facility incidents see a marginal statistical relationship with new airspace designs, lowering the likelihood of a more severe incident, and a strong statistical relationship with special use airspaces, also lowering the likelihood a severe incident. Special use airspaces could have certain underlining rules in place that would limit the overall operational activity. This effect could therefore reduce the probability of a more severe loss of separation incident.

Table 67 - Logit Estimate of Airspace Design Issues for Tower Incidents

Variable	Odds Ratio	Standard Error
Airspace Design New	-	-
Airspace Design Poor or Outdated	1.482	0.450
Special Use Airspace	1.082	.494
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 **N = 5,466**		

Table 68 - Logit Estimate of Airspace Design Issues for TRACON Incidents

Variable	Odds Ratio	Standard Error
Airspace Design New	0 .473	0.212
Airspace Design Poor or Outdated	1.169	0.265
Special Use Airspace	0.287***	0.078
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001 **N = 11,223**		

6.5.2. Airspace Classification

Airspace type D increases the likelihood of a Catastrophic incident relative to the Minimal outcome for both Tower and TRACON facilities

Airspace classification refers to the standard alphabetical classification of airspace used by the FAA. ATSAP reports the airspace type for all 5 categories, but since this report only focuses on incidents that occurred within terminal airspace, only airspaces B, C, and D will be examined.[63] Figure 43 presents the distribution of airspace type by facility for incidents.

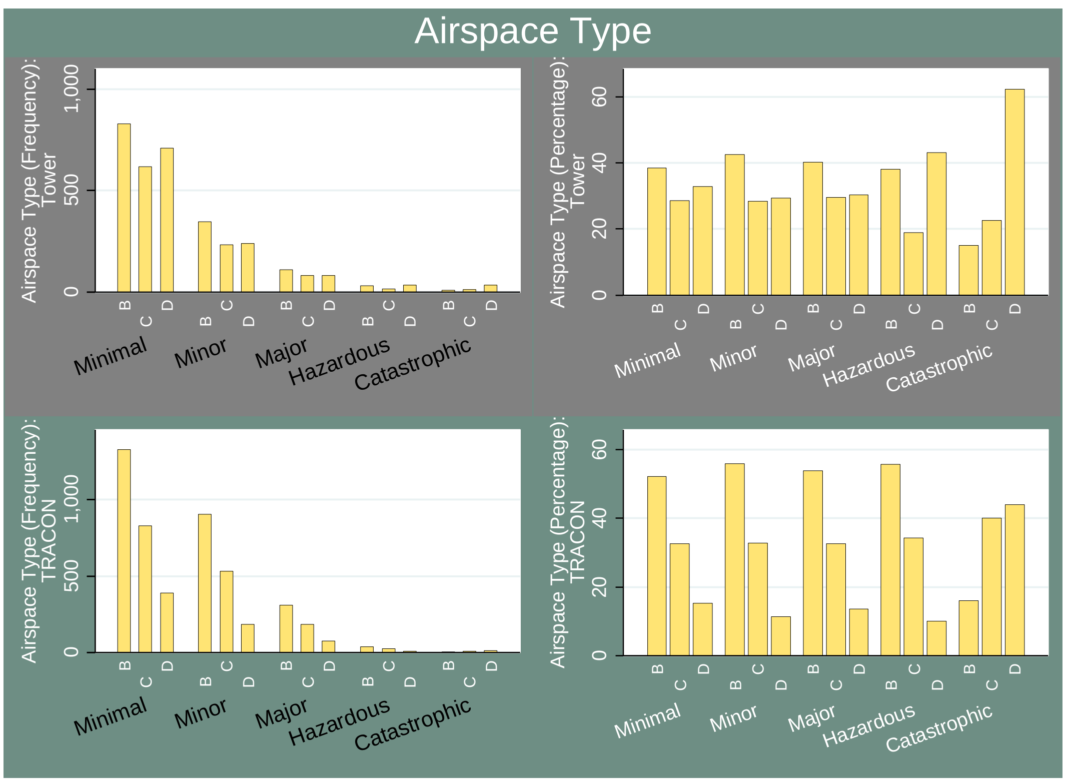

Figure 43 - Distribution of Airspace Type by Facility

[63] For reference, airspace type B is defined as starting at the surface up to FL 100 (10,000 feet), airspace type C starts at the surface up to FL 40, and airspace type D starts at the surface up to FL 25.

Airspace B was the most frequent airspace type for both tower and TRACON incidents; however, the majority of Catastrophic incidents occurred in airspace D for both facility types. The estimation results of a multinomial logit of airspace type over severity level for tower incidents presented in Table 69 accentuate this finding. Airspace type D incidents are more likely to be in the Catastrophic category, while airspace type B appear to decrease the likelihood of being in the Catastrophic category for tower facilities. Estimation results for TRACON facilities presented in Table 70 are more mixed, with no statistically significant results (at the 5% p-level) for the Catastrophic category.

Since airspace type D are associated with low altitude flying aircraft, it is important to control for what type of aircraft were flying and the phase of flight the aircraft was engaged in before coming to any conclusions about the airspace type (see section 6.5.9)..

Table 69 - Multinomial Logit Estimate of Airspace Types for Tower Incidents

	Minor	Major	Hazardous	Catastrophic
Airspace Type B	1.290* (0.161)	1.055 (0.148)	0.911 (0.224)	0.359* (0.146)
Airspace Type C	1.154 (0.105)	1.038 (0.145)	0.611 (0.171)	0.722 (0.261)
Airspace Type D	1.039 (0.106)	0.926 (0.124)	1.206 (0.242)	1.728* (0.381)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 6,916				

Table 70 - Multinomial Logit Estimate of Airspace Types for TRACON Incidents

	Minor	Major	Hazardous	Catastrophic
Airspace Type B	1.534*** (0.192)	1.339* (0.197)	1.647* (0.324)	0.376 (0.197)
Airspace Type C	1.438*** (0.132)	1.294 (0.175)	1.622 (0.409)	1.506 (0.533)
Airspace Type D	1.067 (0.139)	1.139 (0.199)	1.006 (0.383)	3.521** (1.377)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 13,843				

6.5.3. Aircraft Equipment Issues

This variable is a combination of causal factor variables in ATSAP that deal with aircraft equipment issues. Aircraft equipment issues were rarely selected as being a factor in a loss of separation event, with only 102 out of 11,870 incidents noting some sort of aircraft equipment issue for both facility types.

This variable is too rare to allow for any visual representation across severity levels. A Chi-Square test indicates that there is no statistically significant relationship between this variable and severity levels. Although this variable lacks enough variation to be statistically significant, theoretically an aircraft equipment issue could such as transponder malfunction, or something else that would cause a loss in communication with ATC or the loss of control of an aircraft could indeed lead to a loss of separation incident.

6.5.4. Aircraft Performance or Pilot Response

These variables are a set of causal factors defined as the combined performance of the aircrew and the aircraft capabilities that contributed to a loss of separation incident. This set includes the compression on a final approach, untimely aircraft descent/climb, untimely aircraft turn, untimely aircraft roll, untimely runway exit, and untimely speed adjustment. The distributions of these variables are presented in Figure 44-Figure 46.

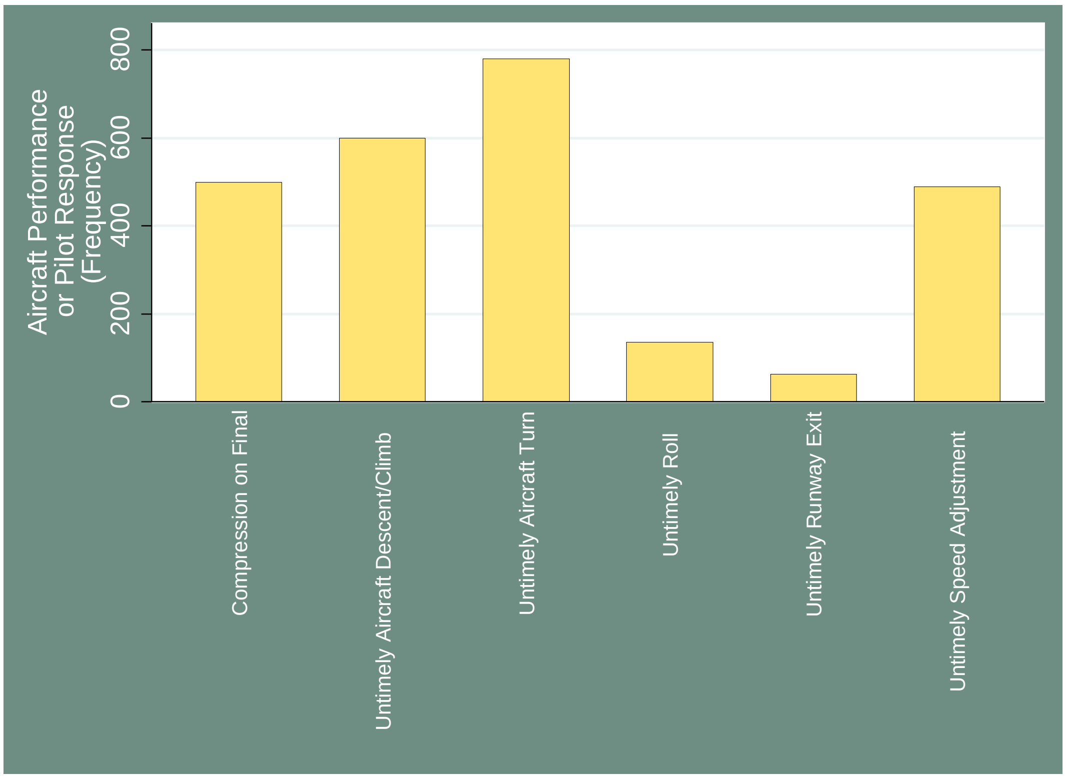

Figure 44 - Distribution of Aircraft Performance or Pilot Response

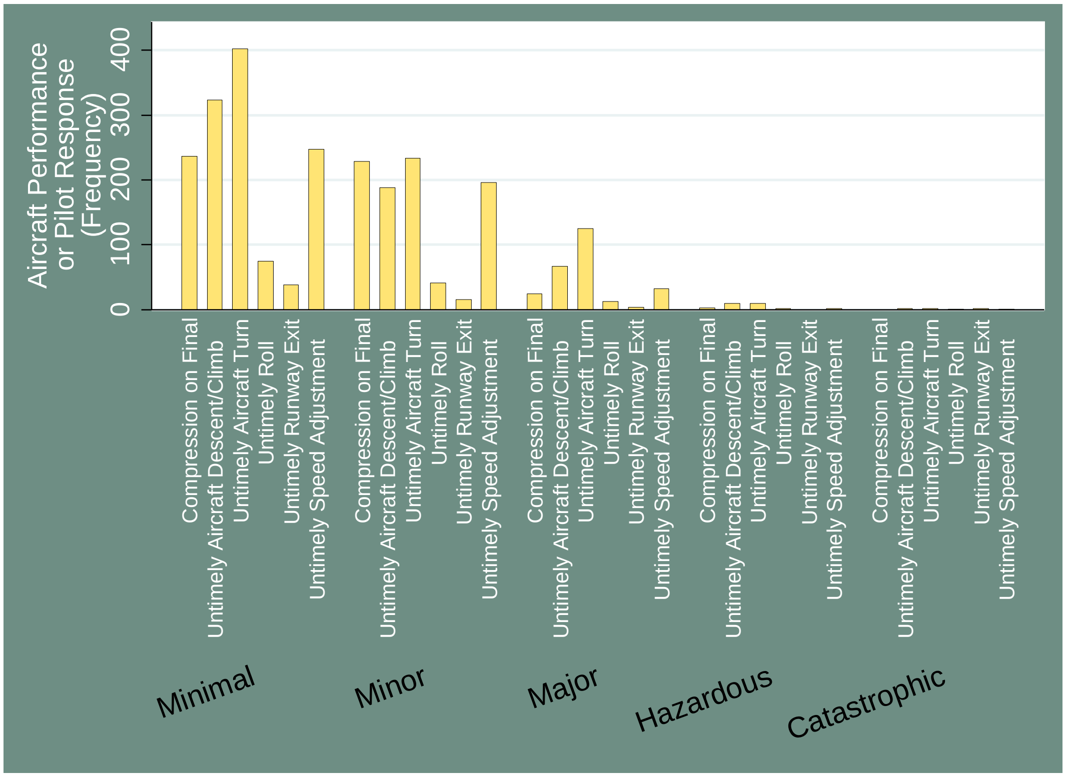

Figure 45 - Distribution of Aircraft Performance or Pilot Response over Severity

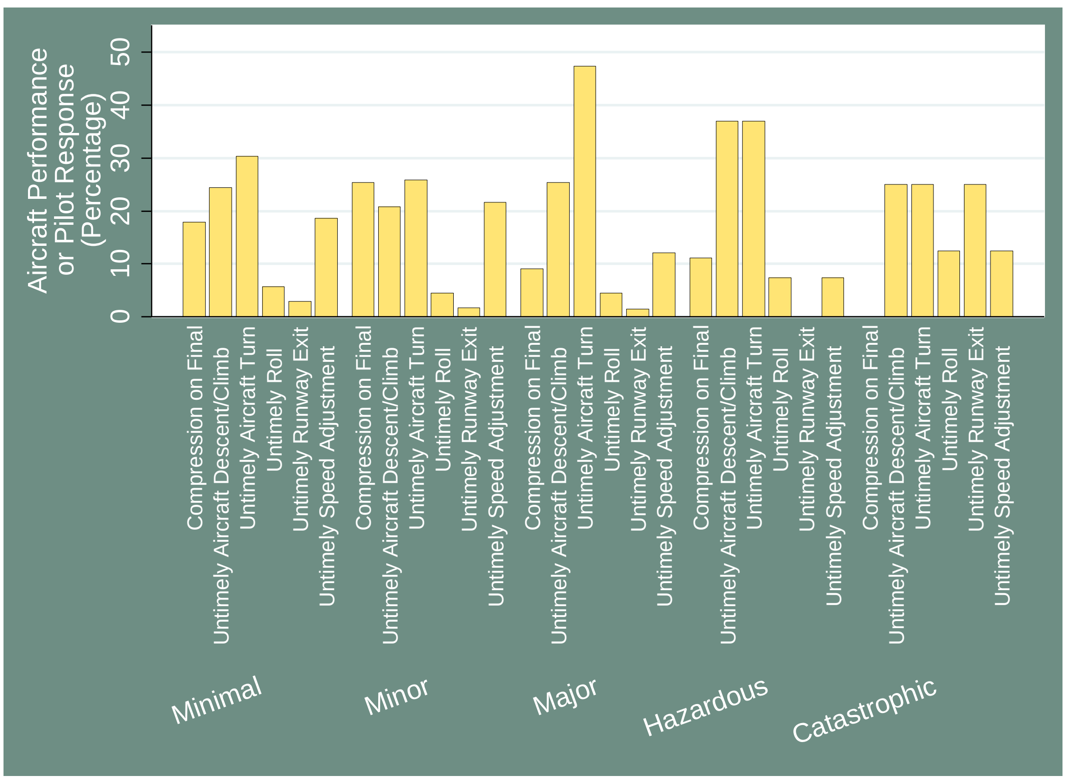

Figure 46 - Distribution (Percentage) of Aircraft Performance or Pilot Response over Severity

Presumably, these variables should be closely related with the variables presented in the previous section. For example if an evasive action occurred, it is possible that it could have caused or contributed to any number of these pilot responses. The correlation between these two sets of variables failed to yield any noticeable relationships, however.

In order to better understand the relationship between these variables and severity, a multinomial logit was estimated. Table 71 presents these results. There appears to be fewer statistical relationships between severity categories and this set of pilot response variables as compared to the previous section. Compression on final approach, untimely aircraft descent/climb, and untimely aircraft turn all appear to increase the likelihood of Minor incidents. Compression on final approach appears to decrease the likelihood of Major incidents, while untimely aircraft turns increase the likelihood of a Major incident. Finally, untimely aircraft runway exits appear to reduce the likelihood of Hazardous incidents, but this is most likely a statistical artifact because zero Hazardous incidents included an untimely runway exit. Unfortunately, not a lot can be taken away from these results, and a full airspace/pilot response model is necessary to better understand the relationships presented here.

Table 71 - Multinomial Logit Estimate of Aircraft Performance or Pilot Response

	Minor	Major	Hazardous	Catastrophic
Untimely Aircraft Descent/Climb	1.3** (.11)	1.1 (.13)	.85 (.27)	.74 (.33)
Untimely Aircraft Turn	1.3** (.11)	1.7*** (.22)	.95 (.3)	.29 (.2)
Compression on Final	2.1*** (.19)	.63** (.11)	.31 (.2)	.17 (.18)
Untimely Roll	1.4 (.28)	.89 (.3)	.92 (.68)	.67 (.69)
Untimely Runway Exit	.64 (.2)	.91 (.57)	4.1e-06*** (8.2e-07)	3 (2.1)
Untimely Speed Adjustment	1.2 (.15)	.83 (.17)	.45 (.34)	.42 (.44)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 **N = 13,843**				

6.5.5. Pilot Expectation Bias

This is a causal factor variable found in ATSAP that is defined as the pilots' strong belief towards a particular outcome based on frequently encountered situations, even when evidence to the contrary contributed to a loss of separation incident. For example, a pilot always crosses a fix at 10,000 feet but is instructed to cross at 11,000 feet and the pilot's cognitive bias causes him/her to cross at 10,000 feet, leading to a loss of separation incident.[64] This variable's legitimacy can surely come under question, since the ATSAP report is almost exclusively filled by controllers. Using the same example from above, it is unclear how the controller knows for certain that this particular pilot ignored ATC or if the pilot simply misunderstood ATC.

[64] ATSAP Data Dictionary

This variable is not filled out often enough to allow for any basic visualization of the data to be meaningful. A Chi-Squared test fails to find any relationship with this variable and severity categories. Given this variables subjectivity noted above, it would have been unadvisable to include this variable in any full model.

6.5.6. Pilot Reaction

No pilot reaction reduced the likelihood of a Catastrophic incident relative to the Minimal outcome. Evasive actions and go-arounds increased the likelihood of a Minor, Major and Hazardous incident relative to the Minimal outcome.

The pilot reaction variables are a set of causal factors found in ATSAP referring to the manner of a pilot's reaction which significantly contributed to a loss of separation incident. These variables include evasive actions to avoid collision with another object, go-arounds, responses to a resolution advisory initiated from a traffic collision avoidance system – resolution advisory (TCAS-RA), and none or unknown reactions. Figure 47-Figure 49 present the distribution of these variables.

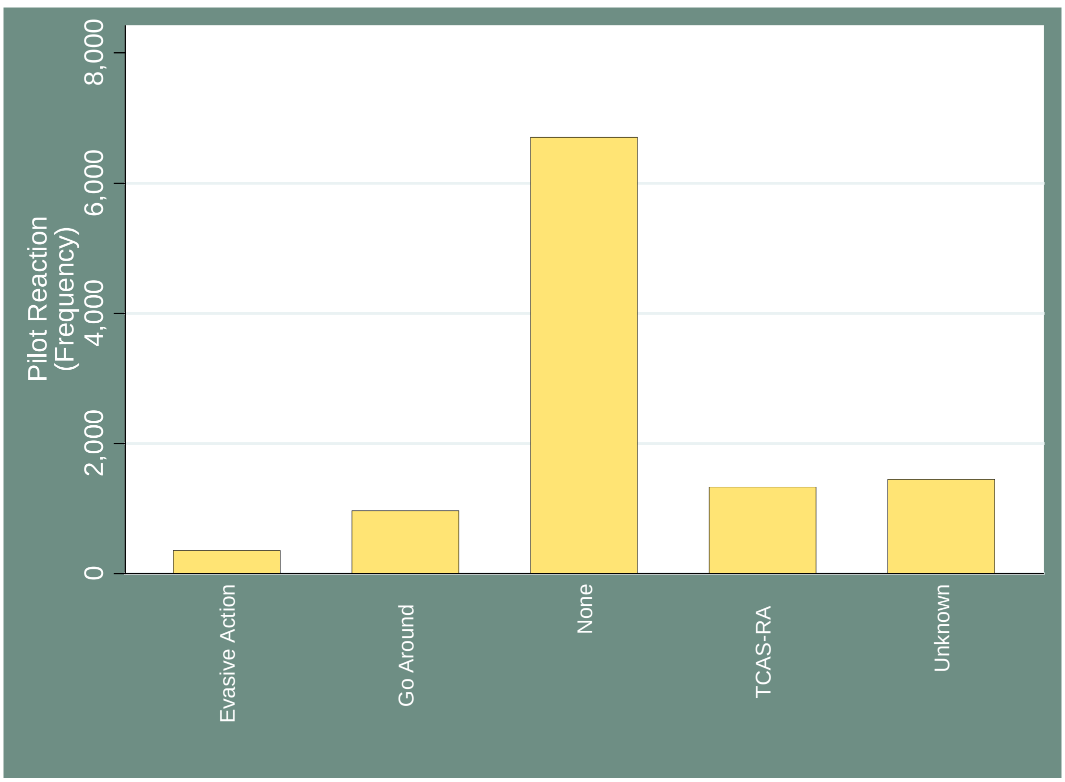

Figure 47 - Distribution of Pilot Reaction

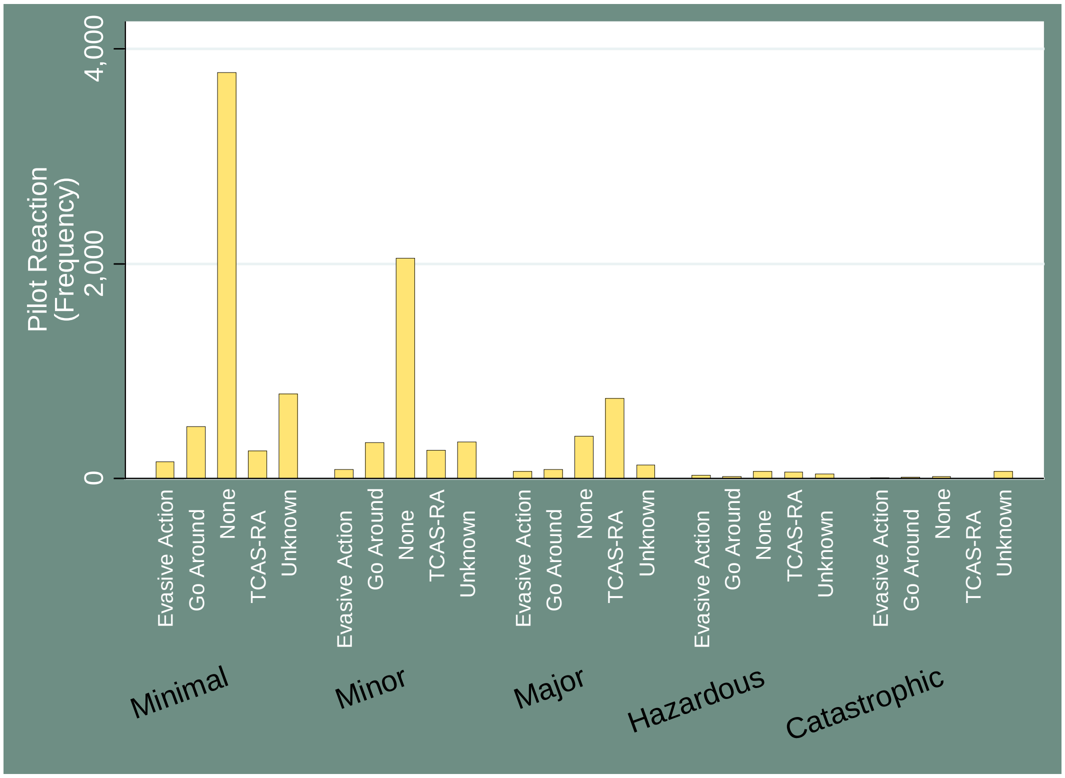

Figure 48 - Distribution of Pilot Reaction over Severity

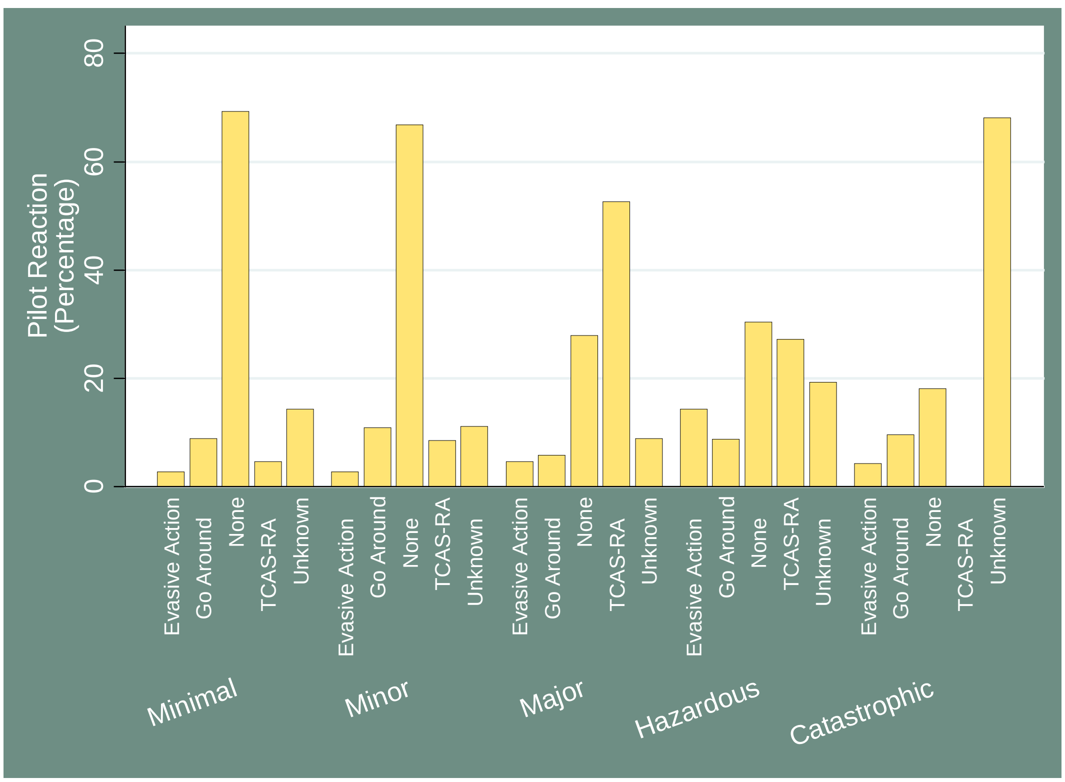

Figure 49 - Distribution (Percentage) of Pilot Reaction over Severity

To better understand the relationship between these pilot reaction variables and the severity categories, a multinomial logit was estimated. Table 72 presents these results. There appears to be strong statistical relationships between different types of pilot reactions and the severity categories. The main findings here include: evasive actions, go-arounds and TCAS-RA all increase the likelihood of severity falling in categories Minor, Major and Hazardous, while TCAS-RA and no (none) reactions reduce the likelihood of a Catastrophic incident. The relationship between the TCAS-RA variable and Catastrophic incidents is suspect, though, because there are zero observations where TCAS-RA was listed as a pilot reaction for Catastrophic incidents. However, this could also mean that the TCAS system is fulfilling its role; TCAS responses should only go into effect when the situation becomes extremely dangerous to avoid the worst possible outcome, so seeing zero observations in the dataset confirms in many ways that the system works.

Finally, it is unclear how to interpret the results for unknown reactions, due to a lack of a formal definition of this variable.[65] Therefore, it should not be included in the full airspace/pilot response model.

Table 72 - Multinomial Logit Estimate of Pilot Reaction

	Minor	Major	Hazardous	Catastrophic
Pilot Reaction: Evasive Action	1.5*** (.17)	3.4*** (.59)	9.1*** (1.9)	2.3 (1.1)
Pilot Reaction: Go-Around	1.9*** (.17)	1.6** (.23)	1.9** (.44)	1.6 (.53)
Pilot Reaction: TCAS-RA	2.8*** (.43)	27*** (5.5)	9.5*** (1.6)	7.4e-07*** (1.5e-07)
Pilot Reaction: Unknown	1.2* (.11)	1.6*** (.2)	3*** (.56)	6.5*** (.99)
Pilot Reaction: None	1.5*** (.1)	1.1 (.088)	.99 (.14)	.36*** (.11)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 21,325				

6.5.7. Non-Conformance with a Clearance

This set of causal factor variables refer to a pilots' improper execution of a clearance which significantly contributed to a loss of separation incident, and includes altitude, altitude crossing, course, and speed. The distribution of these variables are presented in Figure 50-Figure 52.

[65] This variable is not included in the ATSAP Data Dictionary.

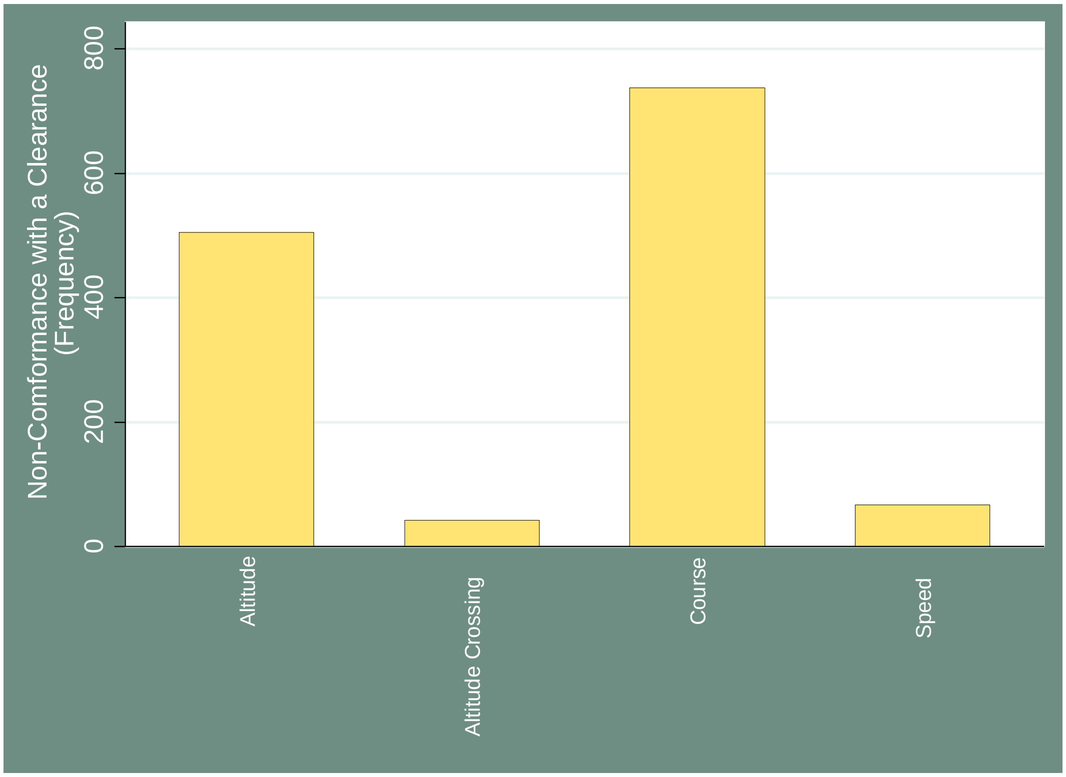

Figure 50 - Distribution Non-Conformance with a Clearance

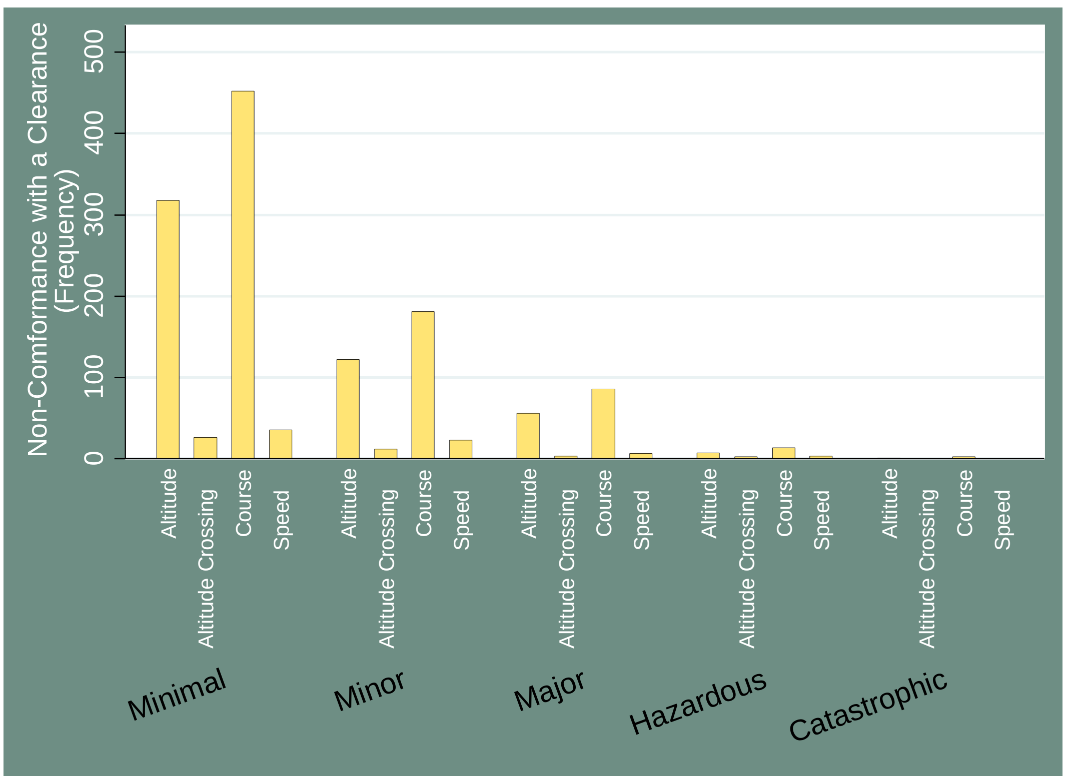

Figure 51 - Distribution Non-Conformance with a Clearance over Severity

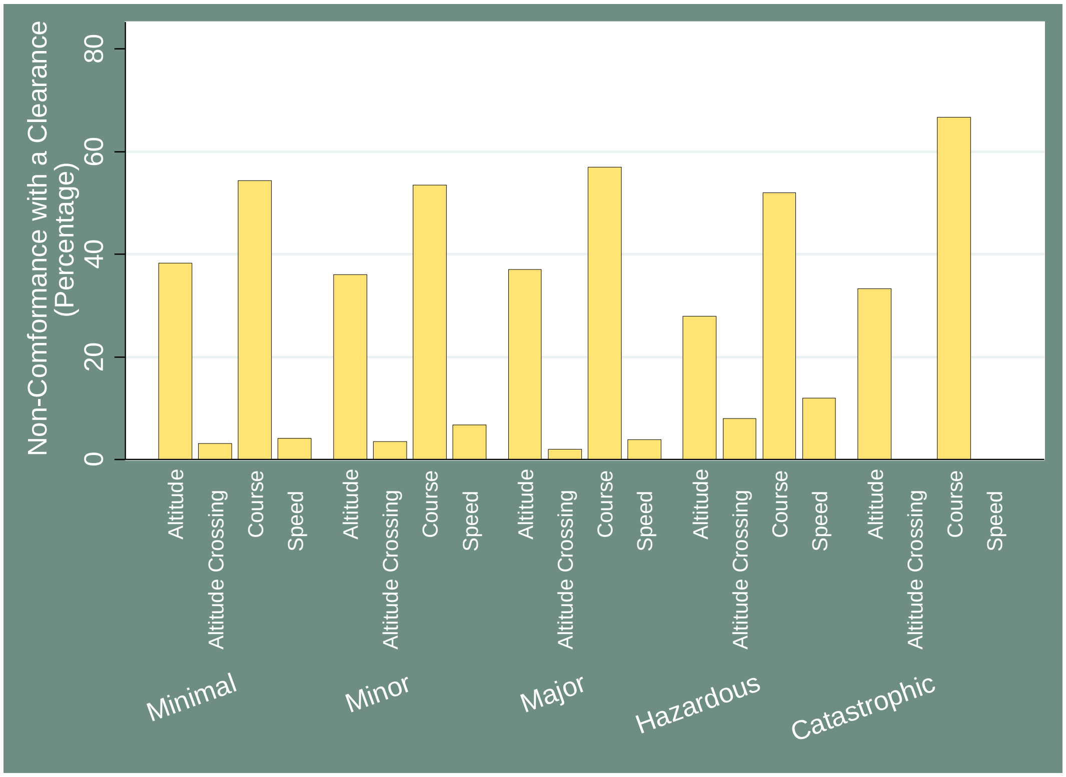

Figure 52 - Distribution (Percentage) Non-Conformance with a Clearance over Severity

It is important to note here that there are no Catastrophic incidents where altitude crossing and speed were a factor. Since the course and altitude variables have significantly more observations than altitude crossing and speed, a multinomial logit was estimated to try and better understand the relationship between these non-conformance variables and severity categories, with the model estimations presented in Table 73. The only statistically significant variables are altitude and speed for the Catastrophic category. This result is undoubtedly from the absence of these variables in Catastrophic incidents. Without any observations, the true relationship between altitude crossing and speed with Catastrophic incidents is impossible to determine.

Table 73 - Multinomial Logit Estimate of Non-Conformance with a Clearance

	Minor	Major	Hazardous	Catastrophic
Non-Conformance with a Clearance: Altitude	.94 (.1)	.96 (.12)	.65 (.22)	.17 (.18)
Non-Conformance with a Clearance: Altitude Crossing	1.2 (.41)	.7 (.38)	2.6 (1.7)	3.9e-06*** (1.4e-06)
Non-Conformance with a Clearance: Course	.9 (.086)	1.1 (.16)	1.1 (.33)	.27 (.2)
Non-Conformance with a Clearance: Speed	1.5 (.5)	.95 (.45)	3.1 (1.9)	2.8e-06*** (8.6e-07)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 **N =11,758**				

6.5.8. Procedure Issues

This set of causal factor variables contains types of policies or practices that contributed to a loss of separation incident. It should be noted that this set of procedure issues are different from the ones found in 6.2.2.3 in that these variables deal with airspace and aircraft control directly. The variables in this set include conventional procedure issues, directive/publication/regulation type problems, and RNAV (Area Navigation) procedure issues. The distributions of these variables are presented in Figure 53.

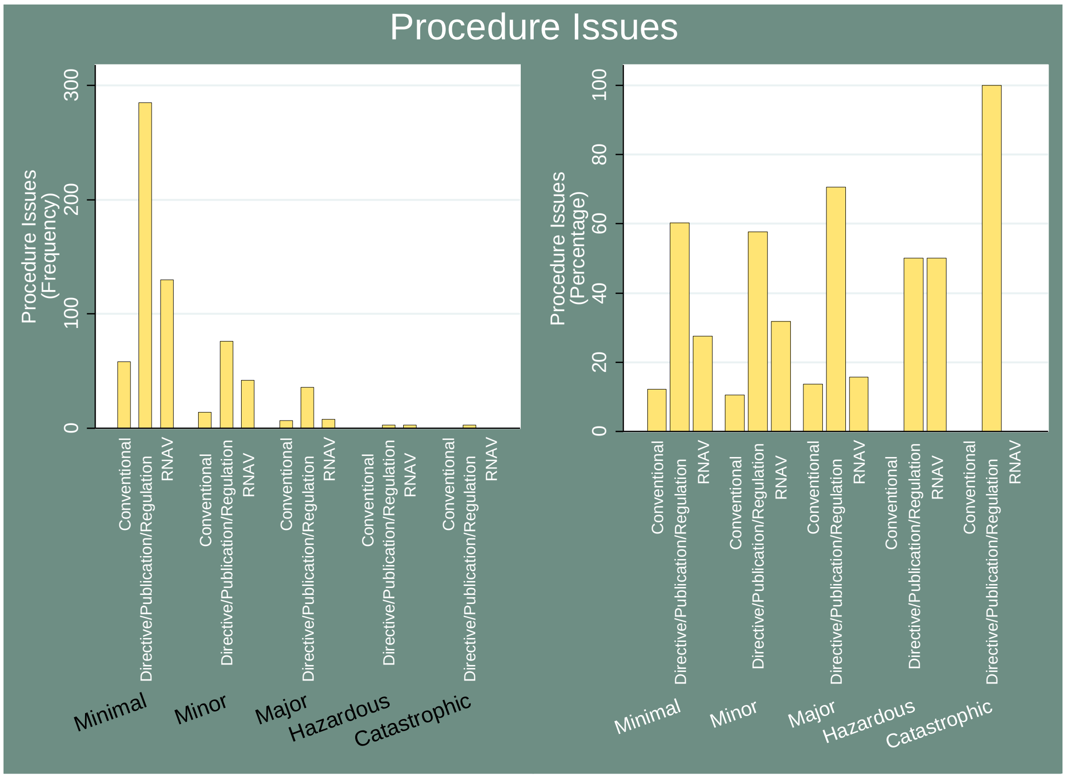

Figure 53 - Distribution Procedure Issues over Severity

There appears to be a lack of observations for the conventional procedure and RNAV issues at the Catastrophic severity level. No logit was estimated due to this subgroups' lack of overall variation. It would not be appropriate to aggregate these causal factors into one 'procedure issue' causal factor category given the distinct difference between these variables. As a final note, only the directive/publication/regulation problem variable will be the included in a final airspace and pilot response model.

6.5.9. Airspace and Pilot Response Model

This section contains the multinomial logit models of airspace and pilot response variables examined in the previous section. These models will not contain variables that lack any observations at several severity levels and also lack the ability to be aggregated into a related variable, due to the statistical impact this has on the models. Models are broken down by facility level first due to the differences in airspace types, and then a full terminal airspace model containing both facility types is estimated for comparison purpose. The natural log of daily operations and year indicators are included in all models as control variables.

6.5.9.1. Tower Multinomial Logit Model

Table 74 presents the multinomial logit estimations for tower facilities. Airspace type is relatively unimportant (in a statistical sense) for Tower facilities, with only airspace type D having a statistically significant relationship, decreasing the likelihood of a Minor incident. This result is not necessarily surprising, given that Tower facilities generally control type D airspace. Evasive actions are significant and increase the likelihood of being in Major and Hazardous categories, while Go-Arounds are significant and increase the likelihood of being in categories Minor and Major. Surprisingly, the No Pilot Action variable is only significant in the Minor group, increasing the likelihood. This seems at odds with logic; if the pilot or aircrew willingly did not respond to a situation, one would think this could increase the chance of a more severe event. This relationship will be important to examine in the other two models.

Table 74 - Multinomial Logit Estimate of Airspace and Pilot Response Variables for Tower Facilities

	Minor	Major	Hazardous	Catastrophic
Airspace Type B	.79 (.11)	1.5 (.37)	1.2 (.42)	.75 (.58)
Airspace Type C	.76 (.16)	1.2 (.32)	.78 (.37)	.84 (.5)
Airspace Type D	.65* (.13)	1.3 (.33)	1.3 (.53)	1.8 (.75)
Daily Operations	1.2** (.069)	1.1 (.095)	1.1 (.12)	.55*** (.074)
Pilot Reaction: Evasive Action	1.5 (.45)	6.4*** (2.1)	10*** (3.5)	1.8 (1.4)
Pilot Reaction: Go Around	1.4* (.2)	1.9** (.4)	.89 (.37)	1.3 (.57)
Pilot Reaction: None	1.4 (.24)	.7 (.17)	.54 (.22)	.35 (.21)
Untimely Aircraft Descent/Climb	1.8** (.32)	2.3** (.66)	1.3 (.76)	2 (1.2)
Untimely Roll	1.7* (.36)	1.2 (.41)	.75 (.57)	.52 (.54)
Untimely Speed Adjustment	2.4*** (.38)	1 (.34)	.65 (.49)	.6 (.57)
Year 2011 Indicator	2* (.65)	1.1 (.39)	4.1 (4.3)	.89 (.57)
Year 2012 Indicator	1.1 (.33)	.99 (.35)	2.5 (2.7)	.57 (.35)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 3,692				

6.5.9.2. TRACON Multinomial Logit Model

The estimation results for the multinomial logit based on TRACON facilities are presented in Table 75. There appears to be a sharp contrast in the role of the airspace type for TRACON's compared to tower facilities. Airspace types B, C and D all appear to increase the likelihood of fall into the Major category. Airspace type C decreases the likelihood of being in the Catastrophic category, while airspace type D increases the likelihood of being in the Catastrophic category (similar to tower facilities). Incidents that occurred in newly implemented airspace design are less likely to fall in the Major category. Pilot Go-Arounds increase the probability of Minor incidents, while decreasing the probability of Hazardous incidents. Again, there is only a marginal relationship between no pilot reaction and severity levels, with the only statistically significant category being Major (reducing the likelihood). This is an unexpected result and lacks a clear answer.

Table 75 - Multinomial Logit Estimate of Airspace and Pilot Response Variables for TRACON Facilities

	Minor	Major	Hazardous	Catastrophic
Airspace Design New	.95 (.21)	.28* (.17)	1.1 (1.1)	2.1 (2.2)
Airspace Type B	1.1 (.16)	2.6*** (.39)	2.6* (1.2)	2.5 (1.6)
Airspace Type C	.81 (.14)	2.8*** (.48)	.69 (.56)	1.9 (1.5)
Airspace Type D	.94 (.18)	3.3*** (.73)	2.5 (1.6)	12*** (6.8)
Daily Operations	1.3*** (.078)	1.5*** (.13)	1.3* (.19)	.89 (.18)
Pilot Reaction: Go Around	2.3** (.58)	.79 (.31)	1.2e-06*** (4.2e-07)	1.4 (1.6)
Pilot Reaction: None	.77* (.083)	.22*** (.058)	.41* (.18)	.12* (.11)
Untimely Aircraft Turn	1.4*** (.14)	1.5* (.22)	.53 (.31)	.75 (.56)
Year 2011 Indicator	3.8*** (.96)	1.8 (.62)	4.3 (4.5)	501404*** (180956)
Year 2012 Indicator	1.8** (.39)	1 (.32)	2.1 (2.2)	319745*** (91482)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 7663				

6.5.9.3. Terminal Airspace Multinomial Logit Model

A terminal airspace multinomial model was estimated, which includes both tower and TRACON facilities. Combining facilities allows for a more complete set of variables to be estimated due to higher number of observations across severity levels. Indicators for tower and TRACON facilities for included as control variables. These results are presented in Table 76. When facilities are combined, airspace types retain their statistical significance with the direction of the coefficients mirroring the TRACON airspace model results. Figure 54 and Figure 55 present the change in probability of severity categories for selected airspace and pilot response variables of interest. All three airspace types increase the probability a Major incident, ranging from 7.5-10% percent. Airspace type D increases the likelihood of Hazardous and Catastrophic incidents by the most, with an over 200% change in the percentage change of a Catastrophic incident. Pilot evasive actions increase the probability of a Major incident by a considerable amount (approximately 18%). Still at large with the general expectations, no pilot reaction appears to decrease the likelihood for all of the more severe incidents (Major, Hazardous, and Catastrophic), while increasing the likelihood of Minimal incidents by 10%.

The important take away from these results are that the airspace types are strictly defined to be within terminal airspace. Therefore, when interpreting the change in probabilities for these airspace variables, it is important to keep in mind the fundamental differences between terminal airspace and other types of airspace. These results confirm the notion that terminal airspaces are generally associated with more severe loss of separation incidents. Moreover, being able to control for possible emergency pilot reactions or errors allows for a better understand of each terminal airspace type.

Table 76 - Multinomial Logit Estimate of Airspace and Pilot Response Variables in Terminal Airspace

	Minor	Major	Hazardous	Catastrophic
Airspace Design New	1 (.23)	.24* (.14)	.65 (.68)	1.1 (1.2)
Airspace Design Poor or Outdated	.9 (.13)	1 (.38)	1.2 (.58)	3.5e-07*** (8.1e-08)
Airspace Design Special	.53* (.13)	.32** (.13)	.52 (.39)	2.2e-07*** (5.5e-08)
Airspace Type B	1 (.11)	1.9*** (.29)	1.7 (.49)	1.4 (.72)
Airspace Type C	.82 (.11)	1.8*** (.3)	.76 (.36)	1.2 (.53)
Airspace Type D	.8 (.11)	2*** (.34)	1.6 (.55)	3.4** (1.3)
Compression on Final	2*** (.17)	.59** (.11)	.23 (.18)	.22 (.23)
Daily Operations	1.3*** (.064)	1.4*** (.11)	1.3** (.13)	.7* (.098)
Directive/Publication/Regulation Issues	.67** (.091)	.8 (.14)	.39 (.22)	.6 (.34)
Non-Conformance with a Clearance: Altitude	.98 (.11)	.84 (.11)	.84 (.29)	.23 (.23)
Non-Conformance with a Clearance: Course	1 (.11)	.97 (.16)	1.1 (.35)	.32 (.23)
Pilot Reaction: Evasive Action	1.6* (.31)	8.3*** (1.7)	9.6*** (2.6)	1.5 (1.1)
Pilot Reaction: Go Around	1.6*** (.21)	1.4 (.28)	.82 (.34)	1.2 (.55)
Pilot Reaction: None	.94 (.091)	.33*** (.072)	.47* (.15)	.2** (.11)
Tower Only Events	1.5 (.31)	2* (.51)	2.9 (1.7)	1.2 (1.3)
TRACON Only Events	1.8** (.4)	2.4*** (.63)	1.1 (.66)	.59 (.61)
Untimely Aircraft Descent/Climb	1.3** (.12)	1.1 (.13)	1.2 (.36)	1.1 (.48)
Untimely Aircraft Turn	1.4*** (.13)	1.8*** (.22)	1 (.33)	.37 (.25)
Untimely Roll	1.9** (.39)	1.3 (.41)	.79 (.59)	.49 (.51)
Untimely Runway Exit	.62 (.21)	1.1 (.67)	1.4e-07*** (3.5e-08)	2.1 (1.6)
Untimely Speed Adjustment	1.3 (.16)	.8 (.17)	.45 (.34)	.41 (.43)
Year 2011 Indicator	3.3*** (.68)	1.5 (.42)	4.2 (3.1)	1.5 (.89)
Year 2012 Indicator	1.6** (.28)	.96 (.25)	2.3 (1.7)	1 (.59)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N =13,832				

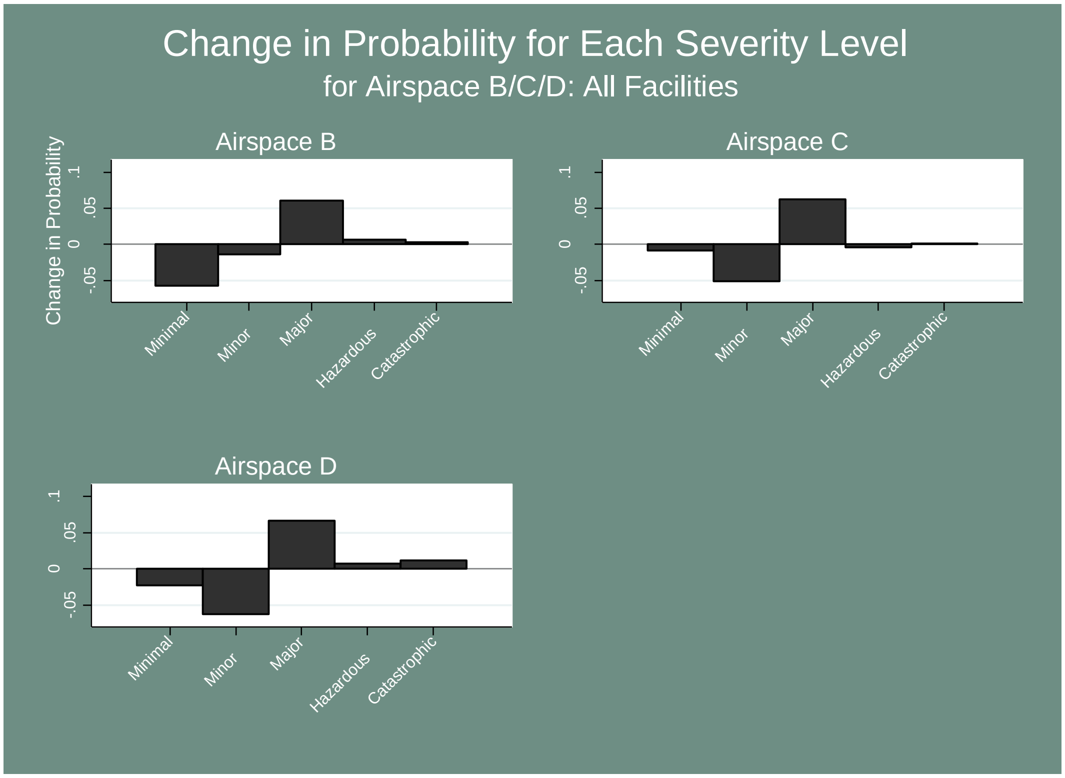

Figure 54 - Change in Probability of Severity Categories for Airspace/Pilot Response Variables: Terminal Airspace

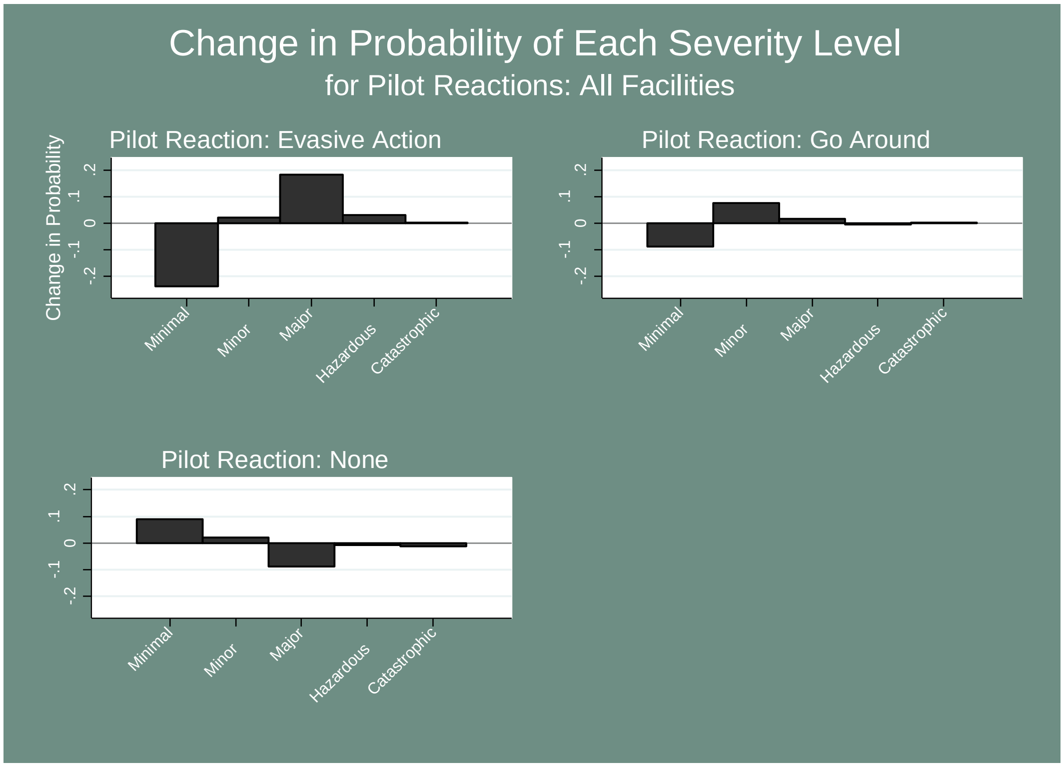

Figure 55 –Change in Probability of Severity Categories for Airspace/Pilot Response Variables: Terminal Airspace

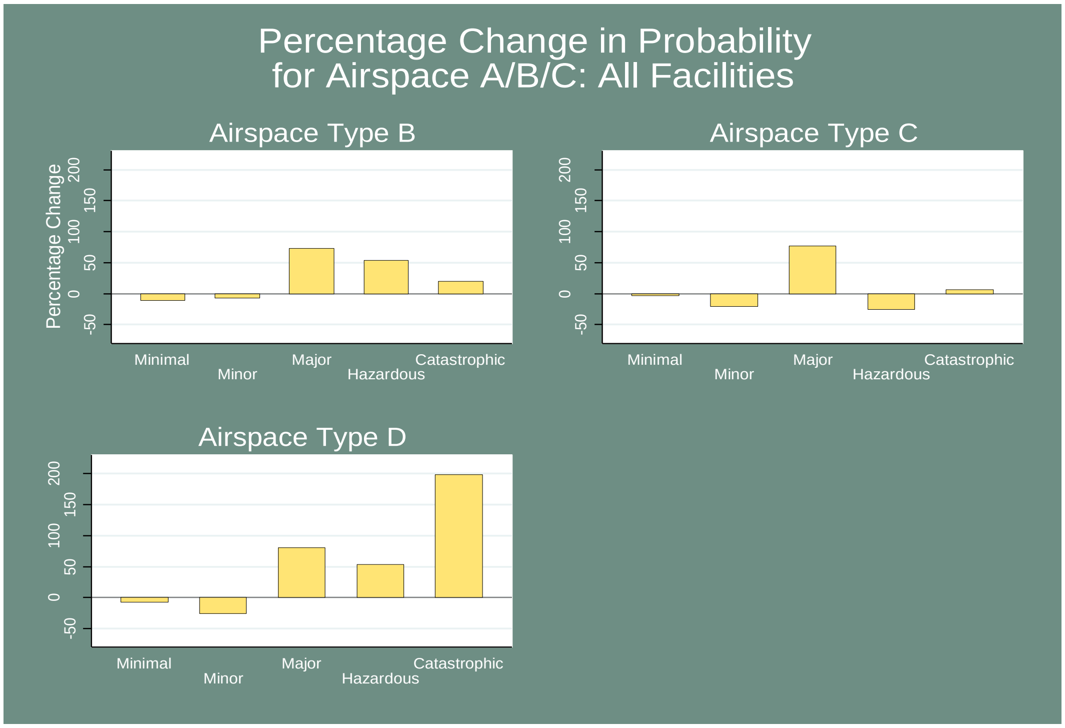

Figure 56 - Percentage Change in Probability of Severity Categories for Airspace/Pilot Response Variables: Terminal Airspace

Figure 57 - Percentage Change in Probability of Severity Categories for Airspace/Pilot Response Variables: Terminal Airspace

6.6. Weather Characteristics Variables

The weather variables characterize the weather conditions during the incident. As described in Section 4.2, the weather data originates from the METAR data archived by Plymouth University. When interpreting the data presented in this section, it is important to note that weather conditions are based on the location of the event. For tower incidents, weather data is always local METAR data. For TRACON events, weather data is either at the positively identified location of the event (identified through the use of ATSAP event location information), or when this is not possible, the weather data the TRACON's primary airport. Due to this discrepancy in how the weather data was assembled based on facility, weather data is presented only by facility type.

Weather causal factors were not examined in this section due data quality issues, aside from the weather complexity factor variable. The weather causal factors lacked the amount of responses needed to be useful for statistical analysis.

6.6.1. Dew Point by Facility Type

This variable provides an estimate of the dew point at the time of the incident. The dew point indicates the temperature at which water vapor in the air condenses into liquid water. Higher dew points are associated with more humid air and severe weather.[66] As with the many of the weather variables, it is unlikely that a higher or lower dew point causes increased or decreased severity. However, factors related to dew point (such as haziness or approaching weather) may contribute to increased or decreased severity. Figure 58 presents the distribution of the dew point by facility type.

[66] National Weather Service Weather Forecast Office (2012).

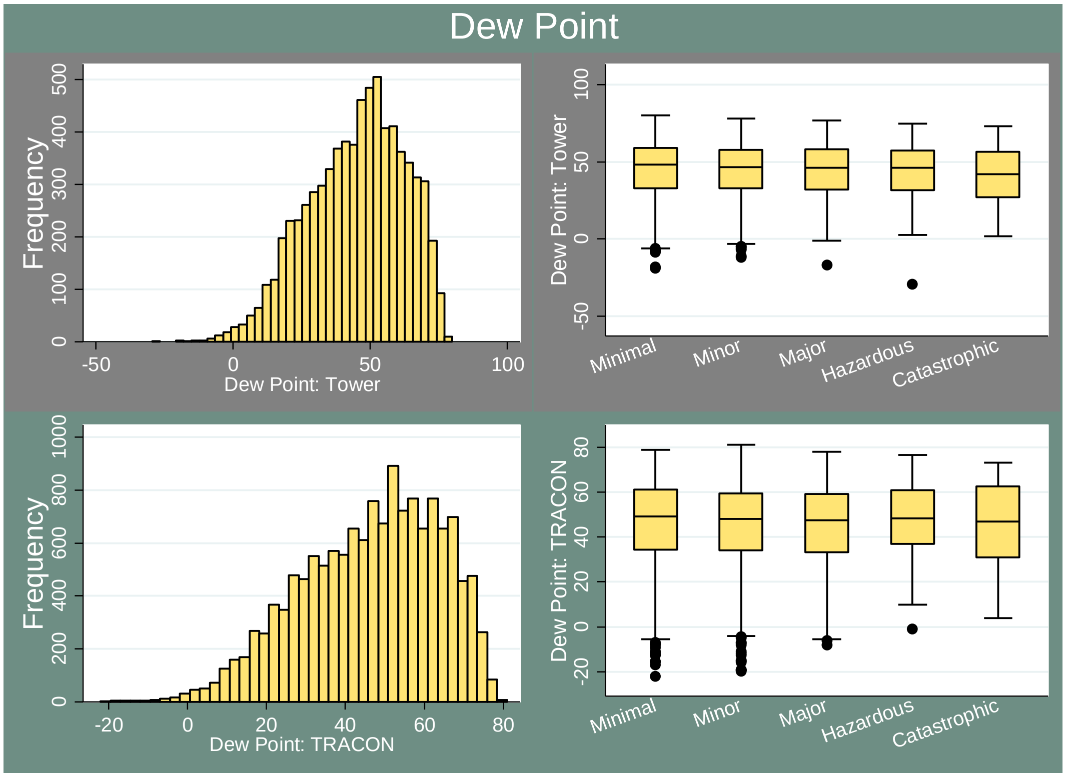

Figure 58 - Distribution of Dew Point by Facility Type

The overall distribution appears similar for both facility types, with both exhibiting a slight leftward skew. Median dew point values across severity levels also appear relatively similar. A Kruskal-Wallis test confirms these results, indicating that these severity levels do not vary jointly across categories for either facility type.

6.6.2. Temperature by Facility Type

The temperature at the time of the incident is interpolated from the closest hourly readings. Figure 59 presents the overall distribution of temperature and the distribution by severity classification by facility type.

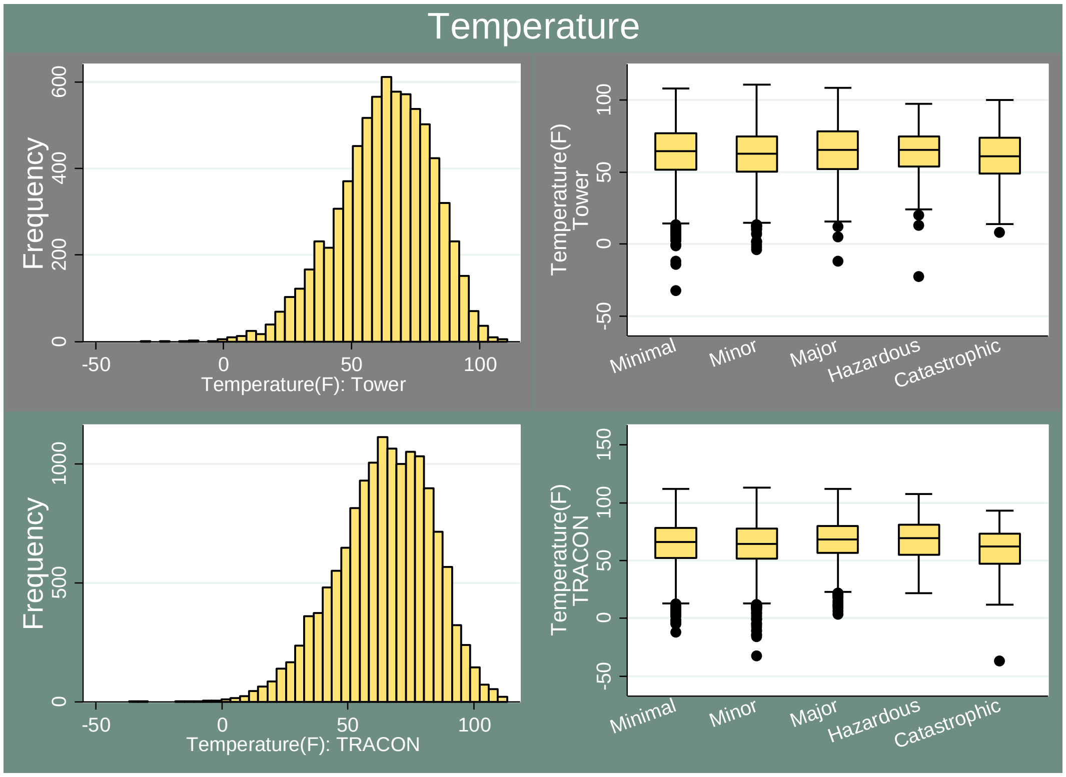

Figure 59 - Distribution of Temperature by Facility Type

The overall distribution is relatively unsurprising and similar between the two facility types. Since this data covers all 50 states over the course of several (5) years, an approximately normal distribution is to be expected. There also seems to be no clear difference across temperature and severity levels. Median temperature levels seem to vary only slightly across severity categories. A Kruskal-Wallis test indicates that the severity levels are indeed jointly different for both facility types and Table 77 and Table 78 present the pairwise comparisons.

Table 77 - Kruskal-Wallis Test of Temperature for Tower Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major		X			
Hazardous					
Catastrophic					

Table 78 - Kruskal-Wallis Test of Temperature for TRACON Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major	X	X			
Hazardous		X			
Catastrophic			X	X	

6.6.3. Temperature Dew Point Difference by Facility Type

The final variable that examines temperature is the difference between temperature and dew point. When the dew point and temperature are closer, fog and precipitation are more likely.[67] Figure 60 details the distribution of the temperature dew point difference by facility type.

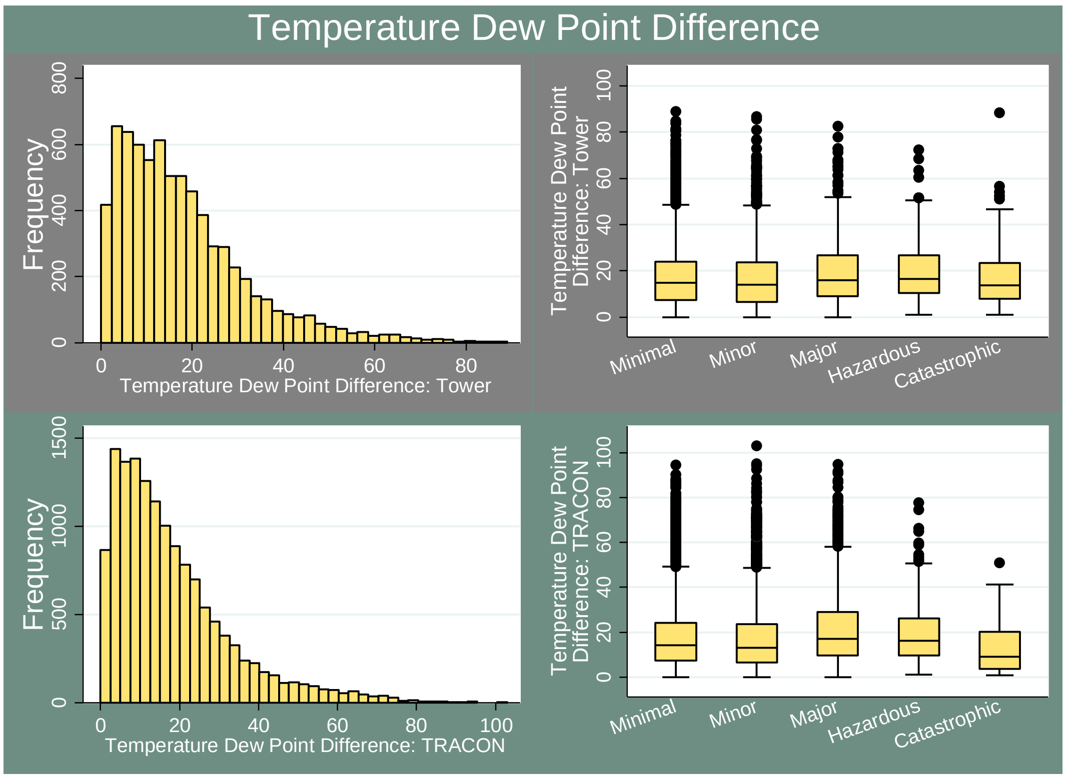

Figure 60 - Distribution of Temperature Dew Point Difference by Facility Type

67 Ibid.

As expected, there are no negative values due to the inherent relationship between dew point and temperature. There also appears to be many cases where the difference between temperature and dew point can be quite large. This effect appears to taper off around the 20 degree mark for both facility types. There looks to be slight variation in the median values across severity levels, with Major and Hazardous categories having slightly higher median values for both facility types. A Kruskal-Wallis test provides some further insight, determining that these categories do jointly vary for both facility types. Table 79 and Table 80 present the pairwise comparisons.

Table 79 - Kruskal-Wallis Test of Temperature-Dew Point Difference for Tower Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor					
Major	X	X			
Hazardous		X			
Catastrophic					

Table 80 - Kruskal-Wallis Test of Temperature-Dew Point Difference for TRACON Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major	X	X			
Hazardous		X			
Catastrophic	X		X	X	

It is worth noting that for TRACON events, the median value for Catastrophic appears to be both statistically different and much lower than the other categories. This suggests that there could be some relationship between fog/precipitation and highly severe/Catastrophic incidents. Further examination with a full weather (multinomial logit) model is warranted to better tease out this possible relationship.

6.6.4. Sea Level Pressure Deviation by Facility Type

This variable indicates the air pressure at the time of the incident, normalized to sea pressure. Pressure varies with altitude, thus it is important to normalize to a standard altitude (in this case, sea level). Thus, it is most helpful to examine this variable in terms of deviation from standard pressure (1013.25 mb). Figure 61 presents this distribution.

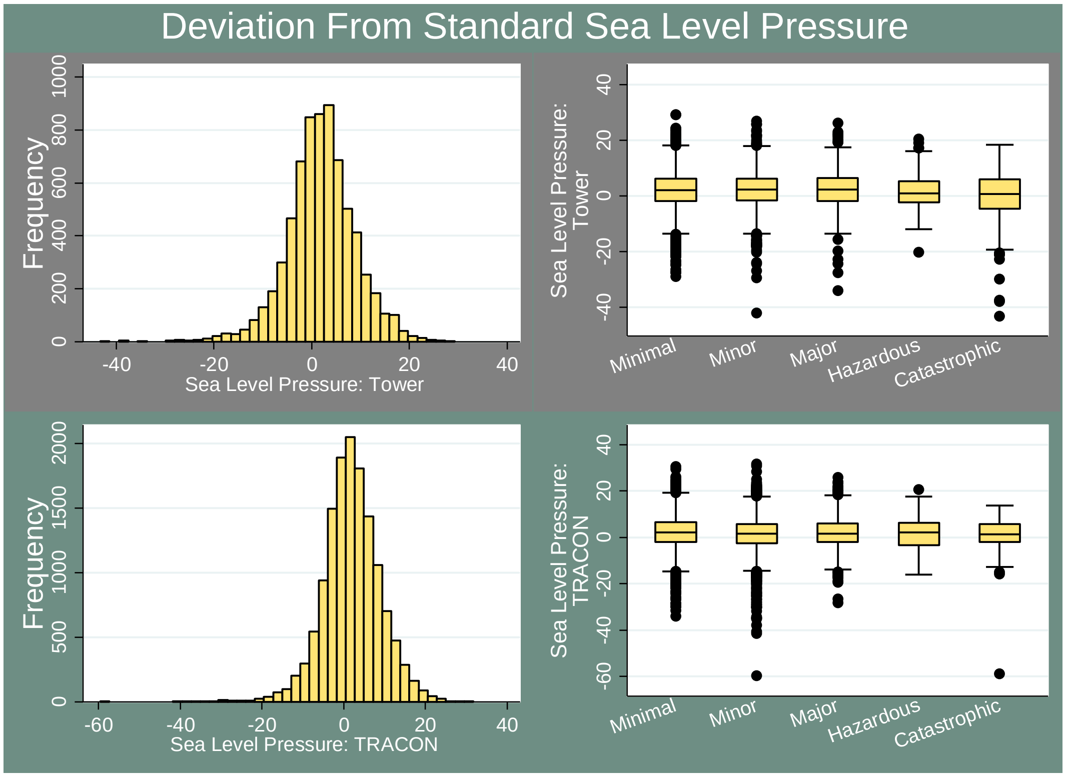

Figure 61 - Distribution of Deviation from Normal Sea Level Pressure by Facility

A Kruskal-Wallis test was conducted, indicating that the categories for tower incidents could not be determined to vary jointly. TRACON incidents did vary jointly, but there was little variation between categories, with only 'Minimal-Minor' varying significantly.

6.6.5. Weather Complexity Factor by Facility Type

The weather complexity factor variable originates from the ATSAP form, where the respondent may indicate if the weather was a significant factor (or not) to the loss of separation incident. Figure 62 presents the distribution of this variable across severity categories.

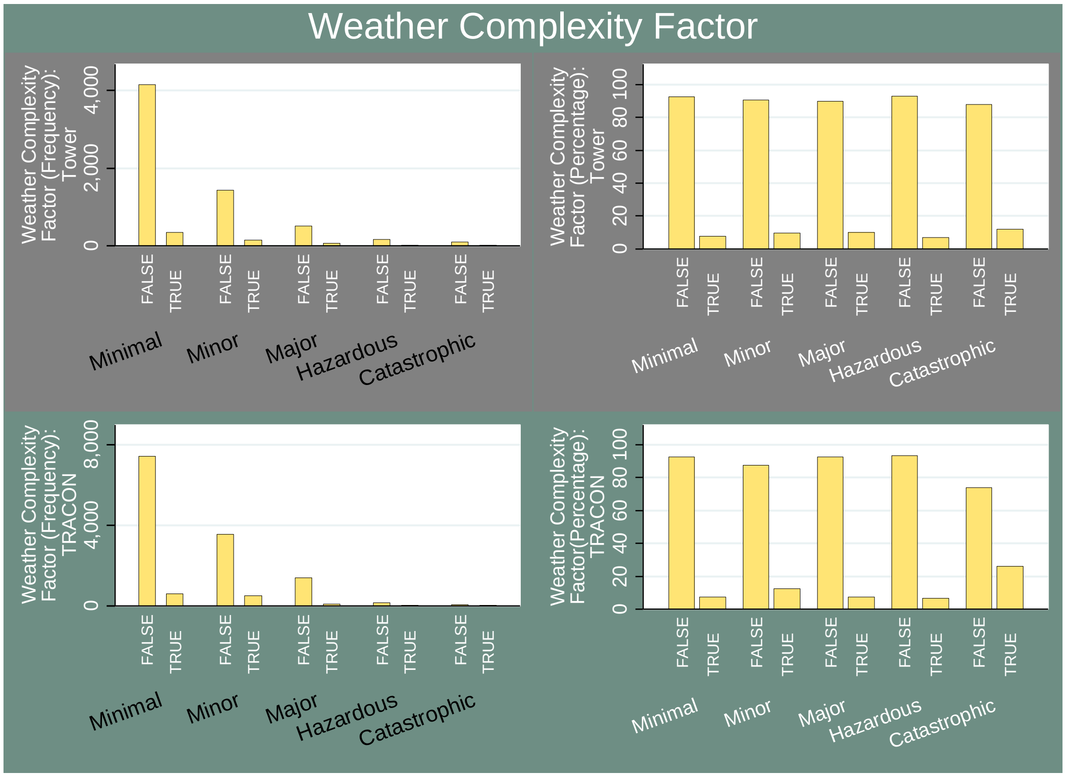

Figure 62 - Distribution of Weather Complexity Factor by Facility Type

Table 81 presents a single variable logit of the weather complexity factor variable on severity. There appears to be no relationships between weather complexity and severity for tower or TRACON incidents.

Table 81 - Logit Estimate of Weather Complexity Factor by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Weather Complexity Factor- Tower	1.227	0.152	6932
Weather Complexity Factor- TRACON	0.891	0.016	13,859
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.6.6. Weather Phenomena by Facility Type

This variable details the type of weather phenomena (provided by the METAR) at the measured time. The bulk of the categories encompass phenomena related to precipitation, in addition to other non-precipitation categories such as fog, smoke, and haze. Figure 63 presents the weather phenomena at the time of the incident by facility type, while Figure 64 presents the weather phenomena excluding the ‘no weather’ category. Weather phenomena data is presented this way due to the high number of no weather incidents. After removing ‘no weather’, the top categories are ‘haze’, ‘rain-light’, ‘rain-moderate’, and ‘snow-light’.

In order to get a sense of how weather phenomena is distributed across incidents and severity levels, all weather phenomena excluding ‘no weather’ categories were grouped together. This creates a binary variable indicating whether there was any type of phenomena during the time of the incident. Figure 65 presents the overall distribution of this variable.

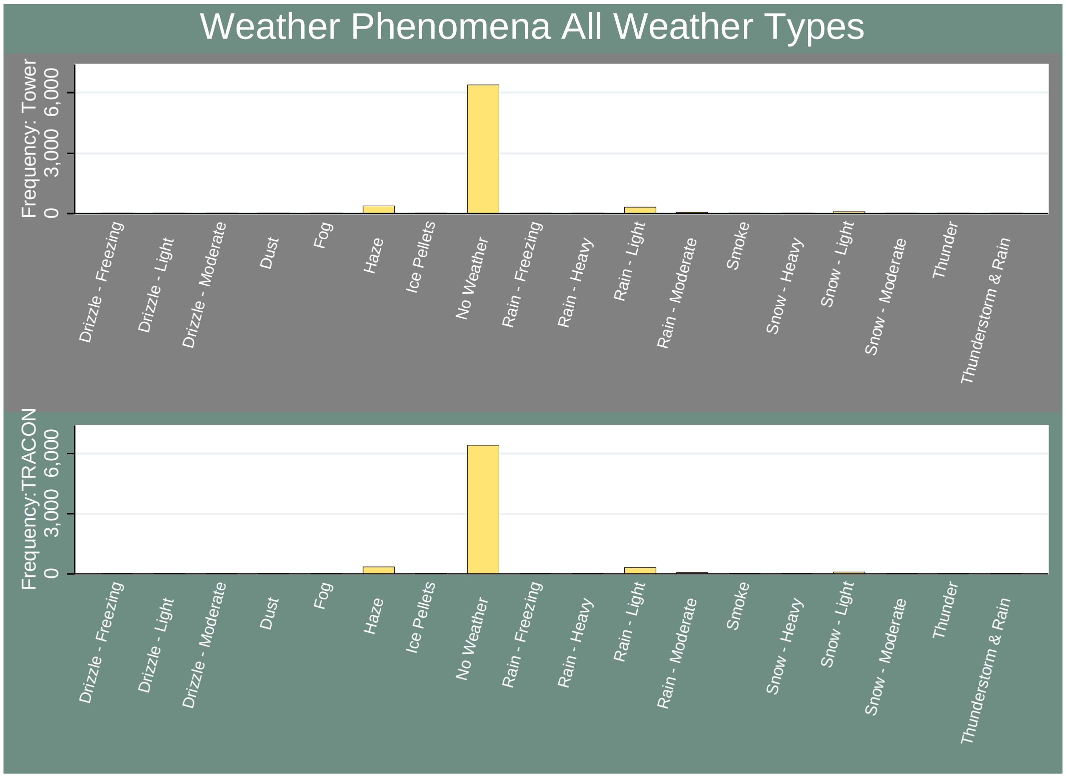

Figure 63 - Distribution of All Weather Phenomena Events by Facility Type

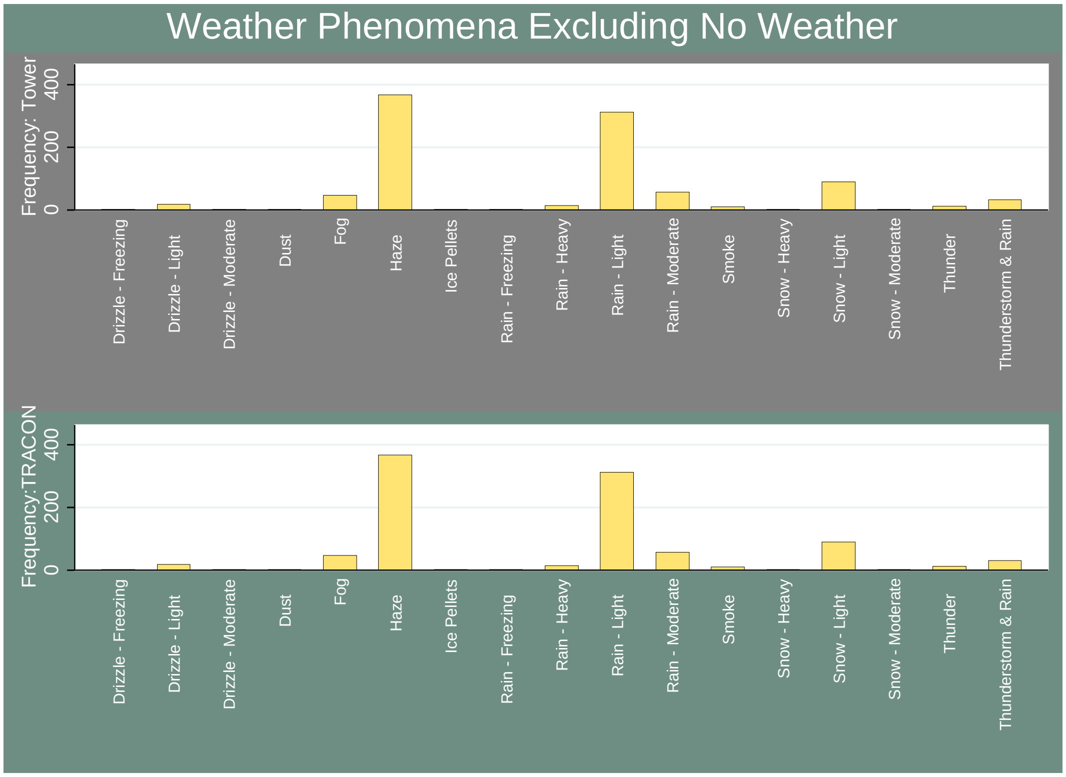

Figure 64 - Distribution of All Weather Phenomena Events Excluding No Weather by Facility Type

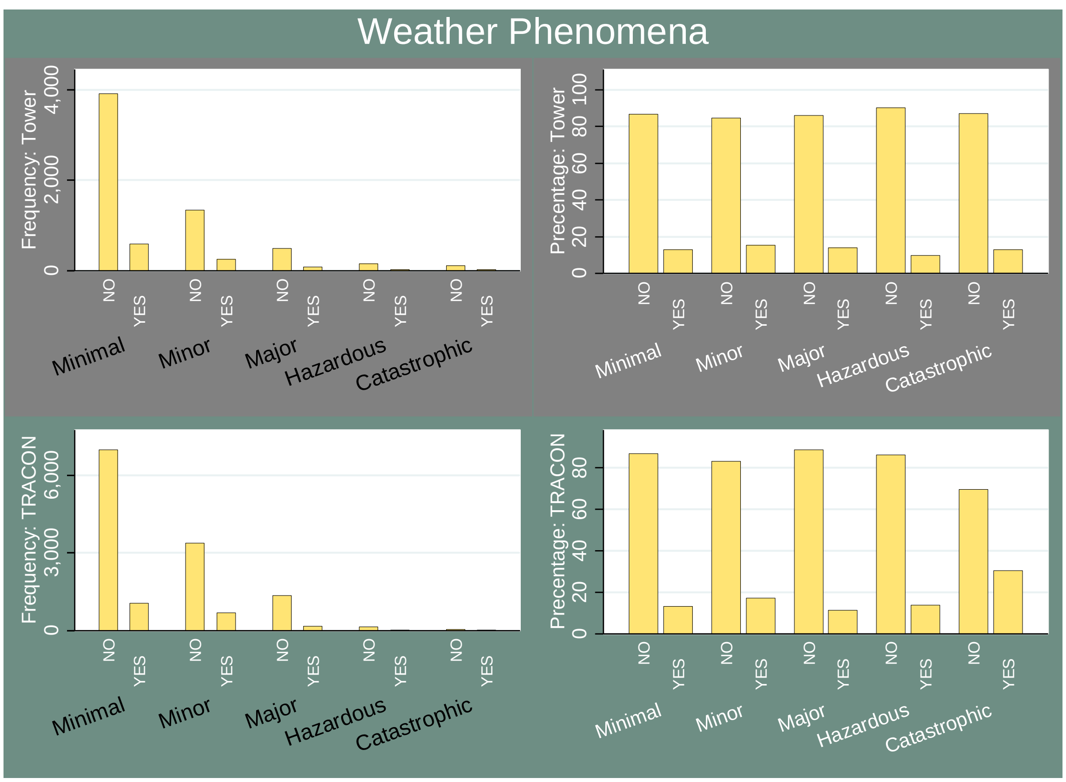

Figure 65 - Distribution Weather Phenomena by Facility Type

The overall distribution appears similar (percentage wise) across severity categories for tower incidents. There is a similar pattern for TRACON incidents, expect for a large increase in weather phenomena events in Catastrophic incidents. To gain more insight into this variable's relationship with severity a single logit on severity was estimated (Table 82 presents the logit estimation). High p-values for both facility types indicate that there is no apparent relationship between the weather phenomena indicator and severity. This suggests that a more comprehensive MNL model is needed to better understand the relationship between Catastrophic incidents and weather phenomena (see section 6.6.12.2).

Table 82 - Logit Estimate of Weather Phenomena by Facility Type

Variable	Odds Ratio	Standard Error	Obs
Weather Phenomena Indicator- Tower	0.945	0.094	6932
Weather Phenomena Indicator- TRACON	0.836	0.096	13,859
Significance Levels: *p<0.05, ** p<0.01, ***p<0.001			

6.6.7. Wind Speed by Facility Type

The wind speed variable is measured by the METAR at the time of the incident (in knots). Figure 66 presents the distribution of this variable by facility type.

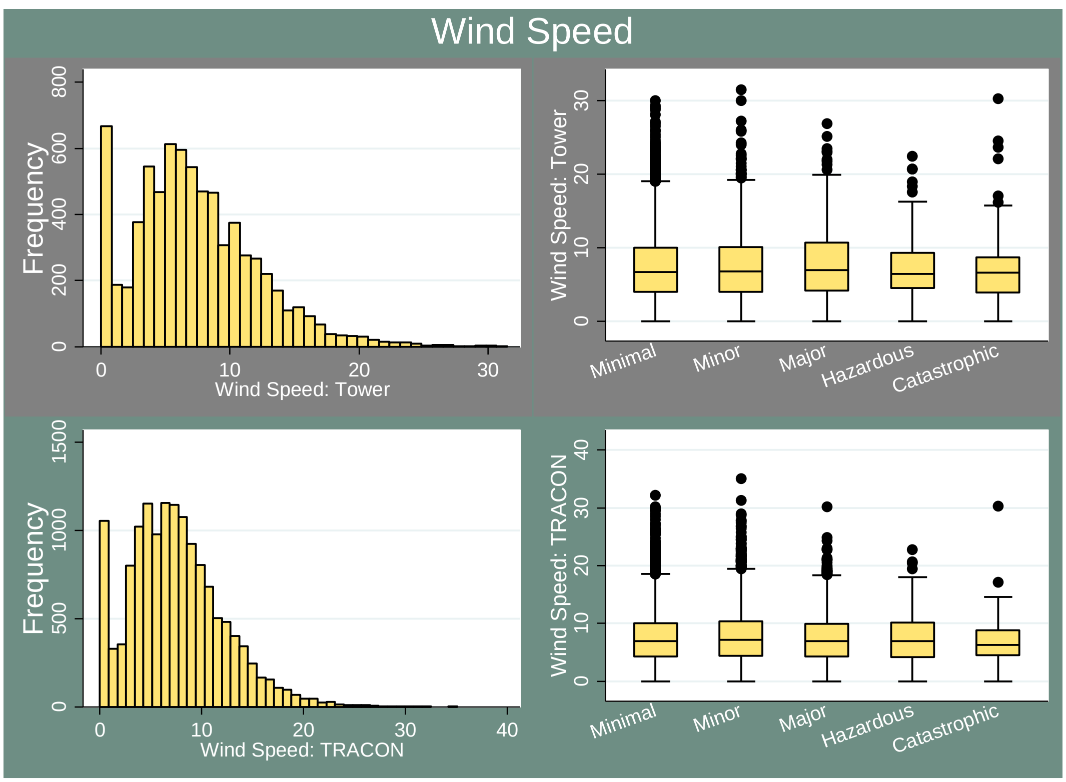

Figure 66 - Distribution of Wind Speed by Facility Type

In both facility types there is clustering around wind speeds of 0 knots, with the peak occurring in between 0 and 10 knots. Wind speeds then quickly taper off to the left to around 30 knots. There is little (if any) variation in the median wind speed values across severity levels. A Kruskal-Wallis test could not jointly distinguish between groups for either tower or TRACON incidents.

6.6.8. Precipitations Last 6 Hours by Facility Type

This variable indicates (in inches) the total precipitations amount in the last 6 hours of the event. This variable should not be confused with the amount of precipitation during the time of the incident. It could be the case that an incident occurred during a period of no precipitation with relatively clear weather, but precipitation did occur 5 hours previously. This creates a caveat that should not be ignored when examining this variable in detail. Figure 67 presents the distribution of precipitation over the past 6 hours.

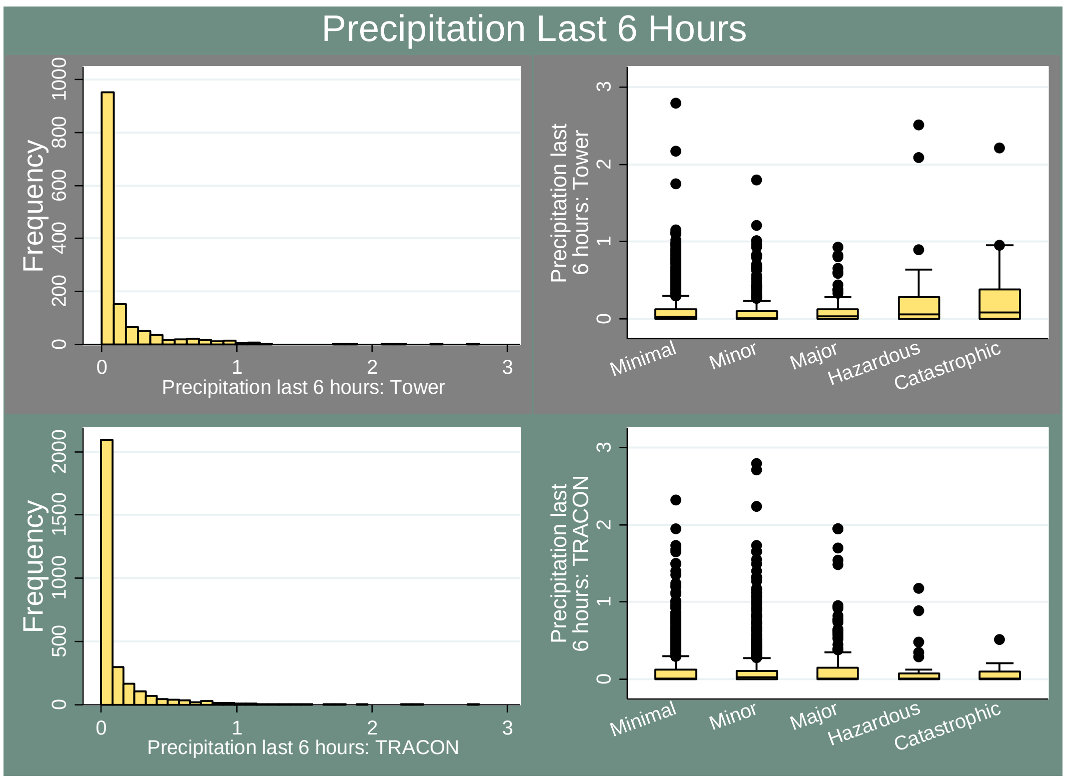

Figure 67 - Distribution of Precipitation the Last 6 Hours by Facility Type

The overall distribution clearly shows groupings around 0 inches of precipitation for both facility types, with a rightward tail of incidents with some amount of precipitation tapering off after 1 inch. The distribution across severity levels is inclusive for both facility types due to the large amount of clustering around 0 inches. A Kruskal-Wallis test indicates that the severity categories cannot be determined to jointly vary for both facility types. Again this result is not surprising given the large amount of clustering. Regardless, any amount of inference to how precipitation over the past 6 hours relates to severity should be taken with a grain of salt due how this variable was constructed.

6.6.9. Cloud Ceiling by Facility Type

This variable measures the height of the cloud ceiling at the time of the incident. It was interpolated in a similar fashion to the other weather variables. Figure 68 presents the distribution of this variable.

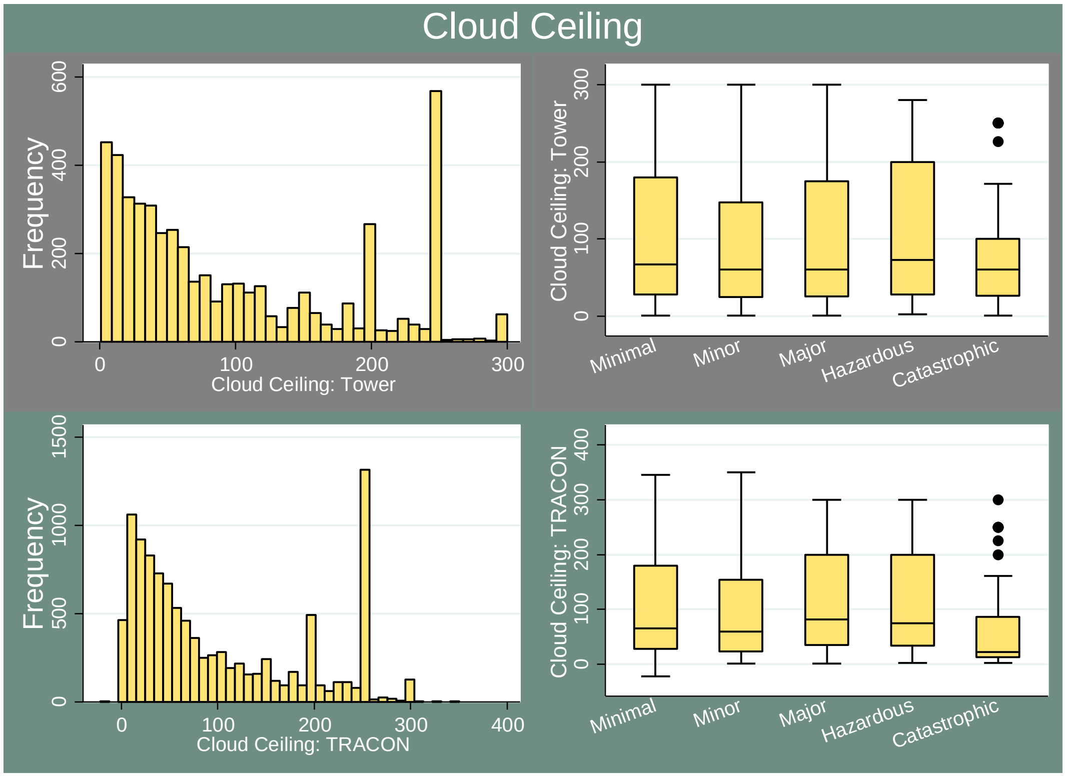

Figure 68 - Distribution of Cloud Ceiling by Facility Type

The first takeaway from the overall distribution is the large spikes at certain values (namely 0, 150, 200, 250, etc.). This is largely due to potential rounding by METAR stations. There is a little variation between median cloud ceiling values across severity levels for tower events, while there does appear to be some variation in TRACON incidents. A Kruskal-Wallis test could not jointly tell the difference between severity categories for tower incidents but could for TRACON incidents. Table 83 presents the pairwise comparison for TRACON incidents.

Table 83 - Kruskal-Wallis Test of Cloud Ceiling for TRACON Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major	X	X			
Hazardous					
Catastrophic	X	X	X	X	

6.6.10. Cloud Coverage by Facility Type

This variable indicates how much of the sky was covered with clouds by facility type. The original rating is presented as a series of increasing fractions from Clear (0/8ths of the sky covered) to Overcast (8/8ths of the sky covered). Due to the sequential nature of these categories (and their approximation to fractions), it was decided to turn this variable into a numeric variable describing how many eighths of the sky is covered. Thus, the variable ranges from 0 to 8. Table 84 presents the mapping from the original categories to the numeric values. As the original categories covered a range of values, the midpoint of each range was used.[68]

Table 84 - Mapping of Cloud Coverage Categories to Numeric Values

Original Category	Numeric Value
Clear (0/8)	0
Few (between 0/8 and 2/8)	1
Scattered (between 2/8 and 4/8)	3
Broken (between 5/8 and 7/8)	6
Overcast (8/8)	8

After conversions to a 0 to 8 scale, values were interpolated between the two points and then rounded. To avoid recoding values with a degree of inaccurate precision, the data were rounded to the nearest half. The final data measures the number of eighths of the sky covered from 0 to 8, measured in steps of 0.5. While the units may seem odd, the variable can still be interpreted as the fraction of the sky (in eights) covered with clouds.

[68] The categories presented in Table 84 present an interesting problem. First, the categories are of differing widths. Clear and Overcast only cover one value while Few, Scattered, and Broken represent ranges. Additionally, some categories overlap, while others are adjacent. Clear indicates 0/8 parts of the sky is covered. The next category, Few, indicates that between 0 and 2 out of 8 parts are covered. This category picks up exactly where clear left off. Scattered begins at 2 where Few left off and ends at 4. Broken, however, begins at 5 – one unit more than where Scattered ends. Overall this is likely a minor quirk in the definition, but it may create artifacts in the data and ends up making the top part of the scale more spaced out than the bottom half.

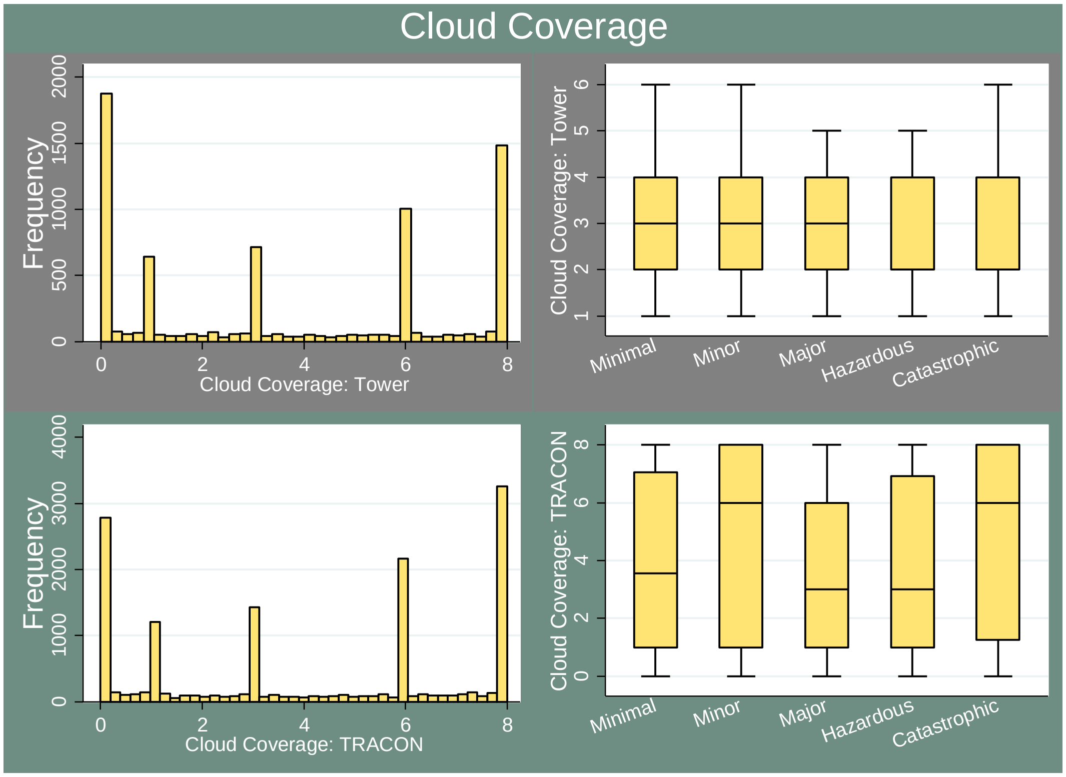

Figure 69 - Distribution of Cloud Coverage by Facility Type, Rounded

Figure 69 shows clear peaks in the overall distrubtion due to the rounding procedure just discussed. Note that in addition to the rounding to the nearest half, there are also distinct spikes are certain values – such as 1, 3, 6, and 8. These values are the midpoints of the original categories, as indicated in Table 85. There are still a fair amount of observations in between these values (generated by interpolation), but this clustering is important to be aware of when considering the impact this variable may have on severity.

The distribution across severity levels is relatively flat in terms of median values for tower events, with the exception of Hazardous and Catastrophic incidents. The distribution across severity levels for TRACON incidents are mixed with the median level swinging in different directions across severity levels. A Kruskal-Wallis test indicates that the categories are not jointly different from each other for tower incidents but are for TRACON. Table 85 presents the pairwise comparison for TRACON facilities.

Table 85 - Kruskal-Wallis Test of Cloud Coverage for TRACON Facilities: Multiple Comparisons Between Groups

	Minimal	Minor	Major	Hazardous	Catastrophic
Minimal					
Minor	X				
Major	X	X			
Hazardous					
Catastrophic					

6.6.11. Visibility at Less than 10 Miles by Facility Type

While the previous two variables dealt with visibility indirectly, this variable measures visibility directly. This variable measures the distance one can see (approximately) in miles at less than 10 miles of visibility. Less than 10 miles was used as a cut off due to the high number of 10 mile visibility observations, which essentially means perfect or unlimited visibility. Using less than 10 miles will allow for a clearer analysis of the distribution of the variable. The distribution is presented in Figure 70.

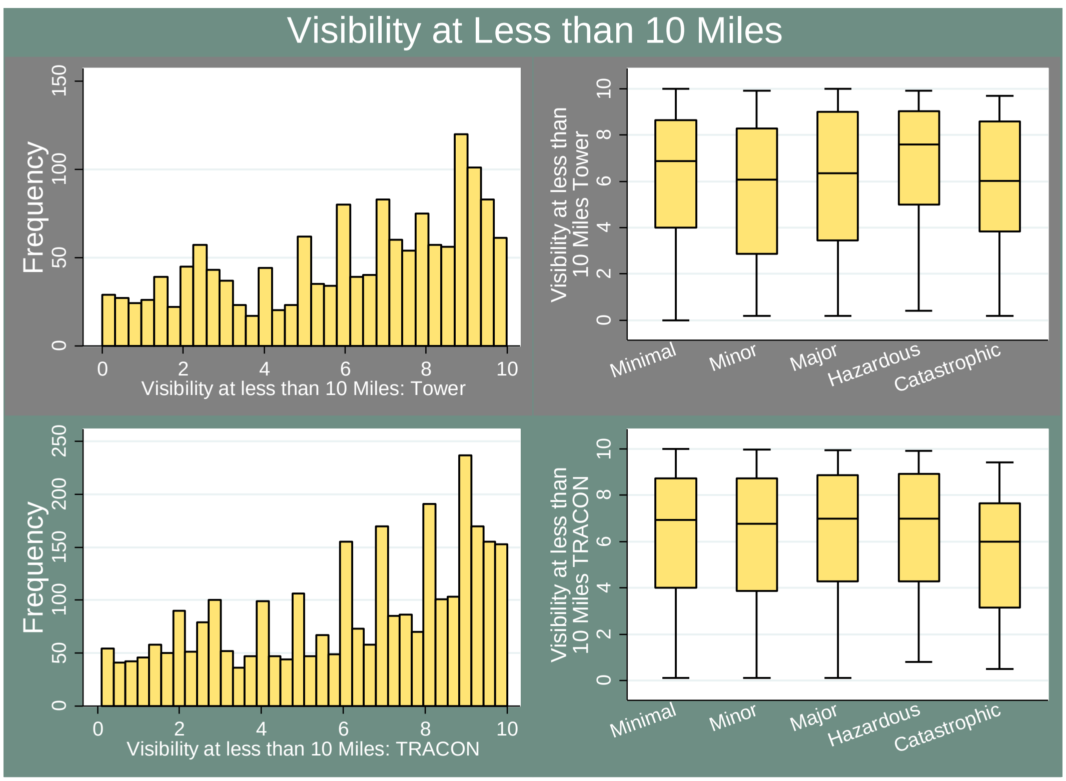

Figure 70 - Distribution of Visibility at Less than 10 Miles by Facility Type

For both facility types, there is an upward trend in visibility, with groupings around whole values (2, 3, 4 miles etc.). This upward trend is not surprising, given that very low visibility is often correlated to less aircraft traffic. This could be due to a change in flight rules (from visual to instrument, grounding most GA aircraft) or aircraft simply being forced to stay on the ground.

The distribution across severity levels is mixed for tower incidents and relatively flat for TRACON incidents. A Kruskal-Wallis indicates that the categories cannot be determined to jointly vary for both facility types. It is unclear what (if any) relationship between severity and visibility exists and it is improper to infer too much without properly controlling for other weather phenomena first.

6.6.12. Weather Model

The following models contain weather variables examined in the previous section. This section lends itself well to the inclusion of interactions between certain weather variables that are closely related. Additionally, the daily operations variable will be used as control variable in this section as a measure of overall traffic volume. There are important underlining relationships between operations and weather variables that need to be controlled for. In other words, it is expected that bad weather days will be associated with less overall traffic and higher severity incidents, and omitting daily operations may place some upward bias on the estimated coefficients of the weather variables.

As a final note, the variables on precipitation, and visibility less than 10 miles will not be used due to both data quality issues and lack of observations. The variable for cloud coverage will be used instead of cloud ceiling, and the dew point temperature difference will be used in place of both temperature and dew point.

6.6.12.1. Tower Multinomial Logit Model

Table 86 presents the multinomial logit estimation of key weather variables on severity categories for tower facilities. Minimal severity level incidents served as the base for the multinomial logit. A low number of weather variables held any statistically significant relationship with severity category variables, with only the temperature dew point difference variable increasing the likelihood of an incident falling in the Major severity category. Total daily operations are significant and have a similar relationship with the severity categories as observed in the facility model.

One plausible reason for the lack of significance in these weather variables is the fact that tower facilities generally have high quality weather data and forecasts, which allow ATC at tower facilities to limit operations or even ground aircraft depending on the type and scale of adverse weather. In this sense, weather conditions would have no significant relationship with severity at tower facilities if aircraft were being appropriately managed (separated) during periods of poor or extreme weather.

Table 86 - Multinomial Logit Estimate of Weather Variables for Tower Incidents

	Minor	Major	Hazardous	Catastrophic
Cloud Coverage	1 (.014)	1 (.021)	.96 (.036)	.98 (.036)
Cloud Coverage X Sea Level Pressure	1 (.0015)	1 (.0023)	1 (.0037)	.99 (.0036)
Daily Operations	1.2*** (.07)	1.1 (.069)	1.1 (.094)	.59*** (.066)
Deviation from Standard Sea Level Pressure	1 (.0083)	1 (.013)	.98 (.019)	.96 (.027)
Temperature Dew Point Difference	1 (.0028)	1*** (.0042)	1 (.0061)	1 (.0081)
Weather Complexity Factor	1 (.16)	1.3 (.31)	.97 (.37)	1.8 (.72)
Weather Indicator	1.1 (.14)	1.2 (.18)	.84 (.3)	.85 (.34)
Weather Complexity X Weather Indicator	1.2 (.26)	1.1 (.37)	1.1 (.68)	.69 (.49)
Wind speed (knots)	.99 (.0082)	1 (.011)	1 (.017)	.97 (.021)
Year 2008 Indicator	.57 (.24)	.38 (.22)	1.6e-06*** (1.8e-06)	7.3e-07*** (5.0e-07)
Year 2009 Indicator	1 (.37)	.5 (.19)	.89 (.95)	.28 (.18)
Year 2010 Indicator	1.4 (.48)	.83 (.31)	2.4 (2.4)	.44 (.26)
Year 2011 Indicator	1.7 (.57)	.94 (.34)	2.8 (2.9)	.48 (.3)
Year 2012 Indicator	1 (.32)	.89 (.31)	1.8 (1.8)	.39 (.24)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 **N = 6445**				

6.6.12.2. TRACON Multinomial Logit Model

Table 87 presents the multinomial logit results for the TRACON facility. Again, the Minimal severity level was used as the base category. The TRACON weather model has quite a bit more action in terms of statistically significant weather variables opposed to the tower weather model. The temperature dew point difference plays a significant role in increasing the likelihoods of being in a Major or Hazardous category while decreasing the likelihood of being in the Catastrophic category. There appears to be some statistical significance between sea level pressure, cloud coverage, and the interaction term and a Wald joint test of significance rejects the null hypothesis of no relationship with severity. The weather complexity factor variable is also significant in Minor and Major categories, increasing the likelihood of Minor incident and decreasing the likelihood of a Major.

Table 87 - Multinomial Logit Estimate of Weather Variables for TRACON Incidents

	Minor	Major	Hazardous	Catastrophic
Cloud Coverage	1** (.014)	.99 (.021)	.98 (.036)	.93 (.046)
Cloud Coverage X Sea Level Pressure	1 (.0011)	1 (.0017)	1 (.0037)	1** (.0049)
Daily Operations	1.3*** (.062)	1.5*** (.1)	1.4*** (.12)	.72* (.1)
Deviation from Standard Sea Level Pressure	.99 (.0057)	.99 (.01)	.98 (.022)	.92* (.031)
Temperature Dew Point Difference	1 (.0048)	1*** (.0047)	1 (.0053)	.98 (.015)
Weather Complexity Factor	1.6*** (.16)	.99 (.14)	.67 (.32)	3.7** (1.7)
Weather Indicator	1.2* (.1)	1.1 (.12)	1.3 (.26)	1.8 (.97)
Weather Complexity X Weather Indicator	.74 (.12)	1.3 (.31)	1.3 (.9)	1 (.64)
Wind speed (knots)	1 (.0063)	.99 (.0091)	.99 (.015)	1 (.029)
Year 2008 Indicator	.66 (.2)	.37 (.32)	.75 (.9)	.78 (.53)
Year 2009 Indicator	1.6 (.45)	.67 (.26)	1.4 (1.5)	418392*** (165288)
Year 2010 Indicator	2.5*** (.66)	1.4 (.51)	4.1 (4.3)	498603*** (154906)
Year 2011 Indicator	3.1*** (.83)	1.7 (.64)	3.4 (3.6)	611432*** (217739)
Year 2012 Indicator	1.7* (.37)	1 (.35)	1.8 (1.9)	516238*** (155380)
SEs in parentheses *p<0.05, ** p<0.01, ***p<0.001 N = 13,239				

Figure 71 presents the impact on probability of severity categories of the temperature dew point difference. This gives a visual representation of the change in probability of falling in a certain severity category as the temperature dew point difference increases. Remember as the temperature dew point difference increases, the likelihood of having fog or precipitation decreases. This means that potentially wetter or less visibility weather conditions are associated with less severe incidents. This finding suggests that during periods of potential precipitation or low visibility, aircraft are most likely being managed in such a way that lowers the chance of a severe event. One obvious example would be increased separation rules for aircraft during periods of poor visibility or heavy precipitation.

What is most striking is the comparison between the two facility types. There is essentially no statistical relationship between any of the weather variables and the severity categories in tower incidents, but there does appear to be some relationship between the two in TRACON incidents. As noted previously, tower and TRACON airspaces are inherently different and need to be treated separately when modelled, especially in light of how the data are collected. Tower facility airspaces are not necessarily subjected to less variable weather conditions but they clearly benefit from a constant stream of detailed localized weather information. Towers also have the advantage of being able to ground aircraft if weather conditions become too hazardous to operate in. TRACONs on the other hand must deal with variable weather conditions as aircraft pass through their relatively larger airspace before being handed off to tower control.

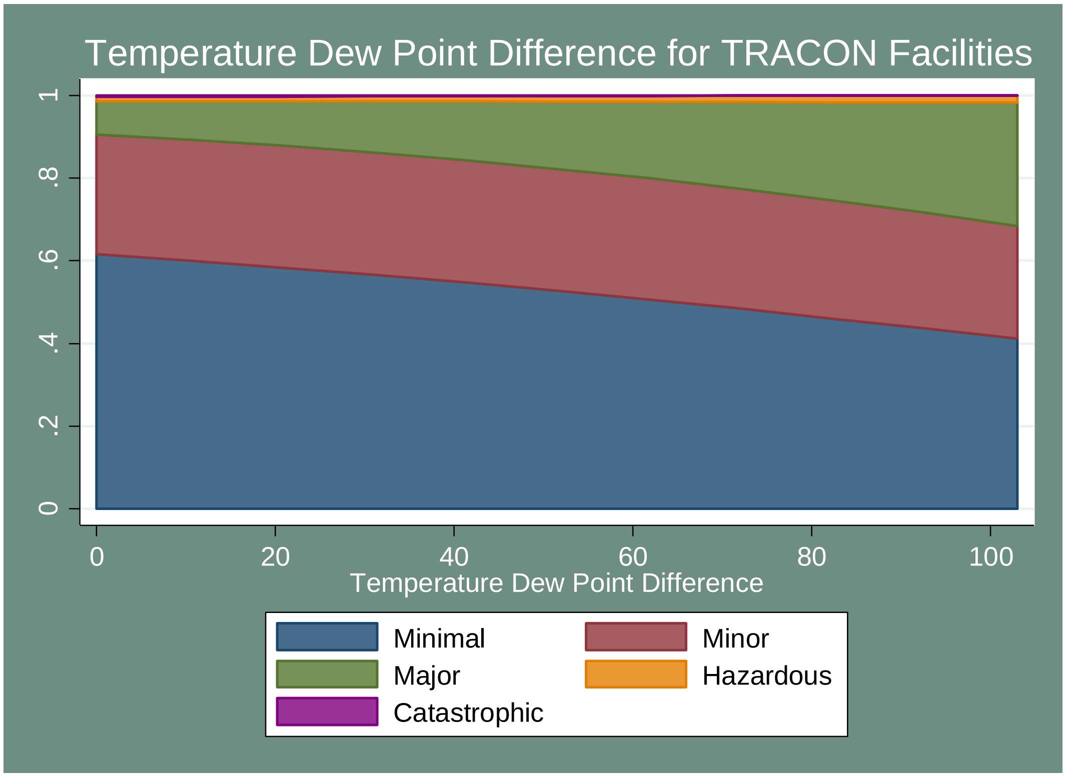

Figure 71- Impact on Probability of Severity Categories of Temperature Dew Point Difference for TRACON Incidents

7. BOUILLABAISSE MODEL

The "Bouillabaisse" Models pull the significant variables from the preceding models into comprehensive models, for Tower and TRACON facilities. It is intended to provide a quick overview of many different variables in one place, but is not intended to be a definitive index of the relationship between each individual variable and severity. It is recommended that readers who are interested in exploring any given variable in more depth reference the more focused models in Section 6. of this report.

7.1. Bouillabaisse Caveats

The bouillabaisse model intended as a preliminary snapshot of how the variables in ATSAP are associated with airborne incident severity. These models are intended to serve as a starting point for determining, in collaboration with the FAA, which variables ought to be looked at more closely, and as a prompt for discussions about the underlying reasons for possible relationships between variables and severity.

*At this point, these Bouillabaisse Models are specifically **not** intended to be:*

- *A predictive index of how many incidents there will be in any given category*
- *Guidelines for policy decisions*

If it is determined that FAA has interested in researching the relationships between incident severity and ATSAP variables further and potentially implementing policy actions, these models will need to undergo substantial further refinement, based on input from FAA. The final model would need to incorporate feedback from FAA on the exact nature of how any one variable could contribute to an incident, and on how the ATSAP variables are interrelated.

At present, there are several statistical problems that may be present in these models; these issues include but are not limited to:

- Limited Sample Period: The models presented below focus exclusively on the 2011-2013 time period, because many of the causal factor variables began being reported in ATSAP in 2011. Since the models do not incorporate all of the available observations the estimates are less robust than they would be if the full time series would be used. If FAA is considering making policy recommendations based on these models, it is important that they be compared with a new set of models that use the full set of available observations to ensure that the estimated coefficients are stable.
- Overfit: Models with a large number of variables are prone to overfit, where the model interprets random noise as a correlation.
- Multicollinearity: Many of the variables in the ATSAP database are correlated with each-other. Multicollinearity reduces the accuracy of estimates for individual coefficients and results in large standard errors.
- Inconsistent Regression results: The problems of a limited sample period, overfit, and multicollinearity in these models mean that it is difficult for the model to converge. Thus, regression results are very slightly (inconsequentially) different each time the model is run.

7.2. Model Specifications

The Bouillabaisse models use a multinomial logit structure, following the structure of the preceding models in this report. In this section the significant variables from the individual models are pulled together into one full model. For the categorical variables one category is omitted from each group, to serve as a basis for comparisons and to prevent a singular matrix.

The base case for each set of variables matches the base cases in the earlier models, and in most cases, is the most frequently observed category in the ATSAP database, with the goal of making comparisons to the most "normal" situation.

The base case for categorical variables consists of:

- Single-Aisle Jet
- Control Status: Instrumental Flight Rules
- Phase of Flight: Departure
- Terminal Radar Approach Control: Departure

There was one variable change for the Bouillabaisse: military and civilian helicopters were combined, since the coefficients were very similar and there were not a large number of observations.

The Tower and TRACON models have slightly different variables. This is because some variables in ATSAP are closely correlated with each other, and thus cause problems when they are in the same model together. In these cases, Volpe chose the set of variables which yielded the most robust statistical results; these choices differed between the Tower and TRACON models.

7.3. Model Results

Many of the model results in the bouillabaisse model echo what was found in the focused models earlier in this report. These models can be viewed below, in Table 88 (Tower) and Table 89 (TRACON). Particularly interesting results are pointed out in the text below.

7.3.1. Aircraft Type

Experimental Planes are associated with Catastrophic Incidents. Emergency Situations are associated with increased severity, while Traffic Management Initiatives are associated with low severity incidents in Tower facilities.

In both the Tower and TRACON models, incidents involving experimental aircraft are much more likely to be catastrophic than incidents involving other types of aircraft. These results also appeared strongly in the data exploration and aircraft model sections of this report, which shows that it is robust across different model specifications.

Single engine prop planes were also associated with increased severity (Major, Hazardous and Catastrophic in Tower facilities, and Catastrophic in TRACON facilities.) Possible reasons for this include that experimental and single engine prop planes may be often flown by hobbyists, rather than professional pilots, and that they may more frequently be located in smaller airports. This is a topic that merits further research.

In Tower facilities, emergency situations were associated with a high frequency of catastrophic incidents. This result repeats what was found in the earlier focused models. It makes sense, since an emergency situation by definition involves some sort of unplanned problematic event.

Incidents involving Traffic Management Initiatives typically have a low level of severity in Tower facilities. This provides evidence for their efficacy, since their purpose is to reduce traffic when circumstances (such as weather) are less than ideal.

7.3.2. Facility Characteristics:

Higher ATC levels are associated with a lower likelihood of Catastrophic incidents for Tower facilities.

The Bouillabaisse results for facility characteristics did not show the same level of extreme contrasts that the aircraft type variables show. The only notable result was that in Tower facilities, higher ATC levels were associated with a reduced frequency of incidents being labeled with catastrophic severity. This result did not carry over to TRACON facilities.

7.3.3. Controller Variables:

Controller Experience has no significant relationship with severity.

Controller experience had no statistically significant relationship with severity in the bouillabaisse model, in both Tower and TRACON facilities. This result mirrors the findings in the Runway Incursions Report.

This result, however, varies depending on how the model is specified. Outside of the bouillabaisse model, in the binary models that looked exclusively at controller experience without controlling for other variables, controller experience was associated with lower odds of an incident being severe in TRACON, but not tower, facilities. In the multinomial controller model that included the full time series but only a subset of the variables, more controller experience is associated with a decreased risk of minor and major incidents in TRACON facilities, but again there is no significant result in tower facilities. There was no significant relationship between controller experience and severity in the controller model that included the variables introduced in 2011. The fact that these results vary depending on the specification means that they must be viewed with caution.

As with all results in this report, it is important that to keep in mind that Volpe is only analyzing the distribution of severity rather than the frequency of incidents, so it is possible that controller experience is tied to how often incidents occur.

7.3.4. Pilot Actions

Untimely Rolls are associated with Catastrophic Incidents in Tower Facilities.

In Tower Facilities, incidents with untimely rolls are 7.4 times more likely to be catastrophic than the typical incident. It may be that pilots who are trying to avoid an incident choose to roll only when they are too close to another object to perform any other maneuvers. This is an area that merits further research by FAA.

7.3.5. Weather Characteristics:

Higher air pressure is associated with reduced incident severity in Tower and TRACON facilities. The magnitude of this effect is small, but statistically significant. It is possible that reduced visibility during low pressure weather results in aircraft getting closer together; it is also possible that air pressure impacts the physiology of pilot and controller attention spans.[69]

[69] Reeves (2014)

7.3.6. Model Output

Table 88 - Tower Bouillabaisse Model (2011-2013)

	Minor	Major	Hazardous	Catastrophic
Helicopter	1.3 (.33)	1.2 (.44)	1.6 (.94)	7.7* (6.7)
Corporate Jet	.67* (.12)	1.6* (.38)	.68 (.37)	1.4 (1.2)
Experimental Plane	.46 (.48)	3.3 (2.7)	5.3 (5.3)	63*** (63)
Ground Vehicle	.18 (.18)	6.2e-08*** (2.1e-08)	6.1e-08*** (2.3e-08)	12 (15)
Military Jet	.64 (.26)	2.5* (1.1)	2.5 (1.9)	1.7 (2)
Military Prop	.94 (.69)	4.3 (3.2)	1.3e-07*** (8.3e-08)	3.1e-07*** (2.7e-07)
Multiple Engine Prop	.83 (.14)	1.2 (.26)	1.6 (.57)	2.9 (2)
Regional Jet	.9 (.13)	1.5 (.33)	1.9 (.78)	9.7e-07*** (5.1e-07)
Single Engine Prop	.96 (.14)	1.6* (.33)	2.9** (1)	6.2** (3.6)
Information Exchange	.74* (.11)	.84 (.19)	.86 (.32)	4.7e-07*** (3.6e-07)
Aircraft 1 Control Status: NORDO	2.2e-08*** (1.0e-08)	2.2 (2.3)	4.3e-08*** (3.6e-08)	1.1e-07*** (1.7e-07)
Aircraft 1 Control Status: On Route	.8 (.19)	.91 (.31)	.85 (.58)	7.7e-07*** (7.4e-07)
Aircraft 1 Control Status: On SID/STAR	1.2 (.16)	.92 (.2)	1.2 (.55)	.85 (.94)
Aircraft 1 Control Status: Radar Advisories	.54* (.15)	.7 (.26)	.66 (.43)	1.3 (.71)
Aircraft 1 Control Status: Visual Approach	1.2 (.19)	2*** (.38)	1.2 (.38)	.65 (.35)
Total Aircraft Involved	1.6*** (.19)	1.6*** (.22)	1.7** (.33)	.13** (.097)
Aircraft 1 Phase of Flight: Arrival	.99 (.13)	.63 (.16)	1.2 (.47)	1.4 (1.1)
Aircraft 1 Phase of Flight: Climbing	.95 (.13)	.73 (.2)	1.5 (.7)	1.3e-06*** (1.2e-06)
Aircraft 1 Phase of Flight: Descending	1.2 (.25)	.89 (.25)	1.3 (.64)	9.2** (6.8)
Aircraft 1 Phase of Flight: Go Around/Missing Approach	1.3 (.27)	.88 (.25)	1.8 (1)	4.5 (4.7)
Aircraft 1 Phase of Flight: Level Flight	1.1 (.28)	1 (.32)	1.4 (1)	1.5 (1.5)
Aircraft 1 Phase of Flight: Surface	.78 (.12)	.42*** (.099)	1.5 (.59)	9.5*** (6.3)
Aircraft 1 Phase of Flight: Terminal Enroute Transition	.28 (.19)	1.2 (.65)	1.5 (1)	3.7e-07*** (4.0e-07)
Aircraft 1 Phase of Flight: VFC Traffic Pattern	1.5 (.53)	.58 (.25)	1.7 (1.1)	20** (18)
Airspace Type/Limitations for the reported event: Class B	.85 (.15)	1.6 (.4)	1.3 (.55)	5.8 (5.4)
Airspace Type/Limitations for the reported event: Class C	.83 (.19)	1.1 (.3)	.99 (.53)	1.7 (1.2)
Airspace Type/Limitations for the reported event: Class D	.85 (.19)	1.2 (.4)	1.5 (.73)	2.6 (1.3)

	Minor	Major	Hazardous	Catastrophic
Pilot Reaction: Evasive Action	1.2 (.38)	5.1*** (2)	7.7*** (3.3)	3.2 (2.6)
Pilot Reaction: Go Around	1.2 (.17)	1.5 (.33)	.63 (.28)	2.7 (1.9)
Pilot Reaction: None	1.5* (.28)	.74 (.2)	.51 (.22)	.34 (.24)
Total CPC years at facility 1	1 (.007)	.99 (.0096)	1 (.022)	1 (.03)
Loss of Communication	.84 (.14)	1.1 (.22)	.67 (.29)	.83 (.63)
Readback Error	.91 (.16)	.92 (.18)	.85 (.33)	.12* (.11)
ATC Level	1.1 (.052)	1.1 (.076)	1 (.12)	.73* (.11)
Total Operations	1 (.0013)	1 (.0024)	1 (.0033)	1 (.0065)
Traffic Complexity Rating	1 (.04)	1 (.051)	1.2 (.11)	.59** (.12)
Runway Count	.94 (.053)	.98 (.088)	1.1 (.13)	1.3 (.26)
Standard Deviation of Sea Level Pressure	.99 (.007)	.99 (.0096)	.97* (.015)	.93*** (.017)
Temperature Dew Point Difference	1 (.0033)	1*** (.0044)	1 (.0078)	.99 (.013)
Weather Complexity Factor	.75 (.18)	.95 (.37)	1.3 (.68)	.95 (.96)
Weather Indicator	1.2 (.21)	1.6* (.3)	.95 (.41)	.47 (.36)
Weather Indicator and Complexity Factor Interaction	.81 (.32)	1.3 (.58)	1.3 (.89)	2.9 (4.1)
Emergency Situation	1.7 (.51)	2.7** (.91)	6.4*** (2.4)	27*** (12)
Special Event	1.1 (.3)	2.5*** (.69)	1.8 (1)	5.7* (4.7)
Traffic Management Initiative	.56* (.16)	.51 (.38)	5.4e-08*** (1.9e-08)	2.0e-06*** (1.3e-06)
Unsafe Actions	1.7*** (.2)	1.1 (.21)	2.2* (.81)	.19* (.14)
Computer Entry Problem	.26 (.27)	1.4e-07*** (5.2e-08)	1.3e-07*** (6.8e-08)	6.8e-08*** (8.4e-08)
Flight Plan/PDC Processing Problem	.18** (.11)	5.6e-08*** (1.4e-08)	7.7e-08*** (2.4e-08)	1.6e-06*** (1.2e-06)
Radar Misidentification	.42 (.46)	3.1 (2.3)	4.6 (4.4)	2.5e-07*** (2.9e-07)
Acknowledgment	1.1 (.25)	2* (.57)	1.5 (1.1)	3.9 (5.2)
Clearance Problem	1.4*** (.14)	1.1 (.17)	.52* (.14)	.29 (.2)
Aircraft Performance or Pilot Actions: Timely Aircraft Descent/Climb	1.4 (.26)	1.6 (.46)	1.1 (.67)	.47 (.44)
Aircraft Performance or Pilot Actions: Timely Roll	1.3 (.29)	.91 (.32)	.68 (.52)	7.9* (8.2)
Aircraft Performance or Pilot Actions: Timely Speed Adjustment	1.8*** (.3)	1 (.35)	.68 (.5)	2.3 (1.9)
Year = 2011	1.6 (.56)	1.1 (.42)	4.3 (4.4)	.94 (.77)
Year = 2012	.88 (.28)	.92 (.33)	2.5 (2.6)	.57 (.44)
Standard errors in parentheses * p<0.05 ** p<0.01 *** p<0.001 N = 3,407				

Table 89 - TRACON Bouillabaisse Model (2011-2013)

	Minor	Major	Hazardous	Catastrophic
Clearance Problem	1.5*** (.083)	1.1 (.13)	1 (.29)	.049 (.091)
Radar Misidentification	.89 (.32)	2.9*** (.84)	3 (2.4)	6.9e-07*** (7.0e-07)
Unsafe Actions	1.7*** (.16)	1.2 (.16)	1.5 (.54)	2.4e-08*** (3.2e-08)
Information Exchange	.85 (.084)	.95 (.13)	.6 (.27)	1.1e-08*** (2.7e-08)
Emergency Situation	1.3 (.32)	1.5 (.64)	2.9* (1.4)	.82 (.76)
Special Event	1.1 (.21)	2.3*** (.5)	1.3 (.78)	.22 (.31)
Aircraft Performance or Pilot Actions: Timely Aircraft Turn	1.1 (.11)	1.3 (.2)	.61 (.35)	33 (63)
Coordination Complexity Factor	.83* (.068)	.64*** (.076)	.7 (.26)	.033* (.044)
Airspace and Procedures Complexity Factor	.83 (.082)	.76 (.12)	1 (.38)	.017*** (.018)
Communication Complexity Factor	1.1 (.081)	1 (.13)	1.6 (.54)	.45 (.47)
Helicopter	1.8 (.65)	.81 (.46)	8.8e-09*** (3.9e-09)	7.8e-06*** (.000014)
Corporate Jet	1.3* (.13)	1.1 (.17)	.63 (.34)	2.0e-07*** (2.5e-07)
Experimental Plane	.72 (.52)	1.3 (.93)	2.9 (2.2)	228** (427)
Military Jet	1.6* (.31)	1 (.29)	1.6 (1.2)	2.5e-06*** (3.0e-06)
Military Prop	.97 (.29)	.94 (.4)	1.4 (1.6)	.000013*** (.00002)
Multiple Engine Prop	1.2 (.13)	.83 (.15)	.87 (.44)	1.5 (2.2)
Regional Jet	.95 (.11)	.98 (.12)	1 (.42)	2.7e-08*** (7.0e-08)
Single Engine Prop	1.4*** (.15)	.69 (.14)	1.2 (.58)	25*** (23)
Flight Plan: None	.23 (.24)	.91 (.63)	4.7 (5)	3086** (7811)
Flight Plan: Unknown	1.6 (.99)	1.8 (1.5)	4.8* (3.3)	1.8e-07*** (4.0e-07)
Flight Plan: VFR	.53** (.12)	.49* (.14)	1.6 (1.1)	87** (125)
Total Aircraft Involved	1.7*** (.11)	2.2*** (.23)	1.9*** (.29)	3.6e-10*** (1.1e-09)
Phase of Flight: Arrival	1.5*** (.16)	.86 (.14)	1.2 (.56)	15 (31)
Phase of Flight: Climbing	.88 (.095)	.98 (.15)	1.4 (.77)	8.7 (13)
Phase of Flight: Descending	1.3*** (.11)	1.2 (.17)	1.3 (.76)	9.2 (13)
Phase of Flight: Go Around/Missing Approach	1.7* (.42)	1.3 (.54)	1.0e-08*** (4.4e-09)	1.1e-06*** (2.3e-06)
Phase of Flight: Level Flight	1.1 (.12)	1.3 (.23)	1.5 (.72)	3 (4.1)
Phase of Flight: Surface	.93 (.42)	1.2e-14*** (5.2e-15)	2.6e-14*** (1.7e-14)	11912*** (24184)
Phase of Flight: Terminal Enroute Transition	.55* (.17)	.84 (.29)	1.5 (1.1)	1 (2)
Phase of Flight: VFC Traffic Pattern	.69 (.36)	.51 (.47)	3.9e-08*** (2.8e-08)	6 (9.8)
Airspace Type/Limitations for the reported event: Class B	.91 (.29)	2.7*** (.68)	2.5 (1.6)	6.3 (13)

	Minor	Major	Hazardous	Catastrophic
Airspace Type/Limitations for the reported event: Class C	.76 (.29)	2.2 (.94)	1.2 (1.4)	9.1e-12*** (2.5e-11)
Airspace Type/Limitations for the reported event: Class D	.71 (.26)	2.1 (1)	6.8e-09*** (3.7e-09)	2.8e-08*** (3.3e-08)
Pilot Reaction: Go Around	1.5 (.51)	1.2 (.56)	9.7e-09*** (3.5e-09)	320** (617)
Pilot Reaction: None	2 (3.1)	5.5e-09*** (7.6e-09)	3.4e-09*** (5.7e-09)	444 (1600)
Total CPC years at facility 1	.99 (.0052)	.99 (.0079)	1 (.021)	1.1 (.059)
Approach Control Handoff	1 (.19)	.76 (.22)	1.5 (.8)	1.5 (1.6)
Terminal Radar Approach Control Satellite	.86 (.081)	1.3* (.18)	.88 (.25)	10 (12)
Terminal Radar Approach Control FLM	.93 (.21)	.72 (.28)	5.8e-09*** (1.7e-09)	10 (17)
Terminal Radar Approach Control Arrival	.84 (.075)	.85 (.077)	.55 (.2)	1.8 (1.4)
Terminal Radar Approach Control Final	1.2** (.087)	.94 (.12)	.96 (.35)	2 (1.8)
Loss of Communication	.99 (.13)	1.7*** (.22)	.82 (.4)	5.6*** (2.5)
Readback Error	1.2 (.12)	1.4* (.2)	.68 (.37)	.72 (.46)
ATC Level	1 (.043)	1.1 (.058)	1 (.12)	1.2 (.42)
Traffic Complexity Rating	1.1*** (.029)	1.1* (.045)	.95 (.071)	.86 (.27)
Runway Count	1.1*** (.011)	1.1*** (.014)	1.1* (.026)	1 (.064)
Standard Deviation of Sea Level Pressure	1 (.0083)	1 (.011)	.98 (.039)	.9 (.092)
Temperature Dew Point Difference	1** (.0037)	1*** (.0043)	1* (.0086)	.85** (.052)
Weather Complexity Factor	.73 (.14)	.47 (.2)	.43 (.45)	.055 (.094)
Weather Indicator	1.3* (.15)	.9 (.14)	1.7 (.56)	2.1 (2)
Aircraft Performance of Pilot Action Complexity Factor	1.1* (.08)	.94 (.095)	.68 (.21)	.39 (.44)
Cloud Coverage x Sea Level Pressure	1 (.0016)	1 (.002)	1 (.0065)	1 (.02)
Total Operations	1*** (.00086)	1** (.0012)	1* (.0019)	1 (.0064)
Year = 2011	3.2*** (.87)	1.9 (.64)	3 (3)	1.4e+08*** (5.7e+08)
Year = 2012	1.4 (.31)	.99 (.29)	1.5 (1.5)	1.9e+08*** (7.2e+08)
Standard errors in parentheses * p<0.05 ** p<0.01 *** p<0.001 N = 6,119				

8. CONCLUSIONS AND NEXT STEPS

The in-depth analysis of airborne loss of separation incidents covered in this report has made several wide-ranging and interesting findings. As expected, the factors affecting the severity categorization of a loss of separation incident are both complex and extensive. This does not mean, however, that all the results were conclusive and further research is need to fully understand these variables complex nature with severity. The findings presented in this section are the central takeaways from this report.

One important result is that **'Catastrophic' severity incidents seem vastly different from the other four severity levels**. In other words, the interaction between the severity outcomes and variables studied in this report sometimes have the opposite effects going from the first four levels and then to Catastrophic. For example, increasing the traffic complexity rating for both Tower and TRACON facilities increases the likelihood of a more severe incident, but decreases the likelihood of the most severe (Catastrophic). Results like these have immediate policy ramifications: attempting to decrease the likelihood of an incident becoming Catastrophic by focusing on factors that would reduce the overall severity of an incident may have no end effect. This is because the factors that contribute to the first four severity outcomes do not always hold for Catastrophic incidents.

Another major finding is that **causal factor variables tend to only have a marginal statistical relationship with severity outcomes when included in fully specified models**. There were several explanations for this small impact on severity, with the most common cause being an overall lack of variation due to infrequently filled out causal factor data fields. Even after aggregating certain variables, small sample sizes persisted making it impossible to garner any statistical inferences in models. Other common issues were subjectivity and possible reporting bias of the causal factor variable. For example, several causal factors were associated with decreased severity levels; however, it was unclear whether this was a true effect, or just statistical noise due to the subjective reporting process for these variables. Better data collection and data entry for causal factor variables would help mitigate these types of potential issues.

Throughout the report, descriptive statistics and models were separated by facility type (Tower vs. TRACON) due to the inherent differences between the two. This came as no surprise, but results varied significantly between facility types, and this distinction was important to highlight due to the varying policy implications for each facility type. For example, daily operations had no relationship with severity levels in Tower facilities, but did have a statistically significant and *decreasing* likelihood of more severe incidents for TRACON facilities as daily operations increased.

Through the modeling process it was also discovered that there was a possible inflection point around the Major severity outcome. In other words, when severity outcomes were grouped together between the non-severe (Minimal/Minor) and severe (Major/Hazardous/Catastrophic) levels based on FAA's categorization, results can be inscrutable. This is because Major incidents often times exhibit characteristics of non-severe incidents. An important policy implication stemming from this result is the need to keep these 5 severity classification separated due to their nuanced nature.

The following are specific variable findings by model section. Certain findings contain odds ratios that were calculated from binary logit (non-severe vs severe) models not presented in the main document text but are available in the Appendix.

8.1. Aircraft

- **Single Engine Props**: Single engine props are **1.7 times more likely to be associated with severe incidents** than are single-aisle jets. They are **3.6 times more likely to be catastrophic** in Tower facilities.
- **Experimental Aircraft**: In Tower facilities, incidents with experimental aircraft are **6.2 times** more **likely to be severe, and 21 times more likely to be catastrophic**. In TRACON facilities, they are **22 times more likely to be catastrophic**.
- **Visual Approaches**: Incidents with visual approaches are **2.6 times more likely to be associated with severe incidents** in Tower facilities than incidents with instrument approaches.

8.2. Facility

- **Aircraft/Pilot Complexity Factor:** When aircraft/pilot complexity is a factor in a Tower incident, there is **a 120% percentage point increase in the probability for a Hazardous outcome and 100% percentage point increase in the probability for a Catastrophic outcome.**
- **Communication Complexity Factor:** When communication complexity is a factor in a Tower incident, there is a **70% percentage point decrease in probability for Catastrophic outcomes.**
- **Coordination Complexity Factor:** When coordination complexity is a factor in a Tower incident, the percentage change in probability **decreases by close to 100% for Catastrophic outcomes.**
- **Traffic Complexity Rating:** As facility level increases for Tower facilities, the probability of a Catastrophic outcome **decreases from 0.01 to close to zero.**
- **Facility Influences:** When facility influences are a factor for TRACON facilities, the percentage change in probability **decreases by 150% for Catastrophic outcomes**.
- **Operations:** An increase in operations had no effect on severity for Tower facilities, and a decreasing effect on the likelihood of a more severe incident for TRACON facilities.

8.3. Control Status

- **Training is in Progress:** Incidents with training in progress are **1.4 times more likely to be severe** than incidents without training in progress in Tower facilities.

8.4. Communication

- **Flight Plan/PDC Processing Problem**: In Tower facilities, flight plan/PDC processing problems **are overwhelmingly low in severity.**

8.5. Airspace and Pilot Actions

- **Aircraft/Pilot action complexity factor**: In Tower facilities, these are **1.6 times more likely to be associated with severe incidents, and 2.2 times more likely to be catastrophic**.
- **Airspace Type D**: Type D airspace is **2.3 times more likely to be associated with severe incidents** for all facility types, and is **3.4 times more likely to be catastrophic.**
- **Pilot Evasive Actions**: There is a **300% percentage point decrease in the probability of a severe incident** for all facility types if the pilot takes action to avoid a potentially dangerous situation.
- **Radar Misidentification**: Incidents with radar misidentification are **3 times more likely to be severe** than incidents without Radar Misidentification in TRACON facilities.
- **Acknowledgement Problems**: When an acknowledgement problem is cited as a causal factor, incidents are **1.7 times more likely to be severe** in both Tower and TRACON facilities.
- **Loss of Communication**: Incidents with a loss of communication are **1.4 times more likely to be severe** in Tower facilities; in TRACON facilities they are **2 times more likely to be severe, and are 5 times more likely to be catastrophic**.

8.6. Next Steps

Aside from particular areas of interests noted above, a more general area for future research should be considered. This type of statistical analysis is likely best conducted as part of a feedback loop with more traditional human factors research. That is, econometric analysis is quite powerful in differentiating which factors have the most influence on incident severity (or, in a different type of model, incident frequency). What these models do *not* provide, however, is a specific intervention to mitigate the factor. Econometric analysis can then be seen as a first step in priority-setting for human factors research that can follow up with a specific remediation to the most pressing variables identified.

After developing a mitigation, econometric research can then a) verify how successful the mitigation was and b) help set priorities for the next round of human factors research. Teams of cooperating econometricians and human factors researchers can also help FAA achieve the biggest “bang for the buck” by combining information on both the size of the change to risk that can be reduced with a mitigation and the cost and likelihood of a successful mitigation.

This application is best understood with the following example: consider a human factors team that only has time/budget to address one factor; the econometric research can provide an estimate of the impact of having removed that factor on the probability of severe incidents and help the team select the factor likely to have the biggest impact. In reality, the FAA faces a more complex situation wherein a human factors team, instead of having time to address a generic “single factor,” has $300,000 and two years of time; moreover, the probability of creating a successful mitigation likely differs across factors. Utilizing these modeling results, the team can then select the mix of research projects that helps to not only optimize their time but which projects would have the greatest impact on increasing airspace and travel safety.

BIBLIOGRAPHY

Biernbaum, Lee and Garrett Hagemann, “Runway Incursion Severity Risk Analysis” John A. Volpe National Transportation Systems Center: U.S. DOT (2012).

Broach, Dana, and David J. Schroeder, “Relationship of Air Traffic Control Specialist Age to En-Route Operational Errors,” prepared for: Federal Highway Administration, Office of Aerospace Medicine, DOT/FAA/AM-05/22 (2005).

Broach, Dana. “An Examination of the Relationship Between Air Traffic Controller Age and En-Route Operational Errors. Proceedings of the 10th International Symposium on Aviation Psychology, Columbus, OH (1999). Volume 2, p. 934.

Center for Naval Analyses Corporation. “Analysis of Operational Errors for Air Route Traffic Control Centers. Volume I: Final Report. (IPR 95-5108). Alexandria, VA: The Center of Naval Analyses Administration William J. Hughes Technical Center (1995).

Dow, Jay K., and James W. Endersby, "Multinomial Probit and Multinomial Logit: A Comparison of Choice Models for Voting Research," Electoral Studies, 23:1 (2004), pp. 107-122.

Greene, William H, Econometric Analysis: Fifth Edition, Upper Saddle River, New Jersey: Pearson Education, Inc. (2003).

Islam, Samantha, and Fred Mannering, "Driver Aging and Its Effect on Male and Female Single-Vehicle Accident Injuries: Some Additional Evidence," Journal of Safety Research, 37:3 (2006), pp. 267-276.

Kaplan, Sigal and Carlo Giacomo Prato, “Braking News: Link Between Crash Severity and Crash Avoidance Maneuvers,” Transportation Research Record: Journal of the Transportation Research Board, No. 2280 (2012) pg. 75-88.

Kockelman, Kara Maria, and Young-Jun Kweon, "Driver Injury Severity: An Application of Ordered Probit Models," Accident Analysis & Prevention, 34:3 (2002), pp. 313-321.

Long, Scott, J. and Jeremy Freese, “Regression Models for Categorical Dependent Variables Using Stata.” Stata Press (2006), pp. 243-246.

O'Donnell, C. J., and D. H. Connor, "Predicting the Severity of Motor Vehicle Accident Injuries Using Models of Ordered Multiple Choice," Accident Analysis & Prevention, 28:6 (1996), pp. 739-753.

Pfleiderer, Elaine M., and Carol A. Manning, “Prediction and Classification of Operational Errors and Routine Operations sing Sector Characteristics Variables,” prepared for: Federal Highway Administration, Office of Aerospace Medical Institute, DOT/FAA/AM-07/18, (2007).

Pfleiderer, Elaine M., Cheryl L. Scroggins, and Carol A. Manning, “Logistic Regression Analysis of Operational Errors and Routine Operations Using Sector Characteristics,” Federal Highway Administration, Office of Aerospace Medical Institute, DOT/FAA/AM-09/04 (2009).

Pounds, Julia, and Anthony Ferrante, “FAA Strategies for Reducing Operational Error Causal Factors,” prepared for: Federal Highway Administration, Office of Aerospace Medicine (2003), DOT/FAA/AAM-03/19.

Reeves, John and Weil, John, “Ventilatory Acclimatization to High Altitude.” Encyclopedia of Occupational Health and Safety, accessed 12/17/14, chapter 37.

Rodgers, Mark D., Lendell G. Nye, “Factors Associated with the Severity of Operational Errors at Air Route Traffic Control Centers,” prepared for: Federal Highway Administration, Office of Aerospace Medicine, DOT/FAA/AM-93/22 (1993).

Schneider IV, William H., Peter T. Savolainen, and Karl Zimmerman, “Driver Injury Severity Resulting from Single-Vehicle Crashes Along Horizontal Curves on Rural Two-Lane Highways,” Transportation Research Record: Journal of the Transportation Research Board, No. 2102 (2009), pg. 85-92.

Schroeder, David J., and Lendell G. Nye, “An Examination of the Workload Conditions Associated with Operational Errors/Deviations at Air Route Traffic Control Centers,” prepared for: Federal Highway Administration, Office of Aerospace Medicine, DOT/FAA/AM-93/22 (1993).

Vortac, O.U., Edwards, M.B., Jones, J.P., Manning, C.A., and A.J. Rotter. “En Route Air Traffic Controller’s Use of Flight Progress Strips: A Graph-Theoretic Analysis.” Washington, D.C.: Office of Aviation Medicine Report (1993), DOT/FAA/AM-92.21.

Washington, Simon P., Matthew G. Karlaftis, and Fred L. Mannering, Statistical and Econometric Methods for Transportation Data Analysis: Second Edition, Boca Raton, FL: Chapman & Hall/CRC (2011).

Williams, Richard, “Generalized ordered logit/partial proportional odds models for ordinal dependent variables,” The Stata Journal (2006), 6, Number 1, pp. 58-82.

Xie, Yuanchang, Yunlong Zhang, and Faming Liang, "Crash Injury Severity Analysis Using Bayesian Ordered Probit Models," Journal of Transportation Engineering, (2009), p 7.

Yasmin, Shamsunnahar, Sabreena Anowar, and Richard Tay, “Effects of Drivers’ Actions on Severity of Emergency Vehicle Collisions,” Transportation Research Record: Journal of the Transportation Research Board, No. 2318 (2012) pg. 90-97.