5-2022

A Probabilistic Fog Stability Index for Predicting the Occurrence of Radiation Fog at U.S. Airports

Kelisha Ali

A Probabilistic Fog Stability Index for Predicting the Occurrence of Radiation Fog at U.S. Airports

by

Kelisha Ali

Bachelor of Science
Aviation Management
Florida Institute of Technology
2021

A thesis submitted to the College of Aeronautics at
Florida Institute of Technology
in partial fulfillment of the degree requirements for a

Master of Science in Aviation in Airport Development and Management

Melbourne, Florida
May 2022

We, the undersigned committee, hereby approve the attached thesis,
"A Probabilistic Fog Stability Index for Predicting the Occurrence of Radiation Fog at U.S. Airports"
by
Kelisha Ali

Ulreen O. Jones-McKinney, Ph.D.
Assistant Professor,
College of Aeronautics
Major Advisor

Ryan White, Ph.D.
Assistant Professor,
Mathematical Sciences

Deborah S. Carstens, Ph.D.
Professor
College of Aeronautics

John E. Deaton, Ph.D.
Professor and Interim Dean
College of Aeronautics

Abstract

Title: A Probabilistic Fog Stability Index for Predicting the Occurrence of Radiation Fog at U.S. Airports

Author: Kelisha Ali

Advisor: Ulreen O. Jones-McKinney, Ph.D.

Fog is a phenomenon that is widely known to affect the aviation industry adversely, as evidenced by economic losses due to the hindrance of fog on airport operations. This is because fog has been a consistent cause for delays, diversions, and cancellations of scheduled commercial airline flights that have subsequently resulted in a substantial negative economic impact on the transportation industry as well as society. This study's purpose was to determine suitable predictors of the occurrence of radiation fog at U.S. airports. It assessed an extant Fog Stability Index (FSI) and 1-Day Persistence model for their reliability in predicting radiation fog. This research study addressed the overarching question: is it possible to predict the occurrence of radiation fog at airports using a probabilistic methodology? Therefore, the study's objective was to compare the reliability of using a theoretical or traditional approach to FSI, a probabilistic FSI, and 1-Day persistence as predictors of radiation fog.

The research utilized data period spanning the years 1973 through 2020, at six airports in east-central Florida. The study utilized archival data from Iowa State University's Environmental Mesonet and radiosonde data provided by NASA's varying weather observation equipment. The study isolated the occurrences of radiation fog as

opposed to advection, sea fog, and other types of fog. This was of specific interest because radiation fog is potentially predictable with such measures and could affect airports to a more noticeable degree, comparatively. Thus, observations were limited to 1000Z to 1500Z and METAR weather codes to BR, FG, MIFG, and BCFG for the occurrence of radiation fog.

A statistical analysis of the data was performed utilizing logistical regression analyses supporting the use of a dichotomous dependent variable: the occurrence or non-occurrence of fog. The preliminary analyses found that 1-Day persistence may or may not be a suitable predictor of radiation fog. This was inconclusive due to the rarity of those events within the sample and resulted in a lack of viable data for logistic regression analyses. Further research will be required confirm the suitability of 1-day persistence for predicting radiation fog.

The primary analyses found that both the theoretical and probabilistic approaches to using FSI were reliable predictors of fog as evidenced by a contingency analysis of predicted fog events versus actual fog events within the sample. However, the probabilistic approach yielded better results with respect to hits – correctly predicting the occurrence of fog when it occurred, and misses – not predicting that fog would occur when it did, as opposed to the traditional FSI high, medium, low fog-event chance model. The traditional or theoretical model yielded a lower percentage of hits and a greater percentage of misses. Thus, the study concluded that using a probabilistic FSI model to predict radiation fog events could positively impact the air transportation industry by providing accurate, additional information to decision-makers reducing the consequent economic impact of delays, diversions, and cancellations.

Table of Contents

List of Figures

List of Tables

Acknowledgment

I would first and foremost like to thank God and my family for being super supportive throughout this entire process. My mom who is a Meteorologist at the Trinidad and Tobago Meteorological Service Center has been a huge help in guiding me through editing this thesis, as well as my dear friend John Offutt. I would also like to thank my Professor Dr. Ulreen Jones-McKinney, for helping me through this process step by step and my committee members Dr. Ryan White and Dr. Debbie Carstens. I would like to thank NASA and the 45th Weather Squadron for approaching me with this field of study to embark upon.

Glossary of Terms and Abbreviations

BCFG	Patchy Fog
BCFG BR	Patchy Fog, Mist/Light Fog
BR	Mist or Light Fog
BR BCFG	Mist/Light Fog, Patchy Fog
BR MIFG	Mist/Light Fog, Shallow Fog
DZ BR	Drizzle, Mist/Light Fog
FG	Fog
hPa	Hectopascal
METAR	Meteorological Terminal Air Report
MIBCFG	Shallow Patchy Fog
MIFG	Shallow Fog
MIFG BR	Shallow Fog, Mist/Light Fog
RA BCFG	Rain, Patchy Fog
RA BR	Rain, Mist/Light Fog
RA FG	Rain, Fog
TS BR	Thunderstorm, Mist/Light Fog
TSRA BR	Light/Moderate Thunderstorm with Rain, Mist/Light Fog
KCOF	"COF" or Patrick Space Force Base Airport
KMCO	"MCO" or Orlando International Airport
KMLB	"MLB" or Melbourne Orlando International Airport
KSFB	"SFB" or Orlando Sanford International Airport
KTTS	"TTS" or NASA Shuttle Landing Facility

KXMR	"XMR" or Cape Canaveral Space Force Station Skid Strip Airport
VCSH BR	Vicinity Showers, Mist/Light Fog
VCTS +RA BR	Vicinity Thunderstorm, Heavy Rain, Mist/Light Fog
VCTS RA BR	Vicinity Thunderstorm, Rain, Mist/Light Fog
VCTS-RA BR	Vicinity Thunderstorm, Light Rain, Mist/Light Fog
Z	Zulu Time or also referred to as Greenwich Mean Time

Chapter 1

Introduction

Purpose of Study

This study aims to determine suitable predictors of the occurrence of radiation fog at U.S. airports. Therefore, the study has explored aspects of radiation fog, including its occurrence, formation, persistence, and its relation to the physical environment. Specifically, the study has aimed to develop a probabilistic FSI assess if the FSI and 1-Day Persistence of the occurrence of fog are significant predictors of radiation fog at airports located in east-central Florida.

According to the Federal Aviation Administration (FAA), aircraft delays at U.S. airports due to the weather represented nearly 10 million minutes in 2013 and approximately 69% of all delays (FAA, 2021). An economic analysis of data is presented in the background and rationale below. From October through March, the combination of airport surface winds, low ceiling, and visibility conditions represented about 75 percent of delays from 2008 to 2013. Weather events such as low visibility due to fog can result in costly operational delays at airports. While the occurrence of fog is relatively challenging to forecast, the ability to predict events such as radiation fog using indicators or factors, such as the FSI, may be of practical significance at airports around the U.S.

My analysis focused on the occurrence of fog at airports within a 50-nautical mile radius of Cape Canaveral. Table 1 lists the airports identified due to their proximity to Cape Canaveral Space Force Station Skid Strip (XMR) and NASA Space Shuttle Landing (TTS), where radiosonde data is collected. The underlying assumption is that this radiosonde data is representative of that 50 nautical mile radius. It also recognizes that fog can be a geographically inconsistent phenomenon, so using multiple observing sites increases the likelihood of detecting the event. The issue with radiosonde observations collected from a single location is that the forecasted fog may occur at more than one airport in east-central Florida. This means that forecasted fog may be present at one airport but not another. Once fog lasts for more than a specific forecast timeframe, or the entirety of the day, it is termed "1-Day persistent fog".

Table 1.1

East-Central Florida Airports Experiencing Radiation Fog

Airport Code (K-)	**Airport Name**
MCO	Orlando International Airport
MLB	Orlando Melbourne International Airport
SFB	Orlando Sanford International Airport
COF	Patrick Space Force Base
TTS	NASA Space Shuttle Landing Facility
DAB	Daytona Beach International Airport

Background and Rationale

Meteorologists with the 45th Weather Squadron have proposed the study to develop a probabilistic FSI and determine its utility to predict fog at CCSFS/KSC, as indicated by its detection at various airports in east-central Florida. Fog is a prominent issue in transportation, especially within the aviation industry. When fog occurs, it can lead to diversions, delays, or even cancellations of flights. Bergot (2021) explained that with consideration of the overall observation, simulation, and predictability of fog, fog levels within the atmosphere have significantly increased over the past few decades due to the rapid increase in airline, maritime and vehicular traffic.

The financial implications of fog impacts can be immense within the aviation industry over time. The Gandhi International Airport in India sustained a value of economic losses nearing 3.9 million USD for airlines between the years 2011 and 2016. Observations have shown that the forecasting of fog may be very incomplete. There is a high level of difficulty with accurately predicting a precise amount of fog in the atmosphere at a particular time. This study urged that better forecasts would help guard against the losses associated with the delays at airports.

Other studies have also indicated this same concern of financial losses in the aviation economy in India. Kulkarni (2019) studied the Indira Gandhi International (IGI) Airport, one of India's busiest airports, accommodating more than 900 flight operations per day. The study concluded that to have a maximum economic profit, airlines at the airport should be operating at full capacity with no delay. This implies that any interruption in normal operations caused by weather phenomena such as fog may result in financial losses to the aviation industry.

Economic Analysis

In 2019, the average cost of aircraft block (taxi plus airborne) time for U.S. passenger airlines was $74.24 per minute. The largest line item, fuel costs declined 6.5 percent to $25.26 per minute. Crew costs, the second-largest line item, rose 5 percent to $24.55 per minute. Maintenance and aircraft ownership actually respectively grew 2.2 percent and 3.2 percent, while all other costs rose 1.1 percent (Airlines for America, 2020).

Table 1.2

2019 Economic Impact Analysis

2019 Cost Per minute Delay		
Cost Centers	**Cost Per Minute**	**Total Cost (Billions)**
Airlines	$ 74.24	$8.30
Passengers	$ 161.90	$18.10
Lost Demand	$ 46.81	$2.40
Indirect	$ 23.69	$4.20
Grand Total	**$ 306.64**	**$33.00**

Source: Airlines for America, 2020

The Economic Impact Analysis presented in Table 1.2 breaks down the cost to impacted parties affected by any type of delay to a scheduled commercial flight. Cost centers impacted by delays include airlines, passengers, lost demand, and indirect. The largest impact is to passengers at approximately $161.90 per minute delay. Overall, the monetary impact to all parties/cost centers is $306.64 per minute delay, or $33 billion in 2019.

Definition of Terms

The paragraphs below list operational definitions for the key terms and phrases relative to the proposed study.

1. ***Advection/ Sea Fog***

   Advection fog is a type of fog which is caused by the horizontal movement of warm moist air over a cold surface. For example, warm moist air flowing over a cold body of water can result in the formation of advection fog, which is also known as sea fog (AMS, 2012).

2. ***Fog***

Fog can be defined as water droplets that are suspended in the atmosphere. They are located within the proximity of the earth's surface that it affects visibility levels. According to the international definition, fog is said to reduce visibility below 1 km, also known as 0.62 miles (AMS, 2012). Fog and clouds are very similar. However, a distinct difference is that the base of fog is at the earth's surface, whereas the base of clouds are above the earth's surface. The level of visibility is affected by the concentration of cloud condensation nuclei and the resulting distribution of droplet size. All the types of fog discussed form when the temperature and dewpoint levels of air become the same, or nearly identical (AMS, 2012).

3. ***Fog Stability Index (FSI)***

FSI is calculated by the addition of the dew point deficit (the overall temperature minus the dew point temperature), the atmospheric stability (the overall temperature minus the temperature at 850hPa altitude), and the wind speed at 850hPa altitude. FSI is an overall indicator of fog occurrence. According to an Air Force weather technical note, the origin of the Fog Stability Index came from developer Herr Harald Strauss. It was used during the 2nd World War in Germany during the late 1930s and 1940's (Strategic Air Command Offutt AFB NE, 1989).

In terms of numerical measurement, a value of FSI over 31 will result in a high likelihood, between 31 and 51 will result in a moderate probability, and values of FSI measuring over 55 will result in a low likelihood of fog forming (Holtslag et al., 2010). One of the shortfalls of FSI is that these high/moderate/low forecast categories are not defined. In addition, categorical forecasts can be inherently misleading in that when the forecast is close to a category edge, a tiny change in input can cause an apparently significant change into another category. These shortfalls are what inspired the initial suggestion for this thesis, i.e., to convert the undefined three FSI categories into continuous rigorously defined probabilities.

The most favorable conditions for fog formation are levels of high humidity, stable atmosphere, and low wind speed.

$$\boxed{FSI = 2(T - Td) + 2(T - T_{850}) + WS_{850}}$$

Equation 1.1 Calculation of FSI (*Holtslag et al., 2010)*

Where:

T = Temperature (surface)

T_d = Temperature at dew point (surface)

T_{850} = Temperature at 850hPa

WS_{850} = Wind Speed at 850hPa

4. ***Freezing Fog***

This fog consists of water droplets that remain in a liquid state until they meet a surface, where they freeze upon contact. They coat exposed objects with rime and/or glaze (AMS, 2012).

5. ***Ice Fog***

Ice fog forms when the air is so cold that water vapor in the air forms suspended ice crystals at and near the surface.

6. ***Odds***

Odds are defined as a ratio of probabilities. More specifically, odds represent the ratio between the probability of an event occurring to not occurring. If this ratio is the probability of an event occurring relative to the probability of the event not occurring, then it is referred to as *odds in favor*. If this ratio is the probability of an event not occurring relative to the probability of the event occurring, then it is referred to as *odds against*. For example, the probability of getting a 5 on a single roll of a die is 1/6 and the probability of not getting a 5 is 5/6. Therefore, the odds in favor of getting a 5 are 1/5 (denoted as 1 to 5 or 1:5) and the odds against getting a 5 are 5:1.

7. ***Odds ratios***

Odds ratios are defined as a ratio of odds. An odds ratio of 1 implies that the odds of both groups being compared are the same. Odds ratios greater than 1 imply that the odds for the first group are greater than those of the second group. For example, if a study finds that three of every four male smokers develop lung cancer, then the odds of male smokers developing lung cancer are 3:1. If the same study finds that two of every three female smokers develop lung cancer, then the odds for female smokers developing lung cancer are 2:1. Given the respective odds of male and female smokers developing lung cancer are 3:1 and 2:1, then the odds ratio are 3:2, which means that male smokers are 1.5 times more likely to develop lung cancer than female smokers.

8. ***Radiation Fog***

Radiation fog is a common type of fog and is usually produced over an area of land where the levels of radiational cooling reduce the air temperature to or below its dewpoint. It forms overnight as the air near the ground surface cools and stabilizes. Nighttime cooling has the ability to intensify all types of fogs. The factors that may lead to radiation fog formation include a shallow surface area, moist air under a dry layer, clear skies, and light surface winds. (AMS, 2012). Radiation fog is most prevalent during fall and winter seasons (FAA, 2021).

9. ***Temperature Inversion***

A temperature inversion in a layer occurs when the temperature increases with altitude. The principal characteristic of an inversion layer is known as its marked static stability, so that very little turbulence can occur within it (AMS, 2012).

Research Questions

The research questions posed in this study were as follows:

RQ1. Are FSI, a Probabilistic FSI, and 1-Day Persistence reliable predictors of the occurrence of radiation fog at east-central Florida airports?

RQ2. What predictor has causal priority, and in what order?

Significance of Study

This study may be of practical significance for airports in the U.S. The use of the FSI to predict the occurrence of radiation fog in a timely manner could help better prepare airports for the phenomenon and mitigate operational and financial impacts on airports. This study is expected to add to the body of knowledge about using FSI to predict fog using probabilities rather than the categories of high, medium, low as noted in the definitions of terms. Furthermore, the study will assist the 45th Weather Squadron in forecasting fog at Cape Canaveral Space Force Station Skid Strip, Patrick Space Force Base, and the NASA Space Shuttle Landing Facility.

Study Limitations and Delimitations

Limitations and delimitations identified for the proposed study include:

1. ***Time of Day*** – This study focused on radiation fog. To ensure that data utilized only captures radiation fog, the research delimited observations to those occurring between 1000Z and 1500Z, and fog persisting beyond this timeframe was assumed to be a sea or advection fog event for east-central Florida locations. Fortunately, sea fog is relatively rare in the study area.
2. ***Measuring Equipment*** – The research is limited to accessible weather balloons and radiosonde equipment locations within the study area of east-central Florida. Fortunately, an extensive archive of these past observations was readily available.
3. ***Data set/ Study period*** – Radiation fog occurs primarily during the months of November through March. The data set analyzed was limited to these months and analyzed for a period of 48 years (1973 – 2020) with the use of random sampling.
4. ***Sampling Sources/Observation Locations*** **–** This research was concentrated on the airports listed in Table 1.1. However, on any given day with appropriate FSI values that predict fog, fog may have occurred at an airport not listed in the table. That is, the study was not a fully comprehensive study of all the airports in the east-central Florida.
5. ***Use of METAR fog codes*** **–**This research focused on radiation fog; in particular, therefore, the study was limited to only four types METAR fog codes to conduct the analysis on the random generated sample. This was due to the assumption and decided interpretation of what METAR codes were truly related to the occurrence of radiation fog.

Chapter 2

Review of Related Literature

Introduction

This chapter is organized into two main sections: an overview of underlying theory and a review of past research studies. The first section provides the theoretical background and foundation on which this study is established and contains an outline of the different theories related to the formation, distribution, and period of the meteorological phenomenon known as fog. It is aimed at reviewing past theories with specific attention to the aspects of radiation fog and the comparison to other types of fog. The second section of this chapter provides a review of past research studies and related literature, as well as a summary and a discussion of the implications in this study.

Overview of Underlying Theory

Duynkerke (1990) provides examples of the models used to predict fog and explains that fog has been difficult to predict over the past 20 years due to two main aspects. Firstly, the uncertainty of predicting cloud cover exists, which is a major contribution to fog. Secondly, there is a lack of knowledge of the factors that help evaluate the formation and the dispersal of fog within certain weather conditions in a particular location. This phenomenon, depending on the density and boundary thickness layer may be a great hazard to not only aviators, but also to other modes of public transport.

The most important factors for the formation of fog are as follows: the cooling of moist air by radiative flux divergence, vertical mixing of heat and moisture (temperature inversion), vegetation, horizontal and vertical wind, heat and moisture transport in soil, advection and topographic effects. After the fog formation by these factors, the fog may become more widespread, or more prevalent due to longwave radiative cooling at the fog top, gravitational droplet settling, fog microphysics, and shortwave radiation (Duynkerke, 1990).

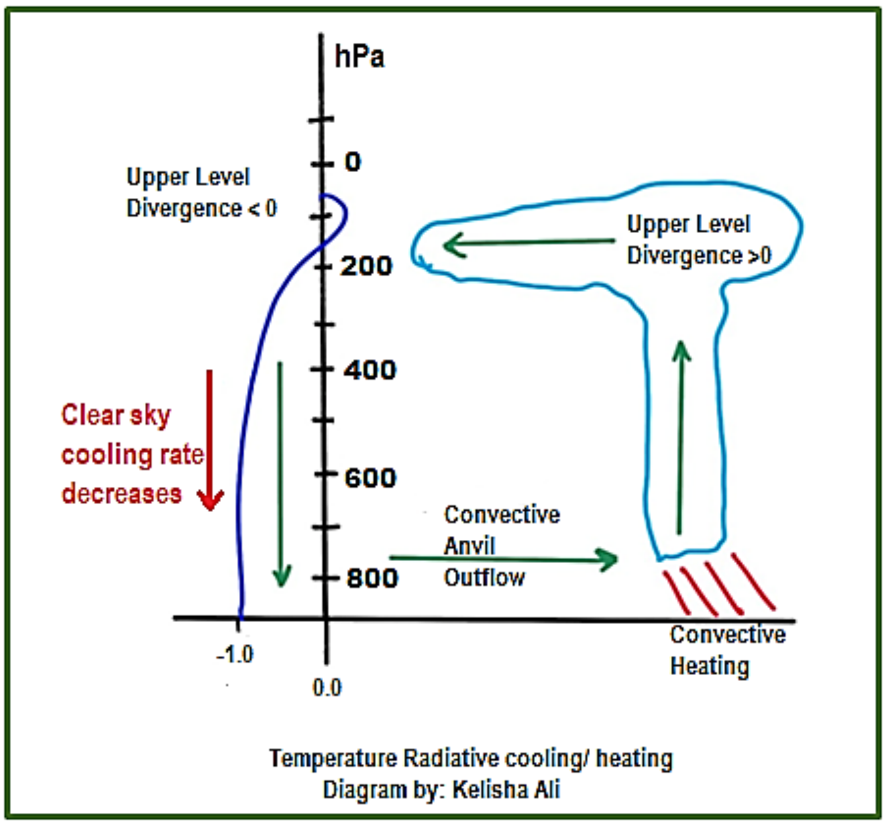


Figure 2.1 *Clear sky radiative cooling due to upper-level divergence and convective anvil outflow.*

Looking into greater detail about the cooling of moist air by Radiation Flux divergence, we find convective anvil clouds usually occur where the radiative cooling of clear skies rapidly decreases with altitude. Divergence usually occurs when air streams begin moving in opposite directions, to be specific, upper-level divergence occurs in the upper/ higher levels of the atmosphere, and it leads to rising air from lower levels of the atmosphere.

A clear sky cooling rate in the tropics decreases above 200hPa, as shown by the dark blue curve in Figure 2.1 (Hartmann et al., 2001). It eventually decreases rapidly, as shown by the lower portion of the curve with the reduction of temperatures. It should be noted that the cooling rate, and the relaxation rate both decrease since water vapor becomes ineffective at that level (Hartmann et al., 2001). Most of the radiative cooling occurs in clear regions. This region where the radiative cooling in clear skies decrease rapidly with altitude is where these anvil clouds usually develop. Due to the thermal dynamical processes involved in the development of these clouds, fog, can also form.

In the topographical regions of valleys, radiation fog also occurs due to the radiative cooling effect. Cold air always has the tendency of sinking since it is denser and will naturally sink into the valley due to gravity. Longwave radiation from solar heating will be lost to space from within the valley (shown by yellow arrows), and this causes temperatures to cool near the dew point. Once the temperatures reach that saturated dew point level, the excess water vapor is forced to condense out into low lying clouds or areas of fog as shown in Figure 2.2.

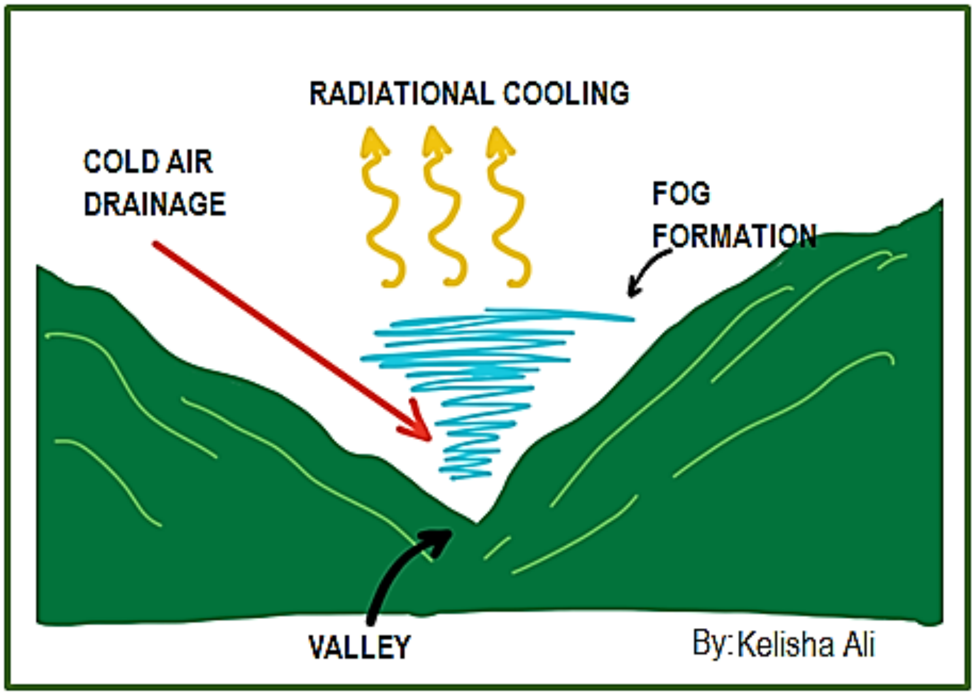


Figure 2.2 *Radiation fog occurring within a valley/Valley fog.*

Temperature inversions also play a significant role in fog formation (Temperature Inversion | Definition & Facts, 2021). Simply put, inversions are the vertical mixing of heat and moisture. A ground inversion can develop when air is cooled by contact with a colder surface, until it becomes cooler than the overlying atmosphere. This phenomenon can occur on clear nights when the ground is naturally cooled off rapidly by radiation. According to (Britannica, 2021), there is abundance of heat and moisture transport in the soil and vegetation with this natural cooling effect. This can lead to high levels of humidity and if the temperature surface of air drops below its dew point, fog may result.

Inversion/advection fogs are very prevalent due to a downward extension of a layer of stratus cloud, which lies just under the base of a low-level temperature inversion. Nighttime cooling may then lead to the formation of a stratus layer and build down to the ground to develop an inversion/advection fog, as shown in Figure 2.3. The appearance of this fog looks like wisps of smoke rising off the surface of the water.

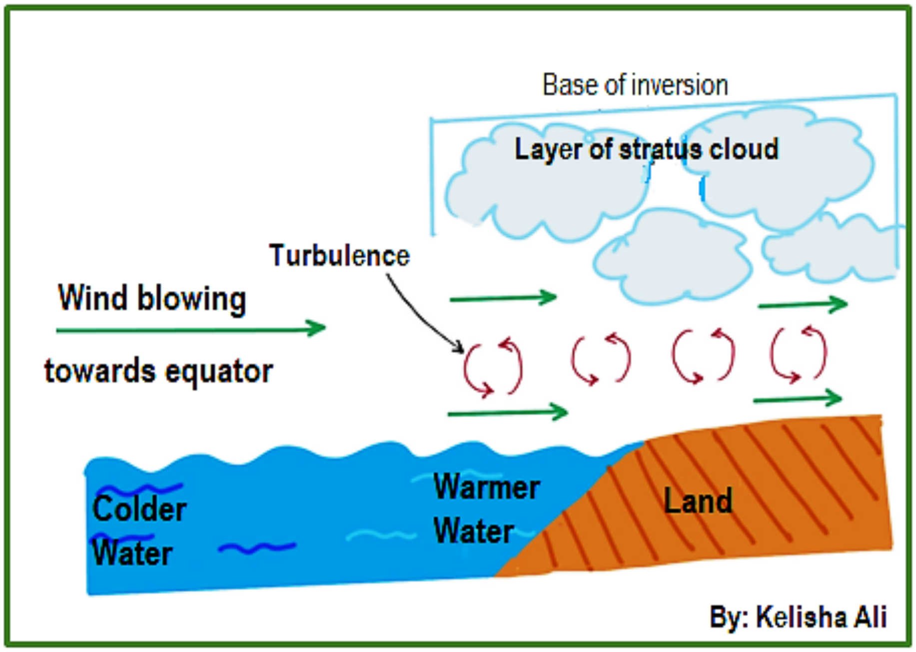


Figure 2.3 *Formation of Inversion/Advection Fog*

Another instance of inversion fog occurs along coastal regions where winds cause upwelling, leading to colder sea surface temperatures. The air which passes over the cold water becomes chilled and the degree of its relative humidity rises, it then becomes trapped underneath the inversion, cools to develop the dewpoint and fog forms.

In east-central Florida, we tend to have the same effect, because for most of the year, we tend to have a warm tropical climate. Cape Canaveral Space Fore Station Skid Strip (KXMR) is located just off the coast of the Atlantic Ocean, and due to its location, KXMR experiences Inversion fogs from time to time. On the other hand, if we are only focusing on location near the Atlantic coast, NASA Space Shuttle Landing Facility (KTTS) may not experience inversion fog as much as KXMR since it is located further inland.

Studies by the National Weather Service (n.d.) report advection and topographic effects can also lead to the formation of fog. Advection fog has similarities to radiation fog, as it is also a result of condensation. The difference between advection fog and radiation fog is that advection fog is caused by the horizontal movement of warm moist air over a cold surface, an example of this is warm moist air flowing over snow. Radiation fog on the other hand, is solely due to a reduction in the surface temperature. An easy way to distinguish between both is that advection fog can be characterized by its horizontal motion along the ground.

This study has put more emphasis on radiation fog at airports in east-central Florida. This is because, according to the (National Weather Service, n.d), radiation fog is most prevalent in fall and winter seasons. Even though fall and winter do not really change Florida's weather, climate, and atmosphere as much as in other states, we can still be affected by radiation fog. Radiation fog forms usually develop gradually overnight with the cooling of air near the surface as it stabilizes, and eventually reaches its saturation point. The fog initially develops at the surface, or just above the surface. It thickens as more air continues to cool. Any air that lies above the initial fog layer that cools will help to deepen the layer of fog. The more the volume of air in this lay, the more the fog will grow and extend upwards.

Fog development will flourish the most in sheltered valleys where there is little or no wind, as well as locations near bodies of water. This makes east-central Florida with numerous swamps, wetlands, lakes and ponds the perfect location for the formation of radiation fog. Fog can be easily disrupted with strong winds and these winds may also hinder the formation of fog. Radiation fog's appearance is usually very patchy and stays stationary in the location it formed until solar warming burns off the fog layer the next day.

Figure 2.4. *Radiation fog at Bengaluru Airport (The Indian Express, 2020)*

On January 22nd, 2020, Bengaluru Airport was able to have their first successful landing after a visibility upgrade. The layer of fog pictured in Figure 2.4 is radiation fog, which has been an issue at Bengaluru airport for proper landing and takeoff conditions. The success of an airplane landing was an accomplishment for this airport due to the harsh reality of radiation fog prohibiting normal operations. A statement issued by Bengaluru International Airport Limited read; "Radiation fog set in during the early part of the day, resulting in a rapid drop in visibility to 200 meters. During this time, IndiGo flight 6E-6389 from Lucknow made a successful touchdown at 0741 hours" *(The Indian Express, 2020, para. 1).* The authority mentioned, "with this upgrade, the South Runway can facilitate aircraft landing with a Runway Visual Range (RVR) as low as 50m and take-offs at 125m. Until now, the permissible visual range was 550m and 300m for landing and take-off respectively."

Apart from our focus on radiation fog, an unusual named category of fog is called "Super fog." Super fog can be found in both cities and forested areas, but its stems from the development of smoke. When smoke in the atmosphere and moisture from nearly saturated air mixes, super fog is formed. This moisture in the atmosphere is released from damp decomposing organic material, for example brushes, leaves and trees.

In addition, other types of fog of interest are ice fog, freezing fog and hail fog. Ice fog occurs when the air is so cold that water vapor in the air forms suspended ice crystals at and near the ground. Ice fog does not form in east-central Florida. Freezing fog occurs when the water drops the atmosphere remain in a liquid state, until they come in contact with a surface that they can bond with and instantly freeze. Therefore, any object that freezing fog encounters would immediately become coated with ice. Freezing fog can easily be identified in forests, where the trees and vegetation are all covered by a thin layer of frost. Lastly, the most unusual type of fog, is known as hail fog, and it usually occurs shortly after a heavy hailstorm. The cold balls of hail, which are large ice pellets, fall into moist, warm humid air near the surface. As the hail accumulates on the ground it causes the air just above the ground to cool and saturate, eventually reaching its dew point, where fog forms. Fog usually forms where the winds are light and under these conditions it is quite patchy and a shallow layer.

Within the aviation industry, inclement weather has been one of the leading causes of economic loss. The impact of weather on airport operations can range from airport closures, serious disruptions of flight schedules, to deadly accidents. The weather phenomenon: fog, contributed to one of the worst and deadliest accidents recorded in aviation history.

Review of Past Studies

On March 27th, 1977, the presence of intermittent fog at the Tenerife airport resulted in low visibility for the pilots of KLM Flight 4805 and Pan Am Flight 1736. The fog led to such poor visibility that the aircraft parked on the airport pavement and those on the taxiway were not visible by the air traffic control tower. Due to improper communication, confusion of both pilots over the intercom, as well as the presence of thick fog, there were 583 fatalities resulting from the collision (Ziomek, 2020). Apart from deadly accidents, both airport closures and flight delays can lead to declining economic performance at airports.

Bergot (2021) explained that with consideration of the overall observation, simulation, and predictability of fog, the impact of fog within the atmosphere has significantly increased over the past few decades due to the rapid increase in airline, maritime and vehicular traffic. Specifically focusing in on the realm of aviation, the financial implications can be immense over time. An example which was highlighted in this article, were the losses that Gandhi International Airport in India. They sustained a value of financial losses nearing 3.9 million USD for airlines, between the years of 2011 and 2016.

Observations have shown that the process of forecasting of fog may be very incomplete, and there is a high level of difficulty with accurately predicting a precise amount of fog in the atmosphere at a certain time. Kulkarni (2019) urges that better forecasts would help to guard against these losses that are associated with the delays at airports. Another source indicated this same concern of the economic loss of the aviation economy in India. Kulkarni (2019) indicated that at the Indira Gandhi International Airport

(IGI), one of India's busiest airports, there are more than 900 flight operations per day. In order to maintain maximum profitability from airport operations, airlines should be operating at full capacity with no diverted, delayed, or cancelled flights. Any interruption may lead to significant financial losses to the aviation industry. It is important to undertake a quantitative study of the estimated losses, therefore, to reduce the economic losses that aircraft face due to fog, this study has aimed at developing a model to better predict these fog phenomena at different airports.

It is important to undertake a quantitative study of the estimated losses due to fog. Table 2.1 displays the average cost of a disturbance in normal flight operations. A similar study, undertaken for the first time in India in 2019, aimed to evaluate the impact of dense fog at IGI Airport on economic losses that occurred during the winter season between 2011 and 2016. Table 2 shows that the economic cost of fog is very volatile and highly depends on whether flights are delayed, diverted, or cancelled. In the study conducted by Kulkarni (2019), they considered a flight capacity of 150 passengers and an average ticket cost of \$52.23USD for domestic flights. For international flights, they considered a carrying capacity of 350 passengers (e.g., Air India Boeing 777-300 ER from London to Delhi) with an average ticket cost of 1343.28 USD per person. The breakdown of charges for different parts of the flight operations for the domestic and international sectors was obtained from India's Ministry of Civil Aviation and the Center for Asia Pacific Aviation (CAPA) India.

In a summary of their findings, a total of 653 hours of dense fog between 2011 and 2016 at IGI set back the airlines to incur economic losses of approximately 3.9 million USD, which was 248 million Indian rupees. It should be noted that between the current time and when that study was conducted between 2011 and 2016, the value of money has appreciated. The values identified in Table 2 are in 2016 US Dollars, and in today's economy, these diverted, delayed, and cancelled flights would cost much more than that listed in the table.

Table 2.1

The average impact to the Indian Economy due to diverted, delayed, and cancelled flights

BREAKDOWN OF CHARGES FOR INTERNATIONAL FLIGHT OPERATIONS		
CONTENTS	**DOMESTIC FLIGHT (USD)**	**INTERNATIONAL FLIGHT (USD)**
A SINGLE DIVERTED FLIGHT		
Extra Fuel Cost	$2985.00	$7462.68
Landing Charge	$179.10	$2985.00
Ground Handling	$671.60	$1343.28
RNFC	$68.95	$68.95
Parking Charge	$194	$2238.80
Food Charge	$1119.40	$2611.94
Accommodation Charge (in the case of night flight)	$2238.80	$5223.88
Total Cost	**$7456.85**	**$21,934.53**
A SINGLE DELAYED FLIGHT		
Hold Fuel Cost	$1493.00	$2985.07
RNFC Charge	$68.95	$68.95
Food Charge	$1119.40	$2611.94
Total cost	**$2681.35**	**$5665.95**
A SINGLE CANCELLED FLIGHT		
Returning all the Price of the Ticket + Compensation/person	$8171.64	$470,932.83

Source: Atmosphere (2019)

In addition to studies about the financial costs of weather delays, other studies have dealt with the FSI to help with the prediction of fog at airports. Multiple studies were conducted at Korea's Incheon International Airport (IIA). At IIA, there has been a significant amount of flight cancellations due to fog. This study also cited Holtslag et al. (2010), which stated the likelihood categories of FSI. To reiterate, an FSI < 31 states a high likelihood, 31 < FSI < 51 signifies moderate likelihood, FSI > 55 implies low likelihood. They evaluated the FSI at IIA during the period of June 1st, 2011, to December 31st, 2011, and compared it with observed fog occurrence. A similar analysis was done for another place, Osan, which was located 100 km from IIA, to compare the two.

The mean FSI values which were recorded during the fog periods were 34.6 at IIA and 30.1 at Osan. The average FSI for the whole period including both fog and non-fog periods was higher than 31 at IIA, indicating that the threshold of 31 suggested by Holtslag et al. (2010) was inappropriate at IIA. Most of the fog at this airport was sea fog, which is not the focus of my study, which is radiation fog. However, this study at IIA goes to show that the threshold of the variability of fog suggested by Holtslag et al. (2010) be challenged.

A study of morning radiation fog by Trigg (2001), identified fog as "a cloud with its base at the Earth's surface, reducing horizontal visibility to less than one km." Radiation fog, if personified, can be termed as "stubborn." This means that the forecasting of radiation fog may be problematic. This is mainly because radiation fog usually forms in a small area, and then it either increases or decreases with no warning. This is makes it is difficult to predict. Radiation fog also has the characteristic of being 1-Day persistent, i.e., if it occurs on one day it is more likely to occur on the subsequent day.

Chapter 3

Methodology

Population and Sample

Population. The target population consisted of surface observations at six airports in east-central Florida. The accessible population will consist of surface observations at airports within 50 nautical miles of the radiosonde station located at Cape Canaveral Skid Strip (KXMR). The airports investigated are Orlando International Airport (KMCO), Melbourne International Airport (KMLB), Stanford International Airport (KSFB), and Patrick Space Force Base (KCOF). Data extended from the year 1973 to the present day.

Sample. The sample data is comprised of surface observations at the airports of focus, and complementary FSI readings were obtained using simple random sampling from available and accessible archival data. In this case, the simple random sampling method entailed randomly selecting FSI readings from the accessible population data set, thereby ensuring that each reading has an equal chance to be included in the sample data set. This process was essential to perform this sampling method as it served to highly negate biases in the selection process.

Power Analysis. Logistic regression feeds a linear regression model into a sigmoid function so that it predicts probabilities between 0 and 1. It then uses maximum likelihood estimates rather than ordinary least squares to estimate model coefficients, which are the combination of coefficient values that produce the best overall fit for the model. Therefore, the reliability of the estimates tends to be low if there are only a few cases (participants) for each combination of scores on predictor variables (Warner, 2008). To mitigate for low reliability of the estimates related to obtaining a sample with a low number of events per variable, Peduzzi et al. (1996) suggested a minimum sample size at least 10 times the number of independent variables in the model. Thus, the current study's sample size (N=297) exceeded this minimum requirement.

In logistic regression, effect size (*ES)* is synonymous with the size of the treatment's effect on the odds ratio (*OR*). For example, a finding of an *OR* of 1.00 between the occurrence of fog and no fog would imply that FSI had no effect on the DV. This would mean that there is an equal probability of fog or no fog. Therefore, the change in odds above or below 1.00 produced by IV represents its *ES* with respect to the dependent measure. The results of the logistic regression determined that the *ORs* of fog occurring was 2.02 when the Probabilistic FSI was less than 40 and 2.08 when the Theoretical FSI was less than 43. These *ORs* were used to calculate the actual power of the sample size at α = .05 using the computer program *G*Power 3.1.9.7* as presented in Table 3.1.

Table 3.1

Power Analysis for Hypothesized Relationship in Sample at α = .05

Model	Statistic	Effect Size *(ES)*	Number of Predictors	Estimated Power
Logistic Regression				
Probabilistic FSI <40	OR = 2.02	ΔOR = 1.02	1	0.99
Theoretical FSI <43	OR = 2.08	ΔOR = 1.08	1	0.99

Note. N = 297.

The estimated power also was validated by consulting Hsieh (1989), who demonstrated power in logistic regression based on varied sample sizes, alpha levels, overall event proportions, and odds ratios at one standard deviation above the mean. The estimated power was found to be consistent in both methods. As noted in Table 3.1, given the *OR*s, the statistical power related to finding a significant model in the logistic regression was 0.99. Thus, the power achieved by the study's sample exceeded .80 for, which was the minimum posited a priori.

Procedures.

Research Methodology: The study employed a cause-type ex post facto design. This design was appropriate because the effects on the dependent variable, which is group membership, have already occurred and my study determined probable research factors, or causes (potentially FSI, probabilistic FSI, and 1-Day Persistence), for group membership. The group membership variable consists of two pre-existing groups: east-central Florida airports that experienced radiation fog, and airports that did not experienced fog. Ex post facto research (also called causal-comparative or retrospective research) is a systematic, empirical inquiry in which the scientist does not have direct control of independent variables because their manifestations have already occurred or because they inherently do not have manipulability. Inferences about relations among variables are made without direct intervention. In this study, the independent variables of FSI and 1-Day persistence of fog were not controlled.

Description of independent and dependent variables: The study explores factors that could reliably predict the occurrence of fog-related weather phenomena. Thus, the dependent variable is the occurrence of radiation fog, while the independent variables are the FSI and 1-Day Persistence of fog.

Study implementation: The primary dataset provided by the 45th Weather Squadron included upper air (radiosonde) data and the FSI for KTTS. I collected surface observation data from the varying airports from the AWOS/ASOS website provided by Iowa State University's Environmental Mesonet. I created my sample by randomly sampling the raw data of FSI values at KTTS. I chose a random sample of 300 data points based on the sample size posited a priori, as discussed earlier.

Threats to internal validity: "Internal validity refers to the inferences about whether the changes observed in a dependent variable are, in fact, caused by the independent variable(s) in a particular research study rather than some extraneous factors." (Ary et al., 2010, p272) Ary et al. (2010) identify 11 threats to internal validity: history, maturation, testing, instrumentation, statistical regression, selection bias, experimental mortality (attrition), selection-maturation interaction, experimenter effect, subject effects, diffusion, and location. The inherent weaknesses of *ex post facto* design in this study, were location, selection, history, and mortality.

Location is a threat to internal validity when a change in the location of the study takes place and could influence the final results of the study (Ary et al., 2010). Location was considered a weakness due to the variability of sea fog or advection fog that may be more prominent in one airport's location than the next. To control for location threat, the study location was held constant by using only the FSI for TTS and limiting the study of data from six airports in east-central Florida.

Selection bias refers to the threat posed by nonrandom factors that could influence the selection of participants and result in differences between the treatment and control groups even before the experiment begins (Ary et al., 2010). This threat is identified as a study limitation and was mitigated by random sampling of the accessible population data provided by the 45th Weather Squadron.

History was determined to be a threat when considering that most of the archival data could be impacted by changing weather phenomena over the years, though extreme, examples could include global warming, tectonic plate shifting, earthen axis shifting, etc. However, there are simpler explanations, such as weather data collection technology progression or any circumstance that could have happened at the time that the weather data was collected. These variables are called threats because unless they are controlled, they can produce an effect that can be mistaken for the effect of the treatment and hence provide an alternative explanation of the study's findings; this can raise doubts about the accuracy of the findings. History was a threat in this study due to the changing weather codes for fog over the 50 years of data collection. I mitigated for this threat by limiting the data for the occurrence of fog to just four codes (BR, FG, BCFG, and MIFG) as outlined in the data analysis that follows.

Mortality typically occurs when there is a differential loss of participants from the group affects the dependent variable (Ary et al., 2010). The loss could lead to different outputs because if one specific type of participant is lost, then the proportion of the other participant's effects on the final results could increase. In this study, mortality was factor perhaps because an ASOS was not working that day, and that occurrence caused a loss of surface data at one airport. Thus, to mitigate for these threats, I eliminated the data for all six airports on that particular day from the study.

Data Analysis.

Data analysis was accomplished through descriptive and inferential statistics. Descriptive statistics enlisted the collection, presentation, and description of numerical data (Ary et al., 2010). The statistics are utilized in the form of graphical presentations to represent the response of the questions asked in the research material to better illustrate the data through ex post facto. In particular, data points regarding FSI and the occurrence of fog at various airports were implemented in the representations including the average FSI and the frequency. Descriptive and inferential statistical analyses were employed using data with both fog and non-fog days, with the use of ASOS/AWOS data. The AWOS findings included data such as ceiling and sky conditions, visibility, temperature, dew point, altimeter setting and wind speed, gusts, and direction. ASOS data additionally provided the type and intensity of precipitation (rain, snow, freezing rain), and obstructions to visibility such as fog and haze (Skybrary, 2021).

Inferential statistics utilized a logistic regression, performed using JMP software, to determine the probability of fog occurring at east-central Florida airports to develop a null model, meaning no independent variables in the model. This set the baseline for determining whether the independent variables were significant factors in the formation of radiation fog. A logistic regression was also to determine the probability of fog with FSI in the prediction model (full model). Logistic regression addressed the problems of nonlinearity, nonsense prediction, non-normality, and lack of homoscedasticity in the variables because it uses maximum likelihood estimates.

The data supplied by the 45th Weather Squadron included FSI determined at KTTS over the period 1973 to 2020 and the presence of fog or no-fog at this location. There were 1355 data points, which represented radiation fog days since all the sea fog days had been removed from the data set. Microsoft Excel was used to randomly sort the data. The "RAND" function was used where it returns an evenly distributed random real number greater than or equal to 0 and less than 1. The random numbers generated were then sorted from largest to smallest, and the first 300 data points were selected. These 300 random data points were sorted according to the day, month, and year. Raw meteorological (METAR) data was obtained from the Iowa State University website for the selected five remaining airports, which were DAB, COF, MCO, SFB, and MLB.

The coded data types focused on weather (WX) and visibility. The software only had the bandwidth capability to run reports for certain periods at a time. This was due to data sets being too large to download or open all at once. The 300 random sample data points, their METAR observation data, retrieved from the Iowa State University website database, were manually examined to check for the codes which indicating radiation fog occurred. These particular WX codes are as followed in Table 3.2.

Table 3.2

The types of fog associated specifically with the category of Radiation Fog

Weather Code	**Type of Fog**
BR	Light Fog/ mist
FG	Fog
BCFG	Patchy Fog
MIFG	Shallow Fog

These types of fog were recorded between 1000Z to 1500Z to avoid the occurrence of advection/sea fog, as the radiation fog was most prominent at this time. This means that if there was a singular fog event between this timeframe on the sample days, then it was counted as viable data. KMLB had no data or METAR reports in the first three sample points of data for the year 1973 to 1987. Therefore, three data points were omitted due to missing data, and this reduced the effective dataset to 297 sample data points.

To check for 1-Day persistence (ODP), the day following the sample observations was searched for fog occurrence. Four distinct outcomes could be recorded: 1. no fog on the sample day and no fog on the next day, 2. no fog on the sample day and fog on the next day, 3. fog on the sample day and no fog on the next day, and 4. fog on the sample day and fog on the next day. The latter observation was indicative of ODP. All the non-fog and fog days were compiled into a table, which meant that if any of the six airports reported fog, then it counted as a fog day for ECFA. These were recorded in binary and represented as 1 – fog occurred and 0 – no fog occurred alongside the FSI readings for KTTS. In addition, Theoretical and Probabilistic models were based on the middle of the range of the FSI values. The midrange values were FSI $\geq$ 43 for the theoretical and FSI $\geq$ 40 for the probabilistic model, as discussed in Chapter 4.

Chapter 4

Results

Introduction

This chapter is organized and presented in two main sections. The first section presents descriptive statistics and contains the results of the analysis performed on the sample data. The second section presents the results of the inferential statistics consisting of preliminary and primary analyses of the sample data. In the preliminary data analyses, I addressed invalid and missing data, outliers in the sample and tested the sample data for compliance with the assumptions for the logistic regression strategy employed in the analyses. I addressed the relationship between the independent variables and the dependent measure in the primary data analyses.

Descriptive Statistics

This section analyzes the sample data using descriptive statistics. Specifically, the original METAR report data, after any missing or incomplete data was removed, was analyzed for further information to provide consensus on any potentially impactful or insightful statistics. Notably, for the varying analyses performed in the following sections besides descriptive statistics, it is most important to understand the variables that are viable for the study. There were several factors of interest in performance of the descriptive statistics.

First, the number of fog and non-fog days at any given sample airport or the collective group of airports in east-central Florida (ECFA) to include supporting or other related statistics beyond a simple count. This is due to a potential notion that one airport may have characteristics causing high percent chance of fog or non-fog days over another airport within the group. This would require further research and attention as it is not completely within the scope of this research, but it may be a factor of interest in future studies branching off this research, as addressed in Chapter 5.

Second, the occurrence of specific weather codes. The METAR report data generated by the Iowa State University, archival data retrieval site, is expansive. It includes every occurrence of any individual or combined weather code as a separate instance meaning fog (FG) and rain (RN) may be considered separate and also combined (FG RN) into a weather code that causes additional data to populate the query produced by the Iowa State University site. With so many weather codes to go through, it would be helpful to know what weather codes exist, which matter, and understand how often each occurs. This proved valuable to whittle down the list of impactful weather codes indicating radiation fog between the accepted timeframe from 1000Z to 1500Z.

A final notable descriptive statistic, though not the last, is visibility information. Visibility correlates heavily to this research in a few ways. The social and economic impact caused by fog at each airport across the country is immense, as noted in the 2019 economic impact analysis presented in Chapter 3. The monetary value and time lost to fog related events causing delays to scheduled commercial airline flights is extremely problematic. However, it should be noted that not all radiation fog events are the same. Visibility is the key factor when discussing differences between weather codes. While there can be a variance in the range of visibility presented by each weather code, the average at least provides a measurement as to level of threat present to cause delays within a given weather observation area or specifically an airport.

Two, each weather code occurrence will produce a differing average of fog. Since the reduced sample size of N = 297 is sufficient according to the power analysis, it can potentially be further assumed that the correlated visibility to the individual weather code, given enough occurrences, is accurate. That being stated, it can also be cross-referenced with meteorological reports on what the observed average or range of weather code visibility readings are. Given the specific area of east-central Florida, it will vary, but if proximal in mile, nautical mile, or linear feet numerical value, it would be of enormous assistance for the application of this research in daily scenarios. Airports, weather observation stations, and other benefactors and practitioners would be able to state warning level of each FSI, ODP, or other predicted radiation fog variables as discussed in Chapter 5.

The data set was examined to discover and discuss the available factors impacting the study. Of note, Tables 4.1 and 4.2 present the most basic data needed to properly analyze the results of the study and potentially find further required research. The fog and non-fog days count is required to complete the rest of the inferential statistics, logistic regression, and contingency analysis. It also provides insight into the potential frequency of radiation fog, given that the data is occurring from 1000Z to 1500Z as recommended by the 45thWeather Squadron.

Table 4.1
Fog Stability Index at KTTS

Fog at ECFA?	***N***	***M***	***SD***	***Range***
Yes	173	38.6	11.4	7–77
No	124	44.1	14.5	12–81
Total	297	40.9	13.0	7–81

Note. Of the overall sample size of $N = 300$, 3 data points had missing information. The FSI ranged from 7 to 81 with higher scores indicating a lower probability for the occurrence of fog at ECFA.

Table 4.2 ***Count, Percent, and Average Fog Occurrence***

Occurrences of Fog Days by Airport					
TTS	**DAB**	**MCO**	**COF**	**MLB**	**SFB**
72	121	90	46	94	88
Mean Fog Days by Airport					
TTS	**DAB**	**MCO**	**COF**	**MLB**	**SFB**
24.0%	40.3%	30.0%	15.3%	29.6%	29.3%
Mean Occurrence of Fog at East-Central Florida Airports (ECFA)					
28.10%					

Note. N = 297. Of the sample airports included in the data set, ECFA, the mean occurrence of fog is approximately 28 percent for any given day included in the random sample per airport. DAB had the highest occurrence of the accepted weather reporting codes in the METAR reports at 121 fog days or approximately 40 percent.

What is more interesting for future studies is the count and percent of fog days per individual airport as part of the collective group of east-central Florida airports in Table 4.2. DAB presenting the most occurrences at approximately 40 percent while COF had the most drastic difference at approximately 15 percent over the 297 days, shows a potential difference between airports. This is especially relevant given that the mean percent chance for a fog day at all ECFA was approximately 28 percent. This shows the disparity between airports, but the range is still acceptable for combining the airports for the FSI portion of the analysis. The total number of fog events in the sample was 511 because fog occurred at more than one airport on any given day. Figure 4.1 depicts the contribution of each airport to total fog events; this is higher for DAB and lower COF. Perhaps location, the geographic region, environmental aspects, etc., could further explain the differences as discussed in Chapter 5. It may also point to some weather occurrences not being radiation fog due to such locational circumstances. Again, this is not data to be collected or further studied as a part of the scope of this research; however, it may be of interest to future studies or practitioners to delve into.

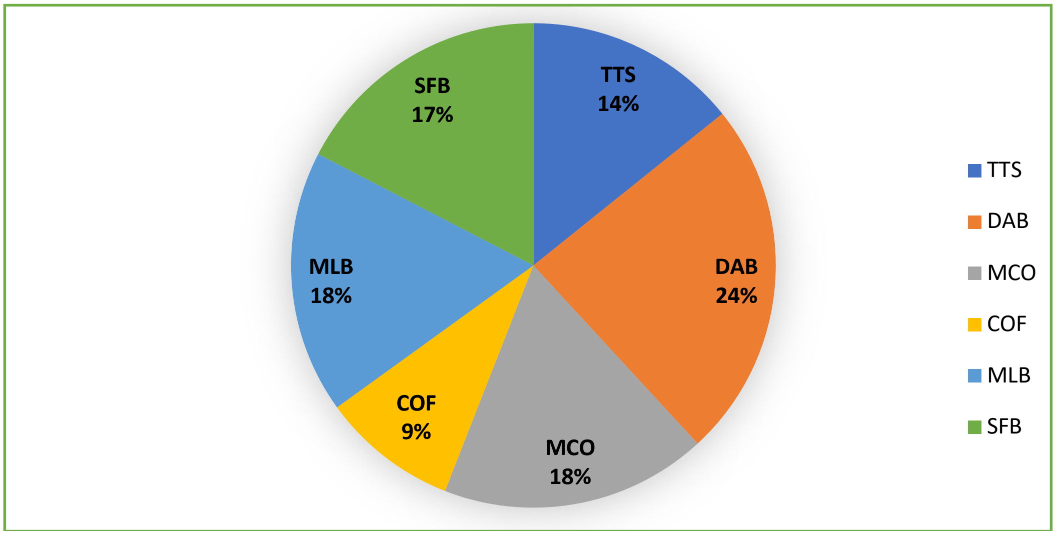


Figure 4.1 – Individual Airport Share (%) of Fog Events in East-Central Florida
Note. Daytona Beach proved to have the highest share of fog events, at 24% of the N = 511.

The analysis based upon weather codes and their corresponding visibility is potentially immense enough to warrant further study and implementation into practical scenarios using FSI and ODP as fog event predictors. As seen in Table 4.3, there is a clear, commonly reported weather code: BR. However, there are many variations of fog weather codes. Whether they include rain, snow, sun, hail, or otherwise, the 45th Weather Squadron and, my own analysis, concluded that it was best to exclude all weather codes besides the four listed in Figure 4.2. BR was by far the most common, with FG, MIFG, and BCFG falling short in terms of reported count. The factors that attribute to this difference are currently unknown but may include the individual reporting differences between weather analysts.

Table 4.3 ***Weather (WX) Codes Occurrences***

WX Codes	Occurrences
BR	3443
FG	809
MIFG	462
RA BR	380

BCFG	188
BR BCFG	74
MIFG BR	24
BCFG BR	12
TSRA BR	6
TS BR	2
VCTS +RA BR	2
MIBCFG	2
RA FG	2
DZ BR	1
VCTS -RA BR	1
VCTS RA BR	1
BR MIFG	1
RA BCFG	1
VCSH BR	1
Grand Total	**5412**

Note. Accepted WX Codes per given METAR data in the Iowa State University archival database were far too wide-ranging and could not all be considered as radiation fog with total assurance. Therefore, per initial instructions from the 45th Weather Squadron, I used BR, FG, MIFG, and BCFG as indicators for radiation fog and cut the list from 19 WX codes to four. Furthermore, the listed weather code occurrences are based on the hourly report during the accepted daily timeframe of 1000Z to 1500Z and were to show the most common types, BR at 3443, FG at 809, MIFG at 462 occurrences, and so on to help with predictability.

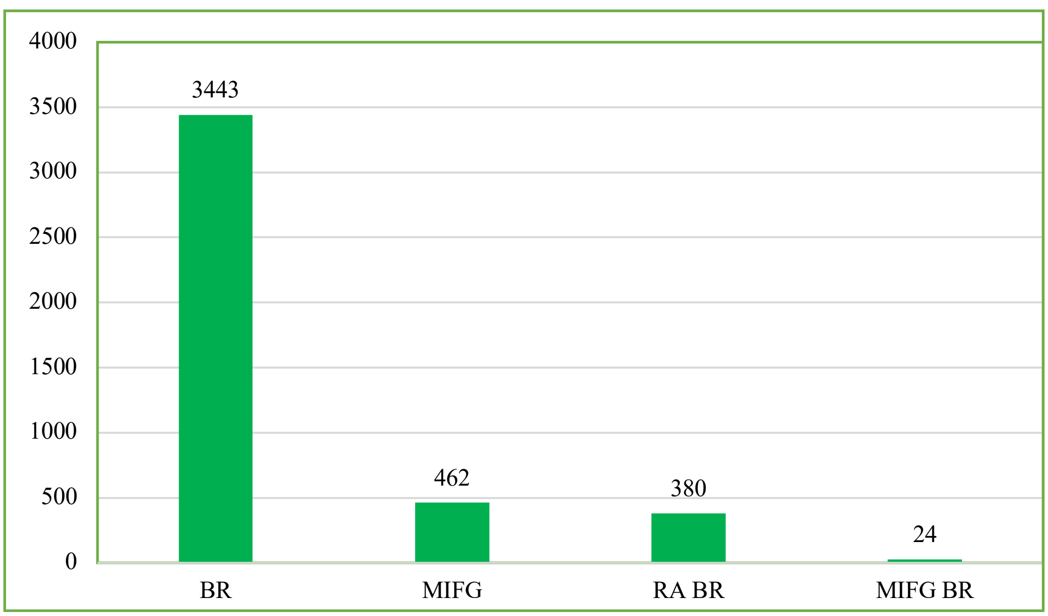


Figure 4.2 ***Graph of Accepted Weather Code Counts***

Note. The graph shows WX codes from the list reduced by the 45th Weather Squadron. Fog, BR, was the most prominent reported on an hourly basis for the accepted daily timeframe of 1000Z to 1500Z as prompted by the 45th Weather Squadron to ensure the reported weather was indeed radiation fog and not simply sea/advection fog.

Table 4.4 presents the mean visibilities associated with weather codes in the sample. In conjunction with visibility, each weather code demonstrates how impactful the event may be to airport operations. FG is by far the most dangerous radiation fog event to occur on average due to the 0.38-mile visibility. It is very likely to cause multiple delays for departures when occurring. BR stands as a potential concern; 3.89 miles of visibility is considerably low, compared with the overall reported weather code average of 3.91 miles, and may present occasions lower than the average. This could also present delays and dangerous situations. Finally, MIFG and BCFG are fairly low threat radiation fog events, although, they are still useful to predict in case of an event where they have a much lower visibility than the average.

Table 4.4 ***Mean Visibility by Weather Code***

WX Codes	**Mean of Visibility (mi)**
BCFG	8.84
MIFG	8.15
VCSH BR	5.00
DZ BR	5.00
MIFG BR	4.54
RA BR	4.07
VCTS -RA BR	4.00
VCTS RA BR	4.00
MIBCFG	4.00
BR	3.89
BR BCFG	3.79
TSRA BR	3.15
TS BR	2.75
BR MIFG	2.50
RA BCFG	2.50
BCFG BR	2.22
VCTS +RA BR	1.25
RA FG	0.50
FG	0.38
Grand Total	**3.91**

Mean Reported Visibility	
Miles	Feet
3.91	20,630

Note. Visibility for each hourly METAR report per weather code was considered due to the reasoning for this study. As visibility is lowered, the closer to 0 is worse visibility per mile, the more negative impact on airline and airport operations there are. With FG being the second most common weather code occurrence per hour during the accepted time period and having the lowest visibility, it is considered a hazardous event to occur. MIFG is fairly low in terms of mean visibility, but further looking at the range of the code may prove to be useful in case of fog events that are unusual to their weather code (i.e., more than 2.5 standard deviations away).

Figure 4.3 depicts the airports that experienced 1-Day persistent fog in the sample. There were a total of 10 days in which fog persisted for more than one day. The airport with the most days of 1- Day persistence was DAB, where there were four sample observations of this phenomenon.

Figure 4.3 - Airports and counts where 1-Day Persistence occurred

Note: There were 10 1-Day persistence events. Two events occurred at different airports on the same day. Thus, there were a total of nine 1-Day Persistence data points.

Inferential Statistics

Preliminary analyses. Prior to formulating the results needed to answer the prompted research questions, the data set was examined to ensure that it did not include invalid or missing data as well as outliers that could unduly influence the results of the study. In addition, the final data set was examined for compliance with the assumptions for logistic regression. The results of these preliminary analyses are discussed below.

Invalid and missing data analysis. The data was examined to ensure validity for conducting the logistic regression analysis as well as for missing and incomplete METAR information in the generated report. In a valid data set for logistic regression, each independent variable should include a minimum of one cell frequency and no more than 20% of cell frequencies should be less than five (Tabachnick, 2013). Validity for the independent variables was confirmed by contingency analyses. ODP was found to be invalid for a logistic regression, as shown in Appendix B. Furthermore, missing and/or incomplete data likely occurred due to METAR reports not being able to be or not accurately being captured by weather observation stations. Of the 300 initial cases, three cases were deleted due to systematically missing information.

Outlier analysis. Outliers are extreme data points that are inconsistent with other data points and should be examined because of their potential to produce results that are not representative of the relationships in the remaining data. Outliers can be labeled either as contaminants or rare cases. The outlier analysis returned seven results that could be rare cases, such as those beyond 2.5 standard deviations, as shown in Figure 4.4. Therefore, the analyses were run with outliers included and excluded.

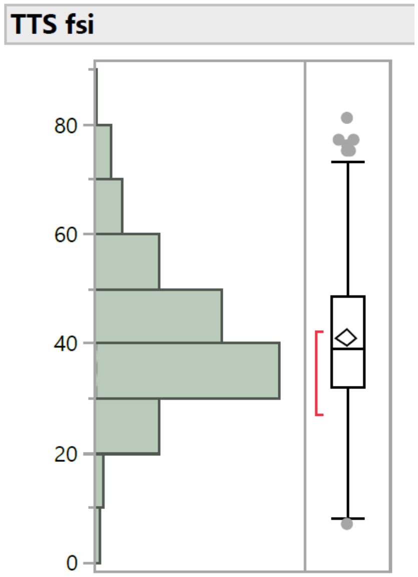


Figure 4.4. Outlier Analysis Box and Whisker Plot

Logistic regression assumptions. The independent and dependent variables were tested for compliance with the assumptions of logistic regression. These assumptions included: a dichotomous dependent variable, mutually exclusive categories on the dependent variable, independence of scores on the dependent measure, and correct specification of the hypothesized model. Although not required for a logistic regression, the absence of multicollinearity in the independent variables, the absence of outliers in the solution, and the linearity of the logit were also addressed because these issues could indicate a poor predictive model.

Dichotomous DV. The requirement for a dichotomous dependent variable was obtained through the METAR data: each airport either had a day with fog or without fog. If fog occurred at one or more of the six ECFA, then it was counted as a fog day. As presented in the discussion of the resolution for missing data section, days where METAR data was either not available or not complete were excluded from the final data set. Therefore, the requirement for a dichotomous dependent measure was fulfilled.

Mutually exclusive categories on the DV. In logistic regression, categories in the dependent measure, which was the fog or no-fog days, are assumed to be exhaustive and mutually exclusive. Therefore, each day should be a categorized as one group or the other, but not both. The random sample days in the study were either assigned to the "Fog" or the "No-Fog" group based their METAR report information between 1000Z and 1500Z, prompting one of four meteorological weather code identifiers: BR, FG, MIFG, and BCFG. Days, where one of the sample airports' METAR reports presented one or more accepted fog event in the timeframe, were assigned to the "Fog" group, and days, where all airports reported zero fog events in the given time frame, were assigned to the "No-Fog" group, thus fulfilling this assumption.

Independence of scores on the dependent measure. According to Tabachnick (2013, p. 445), "Logistic regression assumes that responses of different cases are independent of each other. That is, it is assumed that each response comes from a different, unrelated case." For the current study, the data collection instrument was hosted electronically, and the host site captured separate METAR reports for each sample airport. A review of the data did not reveal any anomalies or duplications. Furthermore, random sampling was used which would mitigate the risk of such. Therefore, it was assumed that the data for each variable associated with each case were unrelated, and therefore this assumption of independence was met.

Correct specification of the model. The assumption that the hypothesized model is correctly specified also is required in logistic regression analyses. This means that the model should only include independent variables that are relevant. The inclusion of FSI in the hypothesized model was based on prior research and theory, as discussed in Chapter 2.

As Warner (2008) recommended, a baseline or null model was developed for the data in the absence of the independent variables. The chi-square test for the fit of the null model was then compared with the fit of the hypothesized model. Because the chi-square test for the fit of the model that included FSI produced a significantly better fit than the null model, it was deduced the model was correctly specified. This is summarized in Table 4.5.

Table 4.5 ***Full to Null Model Comparison***

Model	-Log Likelihood[a]	*df*	χ^2
Null[b]	201.80		
Full[c]	195.47		
Difference	6.34	1	12.67

Note. N = 297. p (χ^2) = 0.004*
[a]Log Likelihood (LL) indicates the agreement between the probabilities of group membership generated by the logistic regression model and the actual group membership within the sample. Larger absolute LL values represent a worse model fit. $\chi^2 = -2(LL_{\text{null model}} - LL_{\text{full model}})$. [b]The null model represents the baseline model without information about the predictor variables. [c]The full model represents the hypothesized model with FSI entered as a continuous IV.

Absence of multicollinearity in the independent variables. Multicollinearity can occur if the independent variables in a regression model are highly correlated. The presence of multicollinearity could result in unstable regression coefficients that are associated with large standard errors (Cohen et al., 2003), whereas the absence of multicollinearity could help ensure that the model is correctly specified and that any redundant variables are removed from the model. Originally, ODP was analyzed to test the potential for it to be used as an accurate predictor of fog events. However, due to ODP not being correctly specified and reducing the IVs to one, it and the absence of multicollinearity in the IVs do not apply.

Absence of outliers in the solution. The solution of a logistic regression model is the predicted probability of each case belonging to a specific group. If a model contains several cases that are poorly predicted then this could be indicative of a poor model fit (Tabachnick, 2013). For example, if the hypothesized model predicts that a day that is actually in the “No-Fog” group has a high probability of being in the “Fog” group then the case would be considered an outlier. Following Tabachnick’s (2013) recommendation, I examined the hypothesized model for outliers by generating studentized deviance residuals (Appendix C). None of the residuals were outliers, and therefore I deduced that the model provided a good fit of the data.

Linearity of the logit. Logistic regression assumes a linear relationship between continuous independent variables and logit (natural log of odds) of the dependent variable. The logit is the function of the predicted probability of the dependent variable that is linearly related to the independent variables (Cohen et al., 2003). The assumption of linearity of the logit is violated when the inclusion of interaction terms between a continuous independent variable and its natural logarithm in the model is statistically significant (Tabachnick, 2013). Under these circumstances, the continuous independent variable should be transformed. As noted in the section on data set modifications, the continuous independent variables included in the hypothesized model were split into dichotomies and expressed as binary data. Because the linearity of the logit cannot be violated by binary data, the assumption of linearity of the logit was upheld in the Probabilistic and Theoretical models.

Data set modifications. A modified version of the final data set (N = 297) was created for the purposes of interpreting the results of the logistic regression analyses as well as comparing the probabilistic FSI (X_p) to the theoretical FSI (X_t). Splitting the independent variable into dichotomies assisted in the interpretation of the odds ratios in the logistic regression analyses. Thus, the likelihood of fog occurring for one dichotomy could be compared directly to the other, rather than comparing odds ratios over a range of values that were more difficult to understand. For example, as shown in Table 4.8, the independent variable X_t = Theoretical FSI was split at the variable's midrange reading of 43.

Based on theory, FSI readings below 31 are considered a high probability, and readings above 55 are considered a low probability for the occurrence of fog. This dichotomy facilitated a comparison of odds ratios between reported fog days with FSIs below 43 and those above 43. In our sample, the probabilistic FSI mean and median values were approximately 40. Therefore, I used this value to split the X_p = Probabilistic FSI into values greater than or equal to 40 (coded as 0) and values less than 40 (coded as 1).

Primary analysis - relationship between the IV and the DV

The first objective of this study was to determine which independent variables (X_t, X_p, and ODP) were reliable predictors for the dependent variable, the occurrence of radiation fog at ECFA.

Full model. Following Warner's (2008) recommendations, I developed a baseline or null model by regressing the dependent variable in the absence of the independent variables. The overall goodness of fit of the null model was then compared to that of the full model with FSI as the IV. A measure that can be used to assess the overall goodness of fit of the logistic regression model is the log-likelihood (LL) function, which is comparable to the sum of the squared residuals in multiple regression (Warner, 2008). Furthermore, the chi-square statistic, which is the difference between -2LL for the full model and -2LL for the null model, should be large for the full model to be judged statistically significant.

As reported in Table 4.6, the full model was statistically significant, $\chi^2(1) = 12.67$, $p = .0004$. In addition to the chi-square statistic, Cohen et al. (2003) recommended reporting the Pseudo-R^2 (R_L^2) for logistic regression, which is the analog for R^2 in multiple regression. This was calculated as $R_L^2 = .0314$, which is provided in the general table note in Table 4.6. The reader is cautioned that the interpretation of R_L^2 is not a proportion of variance accounted for by an independent variable as in multiple regression, but rather the gain in prediction obtained from adding variables to a model. Therefore, the full model provided a predictive gain of approximately 3.14% over the null model ($R_{Lfull}^2 = .0314$, $df = 1$).

Table 4.6
Significance of the Simultaneous (Full) Model

Model	- Log Likelihood[a]	*df*	χ^2
Null[b]	201.80	0	0
Full[c]	195.47	1	12.67
Difference	6.33		

Note. $N = 297$. $R_L^2 = 0.0314$ p (χ^2) = 0.0004*
[a]Log Likelihood (LL) indicates the agreement between the probabilities of group membership generated by the logistic regression model and the actual group membership within the sample. Larger absolute LL values represent a worse model fit. $\chi^2 = -2(LL_{\text{null model}} - LL_{\text{full model}})$. [b]The null model represents the baseline model without information about the predictor variables. [c]The full model represents the hypothesized model with FSI entered as continuous variable.

A summary of the logistic regression estimates for the full and null models is provided in Table 4.7. As reported in the table, the null model was not significant. The null model's logit for predicting the "Fog" group was 0.333, which means that in the absence of information provided by the independent variables, the odds of the sample airports having a fog event on a given day over the period of the sample was 1.39. When the mathematical expression $c^{\text{onstant}} / (1 + c^{\text{onstant}})$ was applied to the null model (Warner, 2008,

p. 954), these odds indicate that approximately 58% of days in the sample were reported as fog days. Because the full model yielded a significant increase in the chi-square statistic, I deduced that the full model was correctly specified. Treating this result as an omnibus test, I examined the relationship between X_p and X_t and the likelihood of fog occurring.

Table 4.7
Summary of Logistic Regression Estimates for the Null and Simultaneous (Full) Models

	B_i[a]	*SE*	χ^2	*p*
Null Model[b]				
Constant	0.333	0.118	8.01	.0047**
Full Model[c]				
Constant	1.693	0.414	16.74	.0001**
X = FSI	0.033	0.009	11.93	.0006**

Note. $N = 297$. $R_L^2 = .0.0314$, $df = 1$ for the full model. Number of correctly classified cases = 195 at a predicted probability cut of .5.
[a]Logistic regression estimates are the natural logarithm of the odds ratio. Therefore, the exponential of the estimate (e^{Bi}) yields the odds ratio. [b]The null model represents the baseline model and predicts the odds for being involved in a predicted fog event without information provided by the independent variables. In the sample, these odds differed significantly from 1; that is, the probability of being involved in a predicted fog event differed significantly from 0.5. For example, in the null model, $B_0 = 0.333$ and therefore $e^{0.333} = 1.39$, which means that the odds of being involved in an predicted fog event was 1.39 at ECFA. This indicates that approximately 3/5 of the days in the sample were involved in a predicted fog event. [c]The full model represents the hypothesized model and predicts the odds for being involved in a predicted event with the independent variables entered into the model simultaneously.
*$p < .05$. **$p < .01$.

As summarized in Table 4.7, the full model logit (L_i) for the occurrence of fog was predicted by the equation $L_i = 1.693 - 0.033X$ and the IV was significantly related to the occurrence of fog. The exponent of each regression coefficient (e^{Bi}) in the prediction equation specifies the change in odds relative to the independent variable (X_i) while controlling for the other predictors in the model. In the context of the current study: (a) if $e^{Bi} < 1.00$, then the odds decrease for prediction of the "Fog" group relative to X_i; (b) if $e^{Bi} > 1.00$, then the odds increase for prediction of the "Fog" group relative to X_i; and (c) if $e^{Bi} = 1.00$ then there is no change in the odds for the prediction of the "Fog" group relative to X_i. (The reader is reminded that the "Fog" group refers to days where sample airports reported one of the accepted weather codes that indicate radiation fog such as: BR, FG, MIFG, and BCFG.

Table 4.8 shows the resulting odd ratios the logistic regression results using the dichotomous FSIs, X_p, and X_t.

Table 4.8
Summary of Odds Ratios for the Independent Variables

Independent Variables[a]	**Odds Ratios (OR)**	**95% CI**	***p***
***Xp* = Probabilistic FSI**			
40 or less vs. 40 or greater	2.02	[1.28, 3.24]	.0030**
40 or greater vs. 40 or less	0.50	[0.31, 0.79]	.0030**
X_t = Theoretical FSI			
43 or less vs. 43 or greater	2.08	[1.29, 3.34]	.0026**
43 or greater vs. 43 or less	0.48	[0.30, 0.77]	.0026**

Note. $N = 297$ [a]Significance tests and confidence intervals (CI) on odds ratios for the independent variables are likelihood ratio (χ^2) based. $*p < .05$. $**p < .01$.

X_t = *Theoretical FSI.* Table 4.8 indicates that the odds of belonging to the "Fog" group decreased significantly for days with an FSI higher than 43, *OR* = 2.08, means that the odds of belonging to the "Fog" group increased significantly for days with an FSI lower than 43. More concretely, this latter result indicates that days with an FSI lower than 43 were 2.08 times more likely to be a fog event than days with an FSI higher than 43.

It also should be noted that the corresponding 95% confidence interval (Table 4.9) indicates that 95% of the time, the odds ratio would vary anywhere between 1.29 and 3.34. In other words, 95% of the time, days with an FSI less than 43 would be between 1.29 and 3.34 times more likely to be involved in a fog event vs. days where the FSI was greater than 43. Given the width of this interval, the corresponding accuracy in parameter estimation is relatively low.

X_p = *Probabilistic FSI.* Table 4.8 indicates that the odds of belonging to the "Fog" group decreased significantly for days with an FSI higher than 43, *OR* = 2.02, means that the odds of belonging to the "Fog" group increased significantly for days with an FSI lower than 40. More concretely, this latter result indicates that days with an FSI lower than 40 were 2.02 times more likely to be a fog event than days with an FSI higher than 40.

It also should be noted that the corresponding 95% confidence interval (Table 4.9) indicates that 95% of the time, the odds ratio would vary anywhere between 1.28 and 3.24. In other words, 95% of the time, days with an FSI less than 43 would be between 1.28 and 3.24 times more likely to be involved in a fog event vs. days where the FSI was greater than 40. Given the width of this interval, the corresponding accuracy in parameter estimation is relatively low.

Predicted probabilities.

I used predicted probabilities to statistically classify cases according to group membership in the full model and to compare the reliability of Theoretical versus Probabilistic FSI. Cohen et al. (2003) suggested that such classifications are useful when a statistical model is used to make decisions among individuals. Classifications also can be used as supplementary analyses to determine the goodness of fit of a logistic regression model. As a result, I compared the statistical classifications of fog days in the full model to actual fog days by determining predicted probabilities for each case and developing a contingency table of predicted versus actual fog days.

Because there was no prior information about the proportion of days that sample airports were involved in fog events in the population, I followed Warner's (2008) recommendation and used a predicted probability cut of .5 for classifying cases in the "Fog" group.

X_t = *Theoretical FSI.* Overall, 181 cases were classified as belonging to the "Fog" group, and 116 cases to the "No Fog" group in the Theoretical model. As illustrated in Figure 4.5, 179 cases (60%) were correctly classified in the model at the predicted probability cut of .5. These correctly classified cases consisted of: 118 out of 173 "Fog" cases (68%) and 61 out of 124 "No Fog" cases (49%). It should be noted that the Theoretical model had a 32% miss rate (55 of 173) and a 51% false alarm rate (63 of 124).

X_p = *Probabilistic FSI.* Overall, 241 cases were classified as belonging to the "Fog" group, and 56 cases to the "No Fog" group in the Probabilistic model. As illustrated in Figure 4.5, 181 cases (71%) were correctly classified in the model at the predicted probability cut of .5. These correctly classified cases consisted of: 149 out of 173 "Fog" cases (86%), and 32 out of 124 "No Fog" cases (26%). It should be noted that the Probabilistic model had a 14% miss rate (24 of 173) and a 74% false alarm rate (92 of 124).

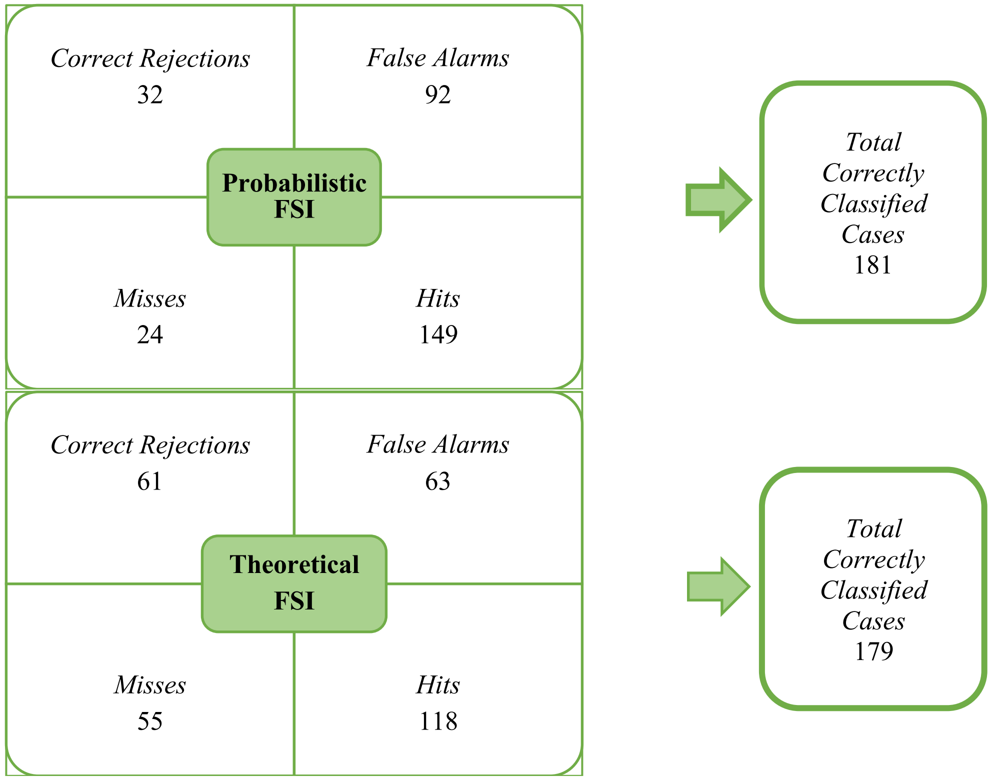
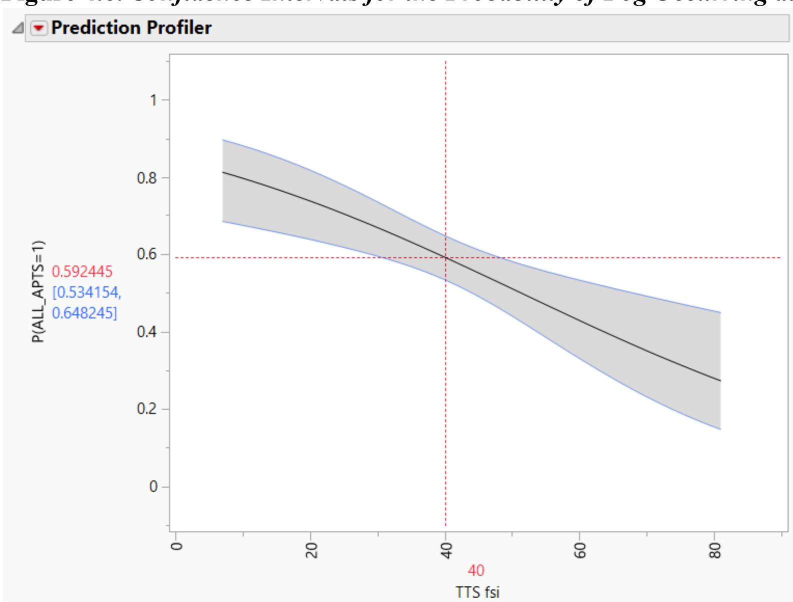


Figure 4.5. **Comparative predicted probabilities for the occurrence of fog.**
Note. Classification results for group membership in the full and stepwise models. Note that the cut was equal to p_i = .05. Correction rejections were the accurate classification of "No Fog" cases to days in the "No Fog" group. False alarms were the misclassification of "No Fog" cases to days in the "Fog" group. Misses were the misclassification of "Fog" cases to days in the "No Fog" group. Hits were the accurate classification of "Fog" cases to days in the "Fog" group.

Confidence Intervals for predicted probabilities.

I generated 95% confidence intervals for the continuous FSI model to evaluate the range of probabilities for the occurrence of fog at various FSI readings ranging 10 to 80. These results are provided in Appendix D. For ease of reference, Figure 4.6 shows the mean FSI value of 40 predicts the probability for fog occurring at ECFA ranges between 53.4% and 64.8%. The figure shows that range of confidence intervals is narrower for FSI values between 35 and 45. The reason for this shape is because the majority of the data in my random sample were between these values of 35 to 45. It demonstrates a greater degree of precision for these mean vales. The shape at the upper and lower ends of the curve show a larger width which indicates a larger confidence level. This suggests that the sample does not provide a precise or accurate representation of the population mean. This is where the outliers from the box and whisker plot in Figure 4.4 above would be located.

Figure 4.6. ***Confidence Intervals for the Probability of Fog Occurring at ECFA***



Chapter 5

Conclusions, Implications, and Recommendations

5.1 Summary of the Study

This study aimed to identify and determine reliable predictors of the occurrence of radiation fog at U.S. airports. Therefore, this study explored several aspects of radiation fog, including its occurrence, visibility factor, formation, persistence, and its relation to the physical environment. Fog Stability Index (FSI) as well as 1-Day Persistence, whether or not the fog persisted to the following day, were two potential predictors of fog examined in this study. They were used to determine if probabilistic measures determining the occurrence of fog would be significant predictors of radiation fog at six airports located in east-central Florida (ECFA). The ECFA examined were NASA Shuttle Landing Facility (TTS), Daytona Beach International (DAB), Patrick Space Force Base (COF), Orlando Stanford International (SFB), Orlando Melbourne International (MLB) and Orlando International Airport (MCO). More than one airport was examined because it was to be ensured that the occurrence of fog was captured and that there would be enough observable, varying occurrences of fog for consideration in the analysis. This is because fog in this region may occur at one airport but not at the other. This study was proposed to Florida Tech's College of Aeronautics by NASA's 45th Weather Squadron to determine the suitability of using a probabilistic FSI model to predict fog levels at ECFA with the use of Automated Weather Observing Systems (AWOS) data at each of the airports and radiosonde data at TTS. The 45th Weather Squadron provided a plethora of archival FSI data from 1973 to 2020 generated by the radiosonde at TTS.

Fog is defined as a visible aerosol made up of tiny water droplets, or ice crystals suspended in the air close to the earth's surface. This study's main focus was on radiation fog, which mainly occurred between the hours of 1000Z and 1500Z, any fog detected thereafter was categorized as advection, or sea fog. The FSI and the METAR observations were utilized to calculate the findings. There were multiple WX codes generated from METAR data, however, only four were pertinent directly to radiation fog. These included Shallow fog (MIFG), Patchy fog (BCFG), Light fog/mist (BR), and Fog (FG).

Fog has proved to be a prominent issue in transportation, especially in the aviation industry. These issues are causal and incidental to delays, diversions, and cancellations of flights that further cause both economic/financial concerns. The 45th Weather Squadron, along with the research performed in this study, indicated that, in order to avoid mixing occurrences of advection/sea fog into the study of radiation fog, it was best to limit the analysis to data between 1000Z and 1500Z. This is a time when advection/sea fog has low chances of occurring, and radiation fog is most prominent.

In order to analyze the data provided and generated, descriptive statistics were used in this study. Several tables and bar graphics were provided to compare the results generated from the complied fog and non-fog days at the six airports. These were recorded in binary and represented as 1 – Fog occurred and 0 – no fog occurred. This data was generated and subsequently was then compared to the FSI readings at TTS, the originally provided archival data from the 45th weather squadron, to see how well fog was predicted.

Theoretical and Probabilistic models were created for analysis and their individual results were compared and contrasted. The models were based on the midrange of the standard and sample FSI values. The FSI midrange value was 43 for the Theoretical model and 40 for the Probabilistic model. A contingency analysis was ran for both models to check if the respective models correctly identified the fog or non-fog days. Essentially, the analysis notated whether or not the FSI was correct when reported or if it produced misses or false alarms. To understand the accuracy and viability of the two models, two separate logistic regression analyses were performed and consequently analyzed. Testing odds ratio, examining model and dependent variable relationships to FSI, and whether the models were statistically significant were among factors scrutinized in the logistical regression. Both the Theoretical and Probabilistic models were proven to be viable for use in this study and the prediction of radiation fog events at ECFA using FSI derived from radiosonde data at TTS.

5.2 Research Findings

RQ1. *Are FSI, a Probabilistic FSI, and 1-Day Persistence reliable predictors of the occurrence of radiation fog at east-central Florida airports?* 1-Day Persistence did not meet the requirements for a logistic regression in the sample due to the low number of occurrences and was therefore eliminated as a potential predictor. Theoretical FSI and Probabilistic FSI are good predictors for determining the occurrence of radiation fog at ECFA. However, Probabilistic FSI proved to be better at predicting fog and non-fog events, with 71% correctly classified cases as compared with 60% in the Theoretical model. The Probabilistic model had a 14% miss rate and but a 74% false alarm rate. In practice, I would infer that a lower miss rate and higher alarm rate would lead to better outcome in preparing airports for fog events. The Theoretical model had a 32% miss rate and a 51% false alarm rate. The higher miss rate in the Theoretical model would be less desirable because more events for which airports would not be prepared for the resulting operational delay etc.

RQ2. *What predictor has causal priority, and in what order?* Causal priority refers to the order in which multiple predictors such as 1-Day Persistence and FSI would be entered in the probabilistic model. However, I was unable to answer this research question because 1-Day Persistence was eliminated as an IV in the preliminary analysis.

5.3 Conclusions and Inferences

As discussed before, the purpose of this study was to identify and determine what the suitable predictors of the occurrence of radiation fog at U.S. airports were. These predictors were FSI values and 1-Day persistence of fog. It was determined that 1-Day persistence was not a strong enough predictor of 1-Day persistence on its own per the scope of this study and the specific, randomly sampled METAR report data used. Out of the 297 sample data points examined after all incomplete or missing data points were removed, there were only nine days when fog persisted to the following day. However, there was no recorded instance where fog occurred after a non-fog day at any of the six sample airports. This may only be a disparaging factor based on the data points generated by the random sample and there could be far more significant counts of 1-Day persistence within the recorded days provided by the archival data. Therefore, while for the scope of this study it was proved 1-Day persistence was not a viable predictor of fog, it should still be researched further, and a larger or differing sample data set may be created to assist in addressing this independent variable more properly to close out any doubts.

Nonetheless, FSI was proved to be a strong predictor of fog. This is because it utilized probabilities to give a percentage chance of occurrence, rather than the categories of high, medium, and low. In testing the Theoretical and Probabilistic models, it was discovered that both models were fairly accurate in their predictions of fog days and non-fog days. However, findings suggest that the Probabilistic model, using an FSI mid value of 40, was more valuable. The Probabilistic model predicted more accurately whether a fog day would occur, and it sided more towards false alarms than to misses which is certainly preferable. Over-preparedness, in this case, is seemingly easier and more desirable to accomplish as compared to under-preparedness resulting from a miss. Nevertheless, the percent chance of occurrence provides airports, airlines, air traffic control, etc. an opportunity for more reliable and informed decision making in terms of operational abilities and capabilities.

5.4 Implications for Practice

This study identified the factors related to assessing the effectiveness of the use of FSI, a probabilistic FSI, and 1-Day persistence as suitable predictors of fog. There are several metrics to quantify the occurrence and intensity/duration of fog in the area. The simplest metric is a binary occurrence of fog at any of the six airports in the study area. This metric has the shortfall of low dynamic range. If light fog was reported at just one of the airports for a few minutes, that counted as fog occurred. If heavy fog was reported at all the airports for many hours, that counted as the same as fog occurred, despite being a much different event. This was the metric chosen for its simplicity in this proof-of-concept study. Other metrics that could be used in future studies include an integer occurrence scale ranging from 0 to 6 depending on the number of airports at which fog occurred. This could potentially show the range, impact, and correlation between fog occurrence at centrally located airports. Another approach would be to develop a continuous fog intensity/duration metric. This metric could add up the amount of time fog was reported and then multiply it by the inverse of the visibility for all the airports. Thus, if no fog were observed, the metric would be zero. If only light fog were observed briefly at only one airport, the metric would give a low score. If heavy fog (low visibility) was observed at all the airports all night long, the metric would yield a large score. The inverse of visibility is needed since lower visibility means a more intense fog event. If fog is defined as a visibility of 1 km or less, then this inverse measure could be (1.1 km – visibility) and only scored if visibility is 1 km or less, i.e., negative measures for visibility > 1 km are set to zero.

5.5 Recommendations for Practice and Research

This study identified the strong correlation between FSI and it being able to predict whether there was going to be fog, or no fog occurring. This is helpful to ECFA since it can help with more accurate forecasting of fog with the use of FSI. This will not only help to reduce the number of delayed, diverted, or cancelled flights, but will also allow for less frequent economic losses over time given the increased information provided to decision makers. Whenever there are disruptions to the normal operations at an airport, an airline and its affected passengers would have incurred operational, monetary, and/or time losses as discussed in Table 3.2, Chapter 3. For those ECFA that are considering implementing the use of FSI as a predictor of fog in the future, this study can prove to be useful.

It should be noted that FSI is probabilistic and may only indicate a likelihood of fog occurring or not. Therefore, each indication must be taken with a grain of salt and assumed that there is a corresponding probability of the FSI inaccurately predicting fog or no fog based on the percent chance of fog occurring. This is why utilization of the

Probabilistic model presented with a mid FSI of 40 value is recommended. It more accurately predicted when fog events were to occur and leaned toward false alarms over misses, unlike its Theoretical model counterpart with a mid FSI value of 43. As discussed, these false alarms are considered preferable, though complete accuracy is desired, due to a prepared nature resulting as comparative to being under-prepared, assuming no fog will occur.

A portion of the descriptive statistics analyzed in this research presented the commonality and average severity of weather code types. It is recommended that this be further studied in the future and assessed in conjunction with FSI values as it would provide decision-makers an added level or information. Mist (BR) may be indicated in a METAR report or indicated by an FSI value to probabilistically occur, but if its average or range is not threatening in terms of visibility, as shown by the descriptive statistics analysis, then operations may be able to continue without added response required.

A recommendation for further research would be to have either or differing or larger sample of data points compared to this study to thoroughly examine 1-Day persistence over a broader scope in order to determine if it can be a reliable predictor of fog. As discussed, this study could not fit 1-Day persistence into a logistical regression or contingency analysis because it was missing one of the four necessary values: a fog day after a non-fog day indicating or simulating a sense of 1-Day persistence. If another study can perform such research and find data points with that indication included, it would at least pave a path toward analyzing the viability and ensuing potential reliability of 1-Day persistence in the probabilistic prediction of fog.

Another recommendation suggested or, rather, a warning presented is to ensure that one has enough time to collect and scrub METAR data for pertinent information. This is one of the most time-consuming portions of this study since it had to be done for every single random data sample point. There is no easy, quick method to collect many data points and analyze them quickly. Given the relatively short time allotted, this study was unable to separate day, month, year, and other information without having to manually write each portion into separate data points for analysis. This would have been an immense undertaking requiring far more time than available. If further research is conducted, time allotted should be considered, and the points should be taken into account. This further points to a clear need for automation of the data preparation process to save time and effort.

Continuing, it would be immensely helpful to have Iowa State University update their Mesonet to produce results with separated date and time data points. Perhaps this is a limitation of their METAR data collection process, program, or coding, but if possible this would eliminate the need for such intense data preparation for all users of the archival data. Nonetheless, it would be of interest to see FSI, weather code types, and 1-Day persistence in relation to date events. Notably, most, if not all, of the data provided occurred between September of one year and March of the following year. Seeing how fog events and FSI act based on the date period, potentially seeing increased frequency, severity, or accuracy, would allow further information provisions to decision-makers. One could write code to extract additional dependent variables from airport weather stations such as a temperature curve during a period or other meteorological data. This would provide for additional collection of timely weather data given that the FSI models are based on weather sensitive events. Finally, by bringing attention and awareness of these capabilities amongst members of the aviation industry as well as awareness of the questions in aviation amongst machine learning industry members, great collaborations could result, and the issues and questions could be solved. In total conclusion, while there is much to research in the future, it can be stated that FSI has proven to be a valuable asset for airports, airlines, and the like in the probabilistic determination of fog.

References

Advection Fog. Advection fog - Glossary of Meteorology. (2012, January 26). Retrieved March 15, 2022, from https://glossary.ametsoc.org/wiki/Advection_fog

Akrherz@iastate.edu, D. H. (n.d.). *IEM :: Download ASOS/AWOS/Metar Data*. Iowa Environmental Mesonet. Retrieved March 18, 2022, from https://mesonet.agron.iastate.edu/request/download.phtml

Akrherz@iastate.edu, D. H. (n.d.). *IEM :: Download ASOS/AWOS/Metar Data*. Iowa Environmental Mesonet. Retrieved March 18, 2022, from https://mesonet.agron.iastate.edu/request/download.phtml

Airlines for America (2020, May 8). *U.S. Passenger Carrier Delay Costs.* Retrieved January 24, 2022, from https://www.airlines.org/dataset/u-s-passenger-carrier-delay-costs/

Asana. (2021, August 6). *Understanding the pareto principle (the 80/20 rule) • asana*. Asana. Retrieved January 25, 2022, from https://asana.com/resources/pareto-principle-80-20-rule

Automated Weather Observing System (AWOS). SKYbrary Aviation Safety. (2021, June 23). Retrieved December 17, 2021, from https://skybrary.aero/articles/automated-weather-observing-system-awos

Bergot, T., & Koracin, D. (2021, February 9). *Observation, simulation and predictability of fog: Review and Perspectives*. MDPI. Retrieved December 17, 2021, from https://doi.org/10.3390/atmos12020235

Duynkerke, P. G. (1991, February 1). *Radiation fog: A comparison of model simulation with detailed observations*. AMETSOC. Retrieved December 17, 2021, from https://journals.ametsoc.org/view/journals/mwre/119/2/1520-0493_1991_119_0324_rfacom_2_0_co_2.xml

Encyclopædia Britannica, inc. (n.d.). *Temperature inversion*. Encyclopædia Britannica. Retrieved December 17, 2021, from https://www.britannica.com/science/temperature-inversion

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical Power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/brm.41.4.1149

Fog stability index: A novel technique for fog/low clouds detection using multi-satellites data over the indo-gangetic plains during winter season. Taylor & Francis. (n.d.). Retrieved December 17, 2021, from https://www.tandfonline.com/doi/full/10.1080/01431161.2018.1483085

Fog. Fog - Glossary of Meteorology. (2012, April 25). Retrieved March 13, 2022, from https://glossary.ametsoc.org/wiki/Fog

Freezing Fog. Freezing fog - Glossary of Meteorology. (2012, January 26). Retrieved March 15, 2022, from https://glossary.ametsoc.org/wiki/Freezing_fog

G, E. (2022). *U.S. passenger carrier delay costs*. Airlines For America. Retrieved March 17, 2022, from https://www.airlines.org/dataset/u-s-passenger-carrier-delay-costs/

Google. (n.d.). Google maps. Retrieved December 17, 2021, from https://www.google.com/maps/place/Lafayette+Airport/@30.2118292,-91.9894856,4603m/data=!3m1!1e3!4m5!3m4!1s0x86249cc9f702a8b7:0x4a9bd7eba9083acf!8m2!3d30.2080487!4d-91.9930795

Hsieh, F. Y. (1989). Sample size tables for logistic regression. *Statistics in Medicine*, *8*(7), 795–802. https://doi.org/10.1002/sim.4780080704

Journal list menu. Wiley Online Library. (n.d.). Retrieved December 17, 2021, from https://doi.org/10.1002/qj.3300

Ju, T., Wu, B., Zhang, H., & Liu, J. (2020, May 8). *Parameterization of radiation fog-top height and methods evaluation in Tianjin*. MDPI. Retrieved December 17, 2021, from https://doi.org/10.3390/atmos11050480

Klemm, O., & Lin, N. H. (2016, April 22). *What causes observed fog trends: Air quality or climate change?* Aerosol and Air Quality Research. Retrieved December 17, 2021, from https://aaqr.org/articles/aaqr-15-05-oa-0353

Kulkarni, R., Jenamani, R. K., Pithani, P., Konwar, M., Nigam, N., & Ghude, S. D. (2019, April 12). *Loss to aviation economy due to Winter fog in New Delhi during the winter* *of 2011–2016*. MDPI. Retrieved December 17, 2021, from https://www.mdpi.com/2073-4433/10/4/198

Maalick, Z., Kühn, T., Korhonen, H., Kokkola, H., Laaksonen, A., & Romakkaniemi, S. (2016, March 9). *Effect of aerosol concentration and absorbing aerosol on the radiation fog life cycle*. Atmospheric Environment. Retrieved December 17, 2021, from https://www.sciencedirect.com/science/article/pii/S1352231016301881

Moore, D. A., & Peterson, R. E. (1984). Air Force Air Weather Service's requirements for Global Meteorological Data. *Applied Optics*, *23*(15), 2474. https://doi.org/10.1364/ao.23.002474

Price, J., & Stokkereit, K. (2020, February 29). *The use of thermal infra-red imagery to elucidate the dynamics and processes occurring in Fog*. MDPI. Retrieved December 17, 2021, from https://doi.org/10.3390/atmos11030240

Radiation fog. Radiation fog - Glossary of Meteorology. (2012, January 26). Retrieved March 15, 2022, from https://glossary.ametsoc.org/wiki/Radiation_fog

Song, Y. Y. (n.d.). *Development and verification of the fog stability index for Incheon International Airport based on the measured fog characteristics*. Atmosphere. Retrieved December 17, 2021, from https://doi.org/10.14191/ATMOS.2013.23.4.443

Stevens, S. (2017, May 31). *7 types of fog you didn't know had names*. Treehugger. Retrieved December 17, 2021, from https://www.treehugger.com/types-fog-you-didnt-know-had-names-4869285

STRATEGIC AIR COMMAND OFFUTT AFB NE. (1989). Environmental impact analysis process. environmental assessment for the force structure actions at March Air Force Base. https://doi.org/10.21236/ada268553

A study of morning radiation fog formation. (n.d.). Retrieved December 17, 2021, from https://scholar.afit.edu/cgi/viewcontent.cgi?article=5875&context=etd

Tabachnick, B. G., & Fidell, T. &. (n.d.). *Using multivariate statistics, 7th edition*. Pearson. Retrieved March 18, 2022, from https://www.pearson.com/us/higher-education/program/Tabachnick-Using-Multivariate-Statistics-7th-Edition/PGM2458367.html

Temperature inversion. Temperature inversion - Glossary of Meteorology. (n.d.). Retrieved March 13, 2022, from https://glossary.ametsoc.org/wiki/Temperature_inversion#:~:text=A%20layer%20in%20which%20temperature,exchange%20can%20occur%20within%20it.

US Department of Commerce, N. O. A. A. (2018, March 28). *Advection Fog*. National Weather Service. Retrieved December 17, 2021, from https://www.weather.gov/safety/fog-advection

Warner, R. M. (2008). Applied statistics: From bivariate through Multivariate Techniques. SAGE Publications.

Appendix A – Raw Data

YEAR	MO	DAY	TTS FSI	MeanVis	MinVis_10_14Z	TTS	DAB	MCO	COF	MLB	SFB	ODP
1973	10	25	57	16000	16000	0	0	0	0	2	0	0
1973	12	11	62	16000	16000	0	0	0	0	2	0	0
1973	12	12	70	16000	16000	0	0	0	0	2	0	0
1975	2	6	37	15360	14400	0	1	1	0	0	1	1
1975	10	16	49	16000	16000	0	0	1	0	0	0	0
1975	11	20	53	16000	16000	0	0	0	0	0	0	0
1975	12	12	38	16000	16000	0	0	0	0	0	0	0
2006	11	21	46	16000	16000	0	0	0	0	0	0	0
2006	11	23	47	16000	16000	0	0	1	0	0	0	0
2006	12	5	35	16000	16000	0	0	0	0	0	0	0
2006	12	12	46	16000	16000	0	0	0	0	0	0	0
2006	12	30	36	16000	16000	0	1	1	1	0	1	0
2007	1	27	33	15680	14400	0	1	0	0	0	0	0
2007	2	9	50	15680	14400	0	0	0	0	1	0	0
2007	2	13	37	11200	6400	0	1	1	0	0	1	0
2007	2	27	44	8050	4000	1	1	1	1	1	1	0
2007	3	23	46	16000	16000	0	0	0	0	0	1	0
2007	11	9	30	16000	16000	0	0	0	0	0	0	0
2007	12	15	32	9050	1600	1	1	1	1	0	1	0
2007	12	27	28	12800	8000	1	1	0	1	1	0	0
2007	12	29	40	10900	4000	1	1	1	1	1	1	0
2008	1	1	37	13050	8000	0	0	1	0	1	1	0
2008	1	3	75	16000	16000	0	0	0	0	0	0	0
2008	2	2	35	16000	16000	0	1	0	0	0	0	0
2008	2	6	39	16000	16000	0	1	1	0	0	0	0
2008	2	8	60	10280	3200	1	0	0	1	1	0	0
2008	2	16	20	10440	2000	1	1	1	0	1	1	0
2008	2	29	37	16000	16000	0	0	0	0	0	0	0
2008	11	6	27	15800	14400	0	0	0	0	1	0	0
2008	11	8	32	13030	8000	1	1	0	0	1	1	0
2008	11	19	53	16000	16000	0	0	0	0	0	0	0
2008	12	6	32	2810	400	1	1	1	1	1	1	0
2008	12	11	35	16000	16000	0	0	0	1	0	0	1
2008	12	16	24	5240	1200	1	1	1	1	1	1	0
2008	12	30	34	16000	16000	0	0	0	0	0	0	0
2009	1	7	58	16000	16000	0	0	0	0	0	0	0
2009	1	17	26	16000	16000	0	0	0	0	0	0	0
2009	1	29	38	16000	16000	0	0	0	0	0	0	0

YEAR	MO	DAY	TTS FSI	MeanVis	MinVis_10_14Z	TTS	DAB	MCO	COF	MLB	SFB	ODP
2009	1	30	47	9180	2000	1	0	1	0	1	1	0
2009	2	14	35	16000	16000	0	1	1	1	1	1	0
2009	2	20	77	16000	16000	0	0	0	0	0	0	0
2009	2	28	45	16000	16000	0	0	0	0	0	0	0
2009	3	6	48	16000	16000	0	0	1	0	0	0	0
2009	3	11	31	13760	4800	1	1	1	0	1	1	0
2009	3	16	44	16000	16000	0	1	1	0	0	1	0
2009	4	1	39	16000	16000	0	1	0	1	1	1	1
2009	11	5	49	16000	16000	0	0	1	0	0	1	0
2009	11	6	38	16000	16000	0	0	0	0	0	0	0
2009	11	10	53	16000	16000	0	0	0	0	0	0	0
2009	11	19	24	9600	4800	1	1	1	1	1	1	0
2009	11	28	29	16000	16000	0	0	0	0	0	0	0
2009	12	1	37	16000	16000	0	0	1	0	0	0	0
2009	12	5	30	14880	4800	1	1	0	1	1	1	1
2009	12	19	60	16000	16000	0	0	0	0	0	1	0
2009	12	30	18	16000	16000	0	0	0	0	0	0	0
2009	12	31	43	16000	16000	0	0	0	0	0	0	0
2010	1	6	50	16000	16000	0	0	0	0	0	0	0
2010	1	15	42	16000	16000	0	1	1	0	0	1	0
2010	1	16	45	16000	16000	0	1	1	1	0	0	1
2010	1	20	22	14860	12800	0	1	1	0	1	0	0
2010	2	2	36	13560	8000	1	1	1	1	0	1	0
2010	2	4	20	13000	8000	1	0	0	0	1	0	0
2010	2	5	48	16000	16000	0	0	0	0	0	0	0
2010	2	11	72	16000	16000	0	0	0	0	0	0	0
2010	2	17	57	16000	16000	0	0	0	0	0	0	0
2010	2	19	57	16000	16000	0	0	0	0	0	0	0
2010	2	24	33	15200	12800	0	1	1	0	1	1	0
2010	2	26	65	16000	16000	0	0	0	0	0	0	0
2010	3	5	73	16000	16000	0	0	0	0	0	0	0
2010	3	7	49	16000	16000	0	0	0	0	0	0	0
2010	11	8	65	16000	16000	0	0	0	0	0	0	0
2010	11	12	59	16000	16000	0	0	0	0	0	0	0
2010	11	17	52	16000	16000	0	1	1	0	1	1	0
2010	11	18	36	16000	16000	0	0	0	0	1	0	0
2010	11	21	42	16000	16000	0	0	0	1	0	0	0
2010	12	3	33	16000	16000	0	0	0	0	0	0	0
2010	12	5	34	15600	14400	0	0	0	0	0	0	0
2010	12	6	48	16000	16000	0	0	0	0	0	0	0
2010	12	16	39	16000	16000	0	0	0	0	0	0	0

YEAR	MO	DAY	TTS FSI	MeanVis	MinVis_10_14Z	TTS	DAB	MCO	COF	MLB	SFB	ODP
2011	1	8	71	16000	16000	0	0	0	1	0	0	0
2011	1	13	45	16000	16000	0	0	0	0	0	0	0
2011	1	30	38	14130	11200	0	1	1	0	1	1	0
2011	1	31	34	16000	16000	0	0	0	0	1	0	0
2011	2	4	26	1900	400	1	1	1	1	1	1	0
2011	2	14	22	16000	16000	0	0	0	0	0	0	0
2011	2	17	26	12800	9600	1	1	1	1	1	1	0
2011	2	20	23	1940	400	1	1	1	1	1	1	0
2011	2	21	40	3240	2000	1	1	1	1	1	1	0
2011	2	26	55	12110	6400	1	1	1	0	1	0	0
2011	3	9	47	16000	16000	0	0	0	0	0	0	0
2011	3	10	43	14930	9600	1	1	1	0	0	1	0
2011	3	19	42	16000	16000	0	1	0	0	1	0	0
2011	11	13	40	16000	16000	0	0	0	0	0	0	0
2011	11	15	38	9330	4000	1	1	1	1	1	1	0
2011	11	16	36	10690	2000	1	1	0	0	1	0	0
2011	11	17	34	14720	12800	0	1	1	1	1	1	0
2011	11	18	36	16000	16000	0	0	0	0	0	0	0
2015	10	1	40	15200	6400	1	0	0	0	0	0	0
2015	10	4	37	16000	16000	0	0	0	0	0	0	0
2015	10	7	31	16000	16000	0	0	0	0	0	0	0
2015	10	26	45	16000	16000	0	0	0	0	0	0	0
2015	10	27	50	16000	16000	0	0	0	0	0	0	0
2015	11	1	44	9500	1200	0	0	0	0	0	0	0
2015	11	2	32	16000	16000	0	1	0	0	0	0	0
2015	11	7	36	15470	12800	0	1	1	0	0	1	0
2015	11	13	33	16000	16000	0	1	1	0	0	1	0
2015	11	14	42	16000	16000	0	0	0	0	0	0	0
2015	11	19	44	14220	6400	1	1	1	0	0	0	1
2015	11	29	55	14400	9600	1	0	0	0	0	0	0
2015	12	3	39	11500	1200	1	0	0	0	1	0	1
2015	12	8	30	16000	16000	0	0	0	0	0	0	0
2015	12	16	36	7330	400	1	1	1	0	0	1	0
2015	12	25	24	10130	2400	1	1	1	1	1	1	0
2015	12	29	47	16000	16000	0	1	0	0	0	1	0
2015	12	30	28	6060	800	1	1	1	1	1	1	0
2016	1	1	33	15200	12800	0	1	1	0	1	1	0
2016	1	3	26	15800	14400	0	1	0	0	0	1	0
2016	1	5	42	16000	16000	0	0	0	0	0	0	0
2016	1	12	20	16000	16000	0	0	0	0	0	0	0
2016	1	19	35	16000	16000	0	0	0	0	0	0	0

YEAR	MO	DAY	TTS FSI	MeanVis	MinVis_10_14Z	TTS	DAB	MCO	COF	MLB	SFB	ODP
2016	1	26	30	13690	9600	1	0	0	0	0	0	0
2016	1	29	33	16000	16000	0	0	0	0	0	0	0
2016	2	1	35	16000	16000	0	0	0	1	0	0	0
2016	3	11	46	16000	16000	0	0	0	0	0	1	0
2016	3	13	42	16000	16000	0	0	0	0	0	0	0
2016	3	14	49	16000	16000	0	1	1	0	0	1	0
2016	3	15	77	16000	16000	0	1	1	1	1	1	0
2016	3	22	67	16000	16000	0	0	0	0	0	0	0
2016	3	24	54	16000	16000	0	1	0	0	0	1	1
2016	10	12	49	16000	16000	0	0	0	0	0	0	0
2016	10	15	58	16000	16000	0	0	1	1	0	0	0
2016	10	17	58	16000	16000	0	1	0	0	0	0	0
2016	10	18	50	16000	16000	0	0	0	0	0	0	0
2016	10	25	41	16000	16000	0	0	0	0	0	0	0
2016	10	29	63	16000	16000	0	0	0	0	0	0	0
2016	10	31	34	14080	11200	0	1	0	0	0	0	0
2016	11	6	65	16000	16000	0	0	1	0	0	0	0
2016	11	11	32	16000	16000	0	0	0	0	0	0	0
2016	11	20	30	16000	16000	0	0	0	0	0	0	0
2016	11	22	24	16000	16000	0	0	0	0	0	0	0
2016	11	28	55	16000	16000	0	0	0	0	0	0	0
2016	12	1	44	16000	16000	0	0	1	0	1	0	0
2016	12	5	30	4360	400	1	1	1	0	1	1	0
2016	12	6	41	16000	16000	0	0	0	0	0	0	0
2016	12	8	24	5760	400	1	0	1	0	1	1	0
2016	12	9	22	16000	16000	0	0	0	0	0	0	0
2016	12	12	45	5100	800	1	1	1	1	1	1	0
2016	12	18	41	15680	14400	0	1	1	0	1	1	0
2016	12	19	32	11280	4000	1	1	1	1	1	1	1
2016	12	24	31	12380	1200	1	1	0	0	1	1	0
2016	12	26	35	16000	16000	0	0	0	0	0	1	0
2016	12	31	35	16000	16000	0	0	0	0	0	0	0
2017	1	1	44	9830	1200	1	1	1	0	1	1	0
2017	1	27	67	16000	16000	0	0	0	0	0	0	0
2017	1	30	55	16000	16000	0	0	0	0	0	0	0
2017	2	2	37	5840	400	1	1	0	0	1	0	0
2017	2	4	38	15680	14400	0	1	0	0	1	0	0
2017	2	8	42	16000	16000	0	1	1	0	0	0	0
2017	2	15	47	10660	2800	1	1	0	1	1	0	0
2017	2	23	28	16000	16000	0	1	1	1	1	1	0
2017	3	1	33	5450	400	1	1	1	0	1	1	0

YEAR	MO	DAY	TTS FSI	MeanVis	MinVis_10_14Z	TTS	DAB	MCO	COF	MLB	SFB	ODP
2017	3	3	35	16000	16000	0	0	0	0	0	0	0
2017	3	4	61	15680	14400	0	0	0	0	0	0	0
2017	3	11	70	16000	16000	0	0	0	0	0	0	0
2017	3	18	52	13030	3200	0	1	1	0	0	1	0
2017	3	21	62	9640	1600	1	0	1	0	0	0	0
2017	3	23	38	14020	6400	1	1	0	0	0	0	0
2017	3	27	45	16000	16000	0	1	0	0	0	0	0
2017	3	29	47	16000	16000	0	0	0	0	0	0	0
2017	10	5	67	13830	4000	1	1	1	1	1	1	0
2017	10	6	42	15820	14400	0	1	0	0	0	1	0
2017	10	7	32	16000	16000	0	1	1	0	1	1	0
2017	10	11	44	16000	16000	0	0	0	0	0	0	0
2017	10	13	54	14930	12800	0	0	0	1	0	0	0
2017	10	17	31	16000	16000	0	1	0	0	1	0	0
2017	10	20	51	16000	16000	0	0	0	0	0	0	0
2017	10	25	27	16000	16000	0	0	0	0	0	0	0
2017	10	27	51	16000	16000	0	0	0	0	0	0	0
2017	11	1	32	16000	16000	0	0	0	0	0	0	0
2017	11	7	27	15400	14400	0	1	1	0	1	1	0
2017	11	11	31	15540	12800	0	0	0	0	0	0	0
2017	11	16	39	16000	16000	0	0	0	0	0	0	0
2017	11	28	28	16000	16000	0	1	0	0	0	0	0
2017	12	9	55	13930	4800	1	1	0	1	1	1	0
2017	12	25	41	16000	16000	0	0	0	0	1	0	0
2017	12	28	32	16000	16000	0	0	0	0	0	0	0
2017	12	30	17	15560	11200	0	0	1	0	1	0	0
2017	12	31	19	16000	16000	0	0	0	0	0	0	0
2018	1	6	33	16000	16000	0	0	0	0	0	0	0
2018	1	20	8	16000	16000	0	1	0	0	0	1	0
2018	2	10	34	14720	9600	1	1	1	0	0	1	0
2018	2	12	33	14080	11200	0	1	1	1	0	0	0
2018	2	13	26	15360	14400	0	1	1	0	1	0	0
2018	2	21	38	16000	16000	0	1	1	0	0	1	0
2018	3	3	46	16000	16000	0	0	0	0	0	0	0
2018	3	9	76	16000	16000	0	0	0	0	0	0	0
2018	3	14	49	16000	16000	0	0	0	0	0	0	0
2018	3	17	52	7230	400	1	0	0	0	0	0	0
2018	3	18	43	16000	16000	0	0	0	0	0	0	0
2018	3	19	43	15820	14400	0	1	0	0	0	1	0
2018	3	20	40	16000	16000	0	1	0	0	1	1	0
2018	3	23	60	16000	16000	0	0	0	0	0	0	0

YEAR	MO	DAY	TTS FSI	MeanVis	MinVis_10_14Z	TTS	DAB	MCO	COF	MLB	SFB	ODP
2018	3	26	50	16000	16000	0	1	0	0	0	0	0
2018	10	7	58	16000	16000	0	0	0	0	0	0	0
2018	10	15	37	16000	16000	0	1	0	0	0	0	0
2018	10	19	39	16000	16000	0	1	0	0	0	0	0
2018	10	24	37	16000	16000	0	1	0	0	0	0	0
2018	11	8	27	6290	600	1	1	0	0	1	1	0
2018	11	15	37	16000	16000	0	0	0	0	1	1	0
2018	11	19	37	15680	14400	0	1	0	0	0	0	0
2018	11	20	28	9230	1400	1	1	0	0	1	0	0
2018	11	23	30	16000	16000	0	0	0	0	0	0	0
2018	12	3	45	8820	1200	1	1	1	1	1	0	0
2018	12	8	31	13200	4400	1	0	0	0	0	0	0
2018	12	11	47	16000	16000	0	0	0	0	0	0	0
2018	12	17	44	16000	16000	0	0	0	0	0	0	0
2018	12	19	7	8760	1000	1	0	0	0	0	1	0
2018	12	22	72	16000	16000	0	0	0	0	0	0	0
2018	12	23	21	14720	9600	1	0	0	0	1	0	0
2018	12	24	9	7930	1200	1	0	0	0	0	0	0
2019	1	1	46	3470	400	1	1	1	0	1	1	0
2019	1	5	81	16000	16000	0	0	0	0	0	0	0
2019	1	15	22	16000	16000	0	0	0	0	0	0	0
2019	1	17	25	16000	16000	0	0	0	0	0	0	0
2019	1	20	49	16000	16000	0	0	0	0	0	0	0
2019	1	21	56	16000	16000	0	0	0	0	0	0	0
2019	1	27	38	13030	8000	1	1	1	1	1	1	0
2019	1	28	52	16000	16000	0	0	0	0	1	0	0
2019	1	29	32	16000	16000	0	1	0	0	0	0	0
2019	2	5	19	3270	400	1	1	1	1	1	1	0
2019	2	11	39	13690	9600	1	1	1	1	1	1	0
2019	2	14	30	16000	16000	0	0	0	0	0	0	0
2019	2	26	34	16000	16000	0	0	0	0	0	0	0
2019	2	28	42	16000	16000	0	1	0	0	0	0	0
2019	3	6	51	16000	16000	0	0	0	0	0	0	0
2019	3	7	37	16000	16000	0	0	0	0	0	0	0
2019	3	9	36	8580	2000	1	1	0	0	0	0	0
2019	3	10	39	6200	600	1	1	0	1	1	1	0
2019	3	13	38	16000	16000	0	0	0	0	0	0	0
2019	3	19	34	16000	16000	0	0	0	0	0	0	0
2019	3	21	31	16000	16000	0	0	0	0	0	0	0
2019	3	23	45	16000	16000	0	1	0	0	0	0	0
2019	3	25	48	6030	400	1	0	0	0	0	0	0

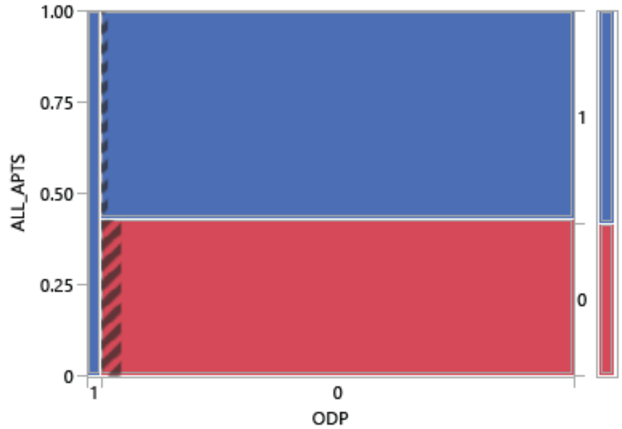
YEAR	MO	DAY	TTS FSI	MeanVis	MinVis_10_14Z	TTS	DAB	MCO	COF	MLB	SFB	ODP
2019	10	1	65	15040	11200	0	0	0	0	0	0	0
2019	10	3	47	16000	16000	0	1	0	0	0	0	0
2019	10	7	46	8930	4000	1	1	0	1	0	1	0
2019	10	10	38	14630	8000	0	1	0	0	0	0	0
2019	10	19	31	16000	16000	0	1	1	1	1	1	0
2019	10	20	57	16000	16000	0	0	0	0	0	0	0
2019	10	28	34	14400	8000	0	1	1	0	0	1	0
2019	11	1	37	16000	16000	0	0	0	0	1	0	0
2019	11	4	48	16000	16000	0	0	0	0	0	0	0
2019	11	8	25	11400	3200	0	0	1	0	1	0	0
2019	11	9	35	15620	9600	0	0	1	1	1	1	0
2019	11	17	35	16000	16000	0	1	0	0	1	1	0
2019	11	19	37	6530	600	1	1	0	0	0	1	0
2019	11	23	27	7120	600	1	1	1	0	1	1	0
2019	12	2	66	16000	16000	0	0	0	0	1	0	0
2019	12	8	28	6920	1400	1	1	0	0	0	0	0
2019	12	14	32	13000	600	1	0	1	1	1	0	0
2019	12	15	55	6500	200	1	0	0	0	0	0	0
2019	12	30	41	16000	16000	0	0	0	0	0	0	0
2020	1	6	28	16000	16000	0	0	0	0	0	0	0
2020	1	7	23	16000	16000	0	0	0	0	0	0	0
2020	1	16	34	16000	16000	0	1	0	0	0	0	0
2020	1	18	64	16000	16000	0	0	0	0	0	0	0
2020	1	19	40	16000	16000	0	0	0	0	0	0	0
2020	1	26	32	15360	12800	0	0	0	0	0	0	0
2020	1	27	12	15040	11200	0	0	0	0	0	0	0
2020	1	30	34	15470	11200	0	0	1	1	1	1	0
2020	2	1	53	11080	2800	1	1	1	0	1	1	0
2020	2	13	40	8200	1400	1	0	1	0	1	0	0
2020	2	16	36	16000	16000	0	1	0	0	0	1	0
2020	2	20	30	15600	12800	0	1	1	0	1	1	0
2020	3	6	75	16000	16000	0	0	0	0	0	0	0
2020	3	7	51	16000	16000	0	0	0	0	0	0	0
2020	3	8	49	16000	16000	0	0	0	0	0	0	0
2020	3	9	59	16000	16000	0	0	0	0	0	0	0
2020	3	10	54	16000	16000	0	1	0	0	0	0	0
2020	3	13	29	6530	1200	1	1	1	0	1	1	0
2020	3	17	39	16000	16000	0	1	0	0	0	0	0
2020	3	20	46	16000	16000	0	1	1	0	0	0	0
2020	3	23	33	16000	16000	0	0	0	0	0	0	0
2020	3	25	49	16000	16000	0	1	1	0	1	0	0

YEAR	MO	DAY	TTS FSI	MeanVis	MinVis_10_14Z	TTS	DAB	MCO	COF	MLB	SFB	ODP
2020	3	28	40	7330	400	1	0	1	0	1	0	0
2020	3	31	31	16000	16000	0	0	0	0	0	0	0
2020	10	1	39	16000	16000	0	0	0	0	0	0	0
2020	10	9	32	16000	16000	0	0	0	0	0	0	0
2020	10	14	41	16000	16000	0	1	1	0	0	0	0
2020	10	18	41	16000	16000	0	0	0	0	0	0	0
2020	10	21	53	15270	9600	0	0	0	0	1	0	0
2020	10	24	37	16000	16000	0	0	0	0	0	0	0
2020	10	25	43	16000	16000	0	1	0	0	1	0	0
2020	10	29	41	16000	16000	0	0	0	0	0	0	0
2020	11	8	62	16000	16000	0	1	0	0	0	1	0
2020	11	27	24	4510	600	1	0	1	0	0	0	0
2020	11	29	32	11640	800	1	1	1	0	0	0	0
2020	11	30	38	16000	16000	0	1	1	0	0	1	0
2020	12	1	58	16000	16000	0	0	0	0	0	0	0
2020	12	3	29	16000	16000	0	0	0	0	0	0	0

Appendix B – Contingency Analyses

Contingency Analysis of ALL_APTS By ODP

Mosaic Plot



Contingency Table

ALL_APTS

ODP	Count	0	1	Total
	1	0	9	9
	0	124	164	288
	Total	124	173	297

Tests

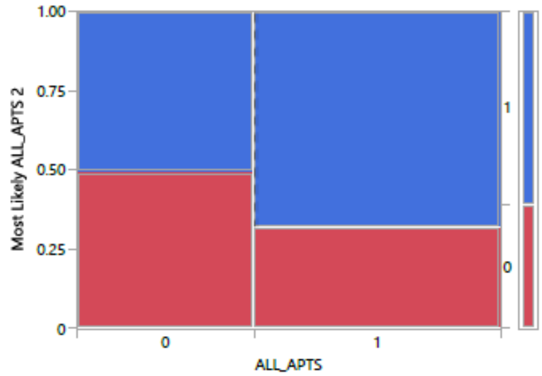
N	DF	-LogLike	RSquare (U)
297	1	4.9644753	0.0246

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	9.929	0.0016*
Pearson	6.652	0.0099*

Fisher's Exact Test	Prob	Alternative Hypothesis
Left	0.0071*	Prob(ALL_APTS=1) is greater for ODP=1 than 0
Right	1.0000	Prob(ALL_APTS=1) is greater for ODP=0 than 1
2-Tail	0.0117*	Prob(ALL_APTS=1) is different across ODP

Contingency Analysis of Most Likely ALL_APTS 2 By ALL_APTS

Mosaic Plot



Contingency Table

ALL_APTS / Count	Most Likely ALL_APTS 2: 0	1	Total
0	61	63	124
1	55	118	173
Total	116	181	297

Tests

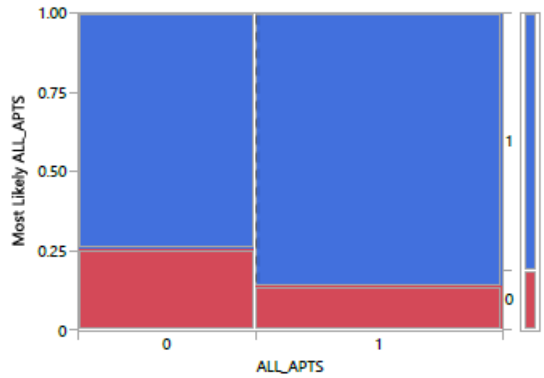
N	DF	-LogLike	RSquare (U)
297	1	4.5845648	0.0231

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	9.169	0.0025*
Pearson	9.189	0.0024*

Fisher's Exact Test	Prob	Alternative Hypothesis
Left	0.9992	Prob(Most Likely ALL_APTS 2=1) is greater for ALL_APTS=0 than 1
Right	0.0018*	Prob(Most Likely ALL_APTS 2=1) is greater for ALL_APTS=1 than 0
2-Tail	0.0027*	Prob(Most Likely ALL_APTS 2=1) is different across ALL_APTS

Contingency Analysis of Most Likely ALL_APTS By ALL_APTS

Mosaic Plot



Contingency Table

Most Likely ALL_APTS

ALL_APTS / Count	0	1	Total
0	32	92	124
1	24	149	173
Total	56	241	297

Tests

N	DF	-LogLike	RSquare (U)
297	1	3.3177191	0.0231

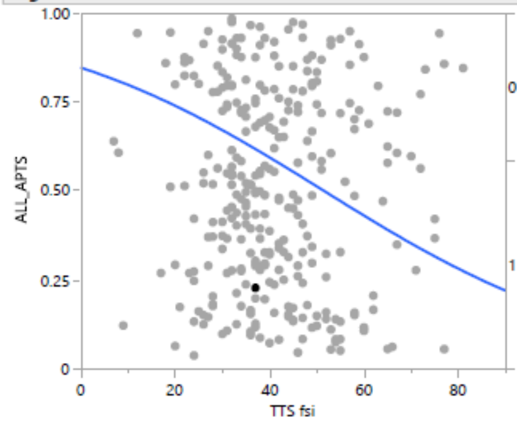
Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	6.635	0.0100*
Pearson	6.723	0.0095*

Fisher's Exact Test	Prob	Alternative Hypothesis
Left	0.9968	Prob(Most Likely ALL_APTS=1) is greater for ALL_APTS=0 than 1
Right	0.0076*	Prob(Most Likely ALL_APTS=1) is greater for ALL_APTS=1 than 0
2-Tail	0.0108*	Prob(Most Likely ALL_APTS=1) is different across ALL_APTS

Appendix C – Statistical Results

Nominal Logistic Fit for ALL_APTS

Logistic Plot



Converged in Gradient, 3 iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	6.33629	1	12.67259	0.0004*
Full	195.46779			
Reduced	201.80409			

RSquare (U)	0.0314
AICc	394.976
BIC	402.323
Observations (or Sum Wgts)	297

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	61	35.46504	70.93009
Saturated	62	160.00275	Prob>ChiSq
Fitted	1	195.46779	0.1803

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq	Lower 95%	Upper 95%
Intercept	1.6929423	0.4137233	16.74	<.0001*	0.88205951	2.5038251
TTS fsi	-0.0329715	0.0095454	11.93	0.0006*	-0.0516801	-0.0142629

Confidence limits are Wald-based.
For log odds of 1/0

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
TTS fsi	1	1	12.6725863	0.0004*

Odds Ratios

Unit Odds Ratios

Per unit change in regressor

Term	Odds Ratio	Lower 95%	Upper 95%	Reciprocal
TTS fsi	0.967566	0.949633	0.985838	1.0335211

Range Odds Ratios

Per change in regressor over entire range

Term	Odds Ratio	Lower 95%	Upper 95%	Reciprocal
TTS fsi	0.08717	0.021833	0.348034	11.471801

Tests and confidence intervals on odds ratios are Wald based.

Nominal Logistic Fit for ALL_APTS

Converged in Gradient, 3 iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	4.58456	1	9.16913	0.0025*
Full	197.21952			
Reduced	201.80409			

RSquare (U)	0.0227
AICc	398.48
BIC	405.827
Observations (or Sum Wgts)	297

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq	Lower 95%	Upper 95%
Intercept	0.26200461	0.1213701	4.66	0.0309*	0.02412349	0.49988573
FSI[<43]	0.36554529	0.1213701	9.07	0.0026*	0.12766417	0.60342641

Confidence limits are Wald-based.
For log odds of 1/0

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
FSI	1	1	9.16912953	0.0025*

Odds Ratios

For ALL_APTS odds of 1 versus 0

Odds Ratios for FSI

Level1	/Level2 ^	Odds Ratio	Prob>Chisq	Lower 95%	Upper 95%
>43	<43	0.4813837	0.0026*	0.2991372	0.7746621
<43	>43	2.0773449	0.0026*	1.2908854	3.3429472

Normal approximations used for ratio confidence limits effects: FSI
Tests and confidence intervals on odds ratios are Wald based.

Nominal Logistic Fit for ALL_APTS

Converged in Gradient, 3 iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	4.38930	1	8.778599	0.0030*
Full	197.41479			
Reduced	201.80409			

RSquare (U)	0.0218
AICc	398.87
BIC	406.217
Observations (or Sum Wgts)	297

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq	Lower 95%	Upper 95%
Intercept	0.32276281	0.1194368	7.30	0.0069*	0.08979353	0.55856235
FSI 2[<40]	0.35133618	0.1194368	8.65	0.0033*	0.11850131	0.58726346

Confidence limits are likelihood-based.

For log odds of 1/0

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
FSI 2	1	1	8.7785993	0.0030*

Odds Ratios

For ALL_APTS odds of 1 versus 0

Odds Ratios for FSI 2

Level1	/Level2 ^	Odds Ratio	Prob>Chisq	Lower 95%	Upper 95%
>40	<40	0.49526	0.0030*	0.3089651	0.7889892
<40	>40	2.0191414	0.0030*	1.2674444	3.2366114

Tests and confidence intervals on odds ratios are likelihood ratio based.

Generalized Linear Model Fit

Response: ALL_APTS
Modeling P(ALL_APTS=1)
Distribution: Binomial
Link: Logit
Estimation Method: Maximum Likelihood
Observations (or Sum Wgts) = 297

Whole Model Test

Model	-LogLikelihood	L-R ChiSquare	DF	Prob>ChiSq
Difference	6.33629315	12.6726	1	0.0004*
Full	195.467792			
Reduced	201.804085			

Goodness Of Fit Statistic	ChiSquare	DF	Prob>ChiSq
Pearson	297.7134	295	0.4448
Deviance	390.9356	295	0.0002*

AICc
394.9764

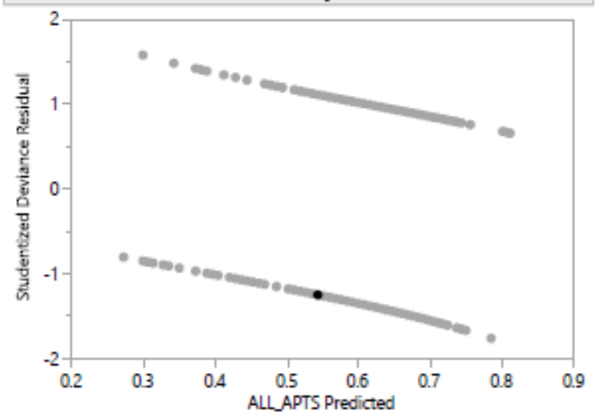
Effect Tests

Source	DF	L-R ChiSquare	Prob>ChiSq
TTS fsi	1	12.672586	0.0004*

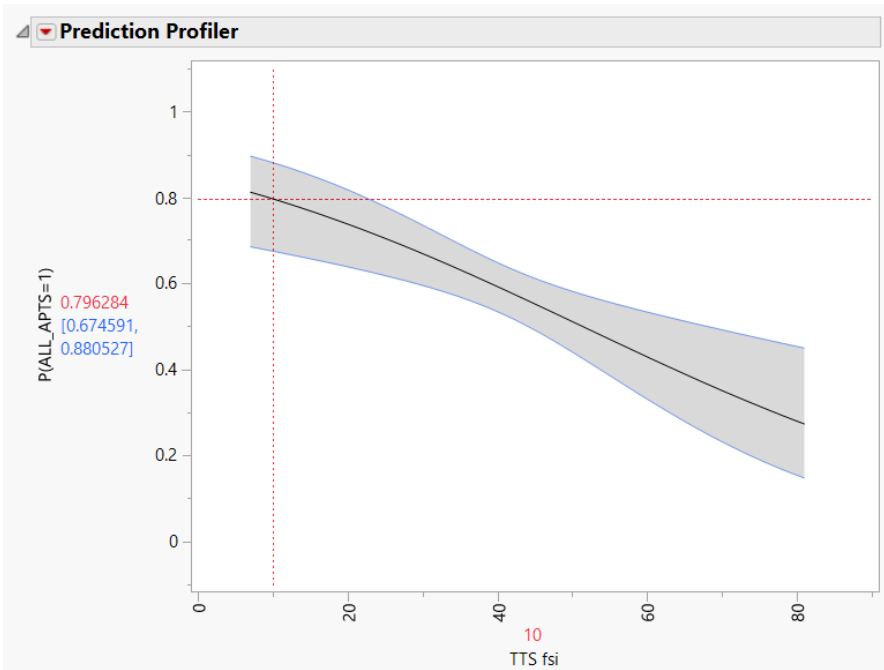
Parameter Estimates

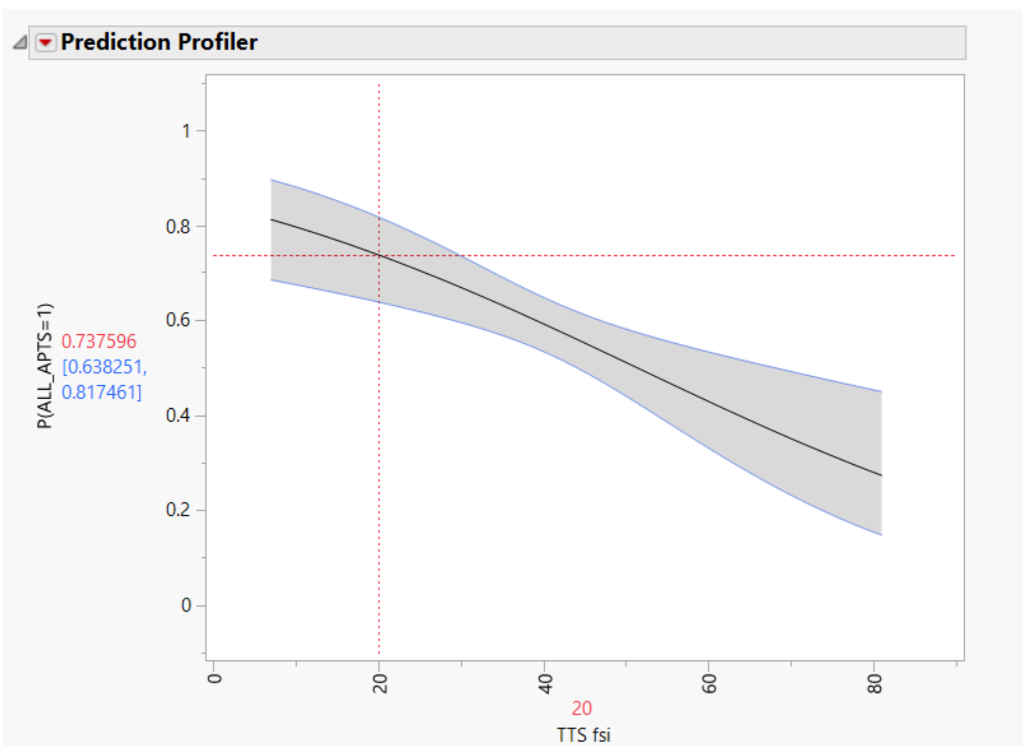
Term	Estimate	Std Error	L-R ChiSquare	Prob>ChiSq	Lower CL	Upper CL
Intercept	1.6929423	0.4137233	18.063182	<.0001*	0.8986891	2.5248958
TTS fsi	-0.032972	0.0095454	12.672586	0.0004*	-0.052168	-0.014623

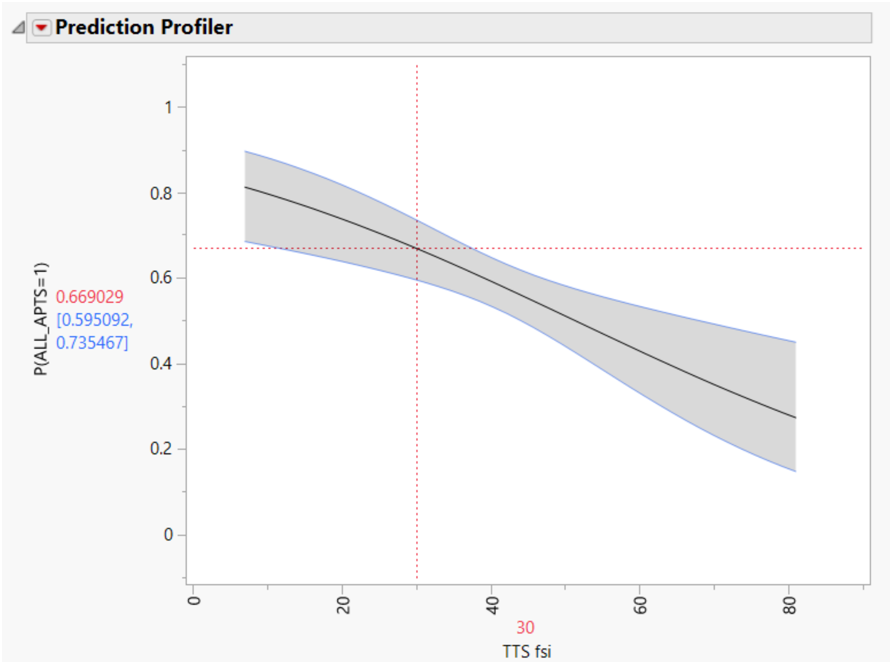
Studentized Deviance Residual by Predicted



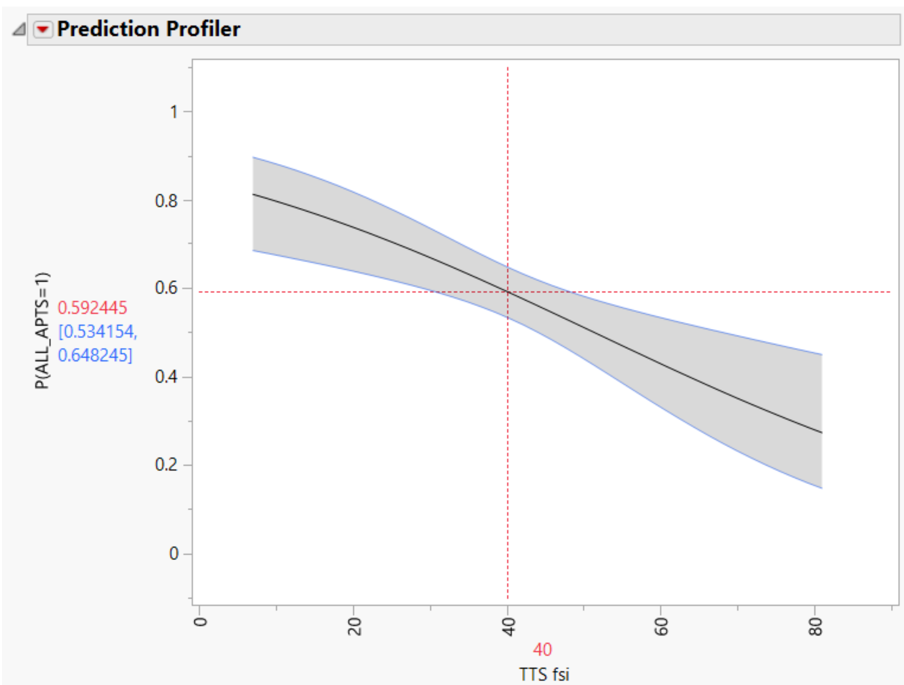
Appendix D – Predicted Probability Confidence Intervals



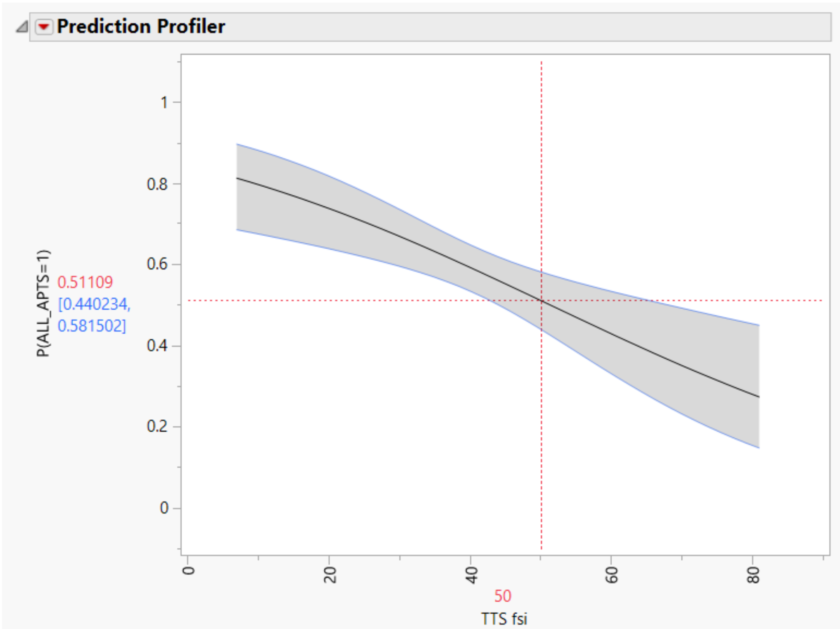


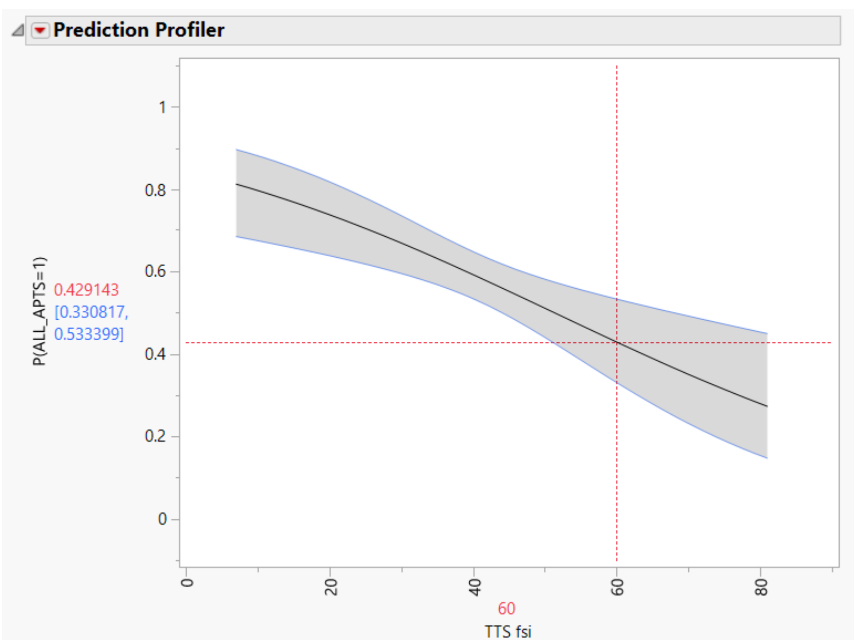


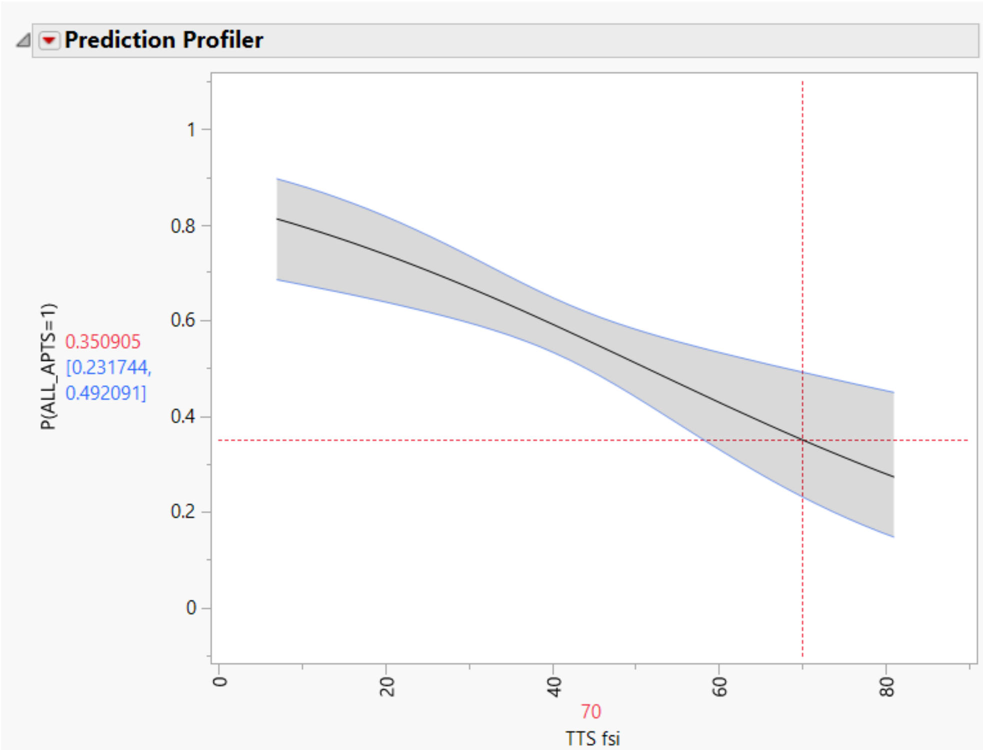
Prediction Profiler
P(ALL_APTS=1)
0.669029
[0.595092,
0.735467]
1
0.8
0.6
0.4
0.2
0
0
20
40
60
80
30
TTS fsi

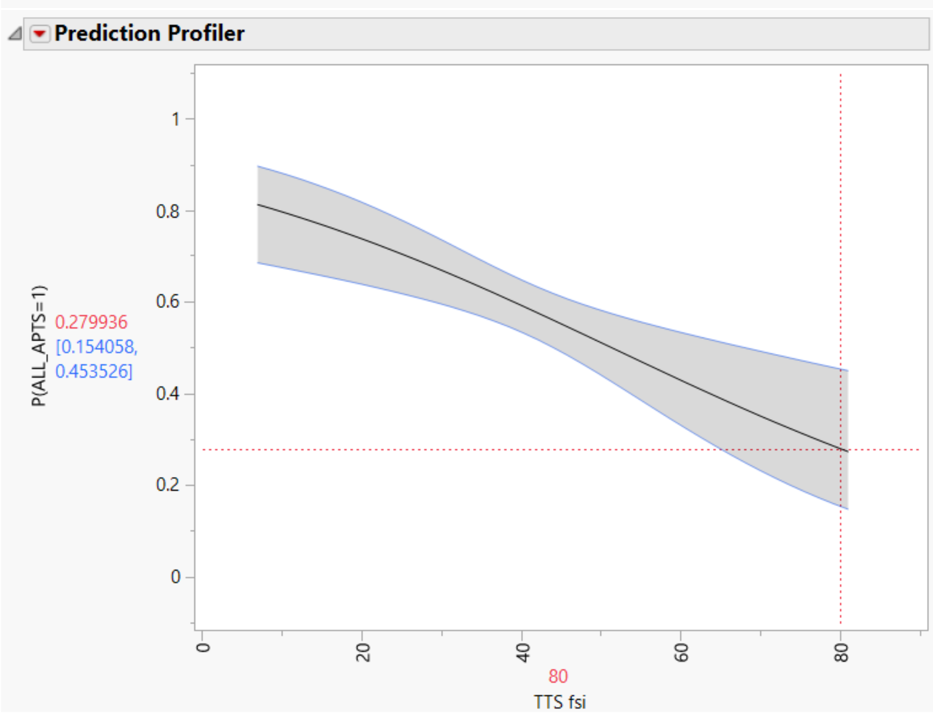


Prediction Profiler
P(ALL_APTS=1)
0.592445
[0.534154,
0.648245]
1
0.8
0.6
0.4
0.2
0
0
20
40
60
80
40
TTS fsi


Prediction Profiler
P(ALL_APTS=1)
0.51109
[0.440234,
0.581502]
1
0.8
0.6
0.4
0.2
0
0
20
40
60
80
50
TTS fsi


Prediction Profiler
P(ALL_APTS=1)
0.429143
[0.330817,
0.533399]
1
0.8
0.6
0.4
0.2
0
0
20
40
60
80
60
TTS fsi


Prediction Profiler
1
0.8
0.6
0.4
0.2
0
P(ALL_APTS=1)
0.350905
[0.231744,
0.492091]
0
20
40
60
80
70
TTS fsi


Prediction Profiler
1
0.8
0.6
0.4
0.2
0
P(ALL_APTS=1)
0.279936
[0.154058,
0.453526]
0
20
40
60
80
80
TTS fsi