AIRPORT PERFORMANCE METRICS ANALYSIS: APPLICATION TO TERMINAL AIRSPACE, DEICING, AND THROUGHPUT

Osama Alsalous

Dissertation Submitted to The Faculty of the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Civil Engineering

Susan Hotle, Chair
Montasir Abbas
Stacey Mumbower
Linbing Wang

April 27, 2022
Blacksburg, VA

Keywords: Air Traffic Management, Arrival Efficiency, Terminal Airspace, Deicing, Simulation, ASDE-X, COVID-19, Clustering Analysis, Airport Key Performance Indicators

Airport Performance Metrics Analysis:
Application to Terminal Airspace, Deicing, and Throughput

Osama Alsalous

ABSTRACT

The Federal Aviation Administration (FAA) is continuously assessing the operational performance of the National Airspace System (NAS), where they analyze trends in the aviation industry to help develop strategies for a more efficient air transportation system. To measure the performance of various elements of the aviation system, the FAA and the International Civil Aviation Organization (ICAO) developed nineteen key performance indicators (KPIs). This dissertation contains three research studies, each written in journal format, addressing select KPIs. These studies aim at answering questions that help understand and improve different aspects of airport operational efficiency. In the first study, we model the flight times within the terminal airspace and compare our results with the baseline methodology that the FAA uses for benchmarking. In the second study, we analyze the efficiency of deicing operations at Chicago O'Hare (ORD) by developing an algorithm that analyzes radar data. We also use a simulation model to calculate potential improvements in the deicing operations. Lastly, we present our results of a clustering analysis surrounding the response of airports to demand and capacity changes during the COVID-19 pandemic. The findings of these studies add to literature by providing a methodology that predicts travel times within the last 100 nautical miles with greater accuracy, by providing deicing times per aircraft type, and by providing insight into factors related to airport response to shock events. These findings will be useful for air traffic management decision makers in addition to other researchers in related future studies and airport simulations.

Airport Performance Metrics Analysis:
Application to Terminal Airspace, Deicing, and Throughput

Osama Alsalous

GENERAL AUDIENCE ABSTRACT

The Federal Aviation Administration (FAA) is the transportation agency that regulates all aspects of civil aviation in the United States. The FAA is continuously analyzing trends in the aviation industry to help develop a more efficient air transportation system. They measure the performance of various elements of the aviation system. For example, there are indicators focused on the departure phase of flights measuring departure punctuality and additional time in taxi-out. On the arrivals side, there are indicators that measure the additional time spent in the last 100 nautical miles of flight. Additionally, there are indicators that measure the performance of the airport as a whole such as the peak capacity and the peak throughput. This dissertation contains three research studies, each one aims at answering questions that help understand and improve a different aspect of airport operational efficiency. The first study is focused on arrivals where we model the flight times within the last 100 nautical miles of flight. Our model incorporated factors such as wind and weather conditions to predict flight times within the last 100 nautical miles with greater accuracy than the baseline methodology that the FAA currently uses. The resulting more accurate benchmarks are important in helping decision makers, such as airport managers, understand the factors causing arrival delays. In the second study, we analyze the efficiency of deicing operations which can be a major source of departure delays during winter weather. We use radar data at Chicago O'Hare airport to analyze real life operations. We developed a simulation model that allowed us to recreate actual scenarios and run what-if scenarios to estimate potential improvements in the process. Our results showed potential savings of 25% in time spent in the deicing system if the airport changed their queueing style towards a first come first served rather than leaving it for the airlines to have their separate areas. Lastly, we present an analysis of the response of airports to demand and capacity changes during the COVID-19 pandemic. In this last study, we group airports by the changes in their throughput and capacity during two time periods. The first part of the study compares airports operations during 2019 to the pandemic during the "shock event" in 2020. The second part compares the changes in airports operations during 2020 with the "recovery" time period using data from 2021. This analysis showed which airports reacted similarly during the shock and recovery. It also showed the relationship between airport response and factors such as what kind of airlines use the airport, airport hub size, being located in a multi-airport city, percentage of cargo operations. The results of this study can help in understanding airport resilience based on known airport characteristics, this is particularly useful for predicting airport response to future disruptive events.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor Dr. Susan Hotle for all the support, encouragement, and guidance that she provided throughout my PhD journey at Virginia Tech. Dr. Susan, I appreciate the balance between independence and guidance that you offered in your advising style. Deepest thanks to Dr. Montasir Abbas for being a part of my graduate school journey, thank you for the continuous support and for being a positive influence academically, professionally, and personally. Monty, I am inspired by your enthusiasm and your love to bring joy into all moments. Thanks to Dr. Satcey Mumbower for your support and feedback, sharing your expertise with me helped improve the quality of my research. I would also like to express my gratitude to Dr. Linbing Wang for his support and feedback, thank you for always making me feel welcome within the family of transportation systems engineering.

I would like to thank John Gulding and Marc Meekma for their support and the Federal Aviation Administration (FAA) for funding parts of this research.

I would also like to thank my parents, sister, and brothers for their unlimited support and love. You are a source of inspiration, I couldn't have done it without your much-needed words of encouragement and advice.

Many thanks to my music family, especially Itraab Arabic Music Ensemble at Virginia Tech. Special thanks to my friends from SalsaTech. My deepest gratitude to all my friends that have been there for me, thanks for your support, you all made a difference and made my journey a memorable one!

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Background and Motivation

The Federal Aviation Administration (FAA) is continuously assessing the operational performance of the National Airspace System (NAS), where they analyze trends in the aviation industry to help develop strategies for a more efficient air transportation system. The FAA and the International Civil Aviation Organization (ICAO) developed nineteen key performance indicators (KPIs) to measure the performance of various elements of the aviation system. For example, there are indicators focused on the departure phase of flights measuring departure punctuality and additional time in taxi-out, which are defined as KPI01 and KPI02, respectively. For arrivals, the key performance indicator KPI08 measures the efficiency of flight times within the terminal airspace, typically the last 100 nautical miles (nmi) of flights. Other indicators measure the airport peak capacity (KPI09) and peak throughput (KPI10) for both departures and arrivals. Monitoring the performance allows the FAA to identify and mitigate issues aimed at improving efficiency.

The FAA performance report shows inefficiencies in different phases of flight according to KPIs. Performance indicators compare reality to benchmarks, we realize that real life operations do not always go as expected due to the complexity of the aviation system. While it is important to look for trends in performance indicators, it is also beneficial to examine the methodology behind the measurement and factors driving the observed trends. Examining methodologies can help in producing more representative benchmarks and more robust metrics. Additionally, studying factors behind observed performance metrics is important for making decisions to improve future performance.

This dissertation contains three research studies, each written in journal format. These studies aim at answering questions that help understand and improve different aspects of airport operational efficiency. In the first study, we model the flight times within the terminal airspace and compare our results with the baseline methodology that the FAA uses for benchmarking. In the second study, we analyze the capacity of deicing facilities and efficiency of deicing operations at ORD by developing an algorithm that analyzes radar data. We also use a simulation model to calculate potential improvements in the deicing operations. Lastly, we present our results of a clustering analysis of the response of airports to demand and capacity changes during the COVID-19 pandemic.

1.2 Terminal Airspace Performance Benchmarking

The FAA reports the efficiency of the descent phase of flights by KPI08. This performance indicator compares actual transit time within 100 nmi around the airport to an unimpeded time that is derived from historical data. Majority of actual flight trajectories are found to be longer in time and distance due to air traffic control (ATC) management procedures. ATC manage the flow of air traffic using tactical procedures such as path stretching or airborne holding patterns that may be necessary to maintain safe separation between aircraft approaching

the airport. Inefficiencies in this phase contribute to arrival delays at the airport. Figure 1.1 shows an example of arrival trajectories into ORD compared to a flight that represents the unimpeded time that is used in KPI08 as a benchmark. The benchmark flight is highlighted in red in the figure.

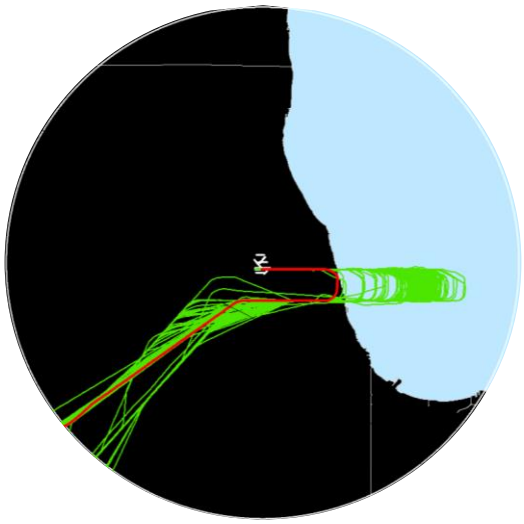

Figure 1.1 Flight trajectories in green shown alongside an unimpeded flight in red

The FAA currently calculates the unimpeded times for KPI08 by grouping flights by four characteristics and then finding the average of observations between the 5th and 15th percentiles. The characteristics used for grouping flights are arrival runway, engine type, approach bearing, and wake turbulence category (WTC). This method assumes that the population of flights between the 5th and 15th percentiles represent the ideal achievable performance. The benchmark shown in Figure 1.1 is for Medium Jets landing on runway 28C approaching from bearing 230 degrees.

1.3 Efficiency of the Departure Phase

The efficiency of the departure phase at an airport is typically measured by two metrics, namely departure punctuality and additional time in taxi-out. The FAA measures departure punctuality by KPI01 which measures the percentage of flights departing from the gate on-time. This indicator is considered airline-focused, it can be used as measure of the quality of service from a passenger perspective too.

Additional time in the taxi-out phase is a key metric that reflects the efficiency of surface operations at the airport. KPI02 measures the deviation from benchmark taxi-out times, the benchmark being the 20th percentile of observed taxi-out times at the airport. Using a percentile for benchmarking is intended to focus on the efficiency of air traffic management at the airport. Taxi-out time is impacted by factors such as queueing due to congestion at the airport or taxiway construction. Additionally, taxi-out time increases significantly during winter months at airports that experience snow and ice conditions. The increase during winter is a result of some aircraft needing to travel extra distance to centralized deicing facilities before they head to the departure

runway. The additional time in taxi-out, i.e. KPI02, historically showed seasonal spikes. It increases in January and February due to deicing operations and in July due to increased air travel demand.

1.4 Airport Capacity and Throughput Indicators

Indicators related to capacity and throughput measure performance from an airport wide perspective. The first metric is KPI09, it calculates the peak capacity at the airport. The FAA defines peak capacity as the 95^{th} percentile of the declared capacity at the airport. The air traffic manager at an airport sets the declared capacity value as the number of aircraft the airport can serve per unit time. Capacity is usually set at 15-minute intervals, this is also referred to as the called rate. Although runway capacity may not always be the limiting factor to how much traffic the airport can accommodate, it is an important indicator to show available capacity. The second indicator is KPI10, the 95th percentile of peak throughput. Similar to KPI09, the FAA calculates it as the 95^{th} percentile of the actual throughput at the airport. Each of KPI09 and KPI10 can be calculated for departures and arrivals.

These indicators are useful to evaluate periods of time that an airport experiences large scale changes in capacity and/or demand. Capacity and throughput indicators can capture the response of airports to events such as economic recessions, airline hub shifts, construction, or pandemics. Additionally, using KPI09 and KPI10 together can provide insight on how close an airport is operating to its capacity. This information is useful in the decision-making process to improve efficiency, for example, projects to increase capacity or strategies to decrease schedule peaking.

1.5 Major Contributions

In the first study, we model the relationship between travel time within the terminal airspace and contributing factors using a multivariate Log-Linear model. Our model quantifies the impact that these factors have on the total travel time within the last 100 nautical miles (nmi) of flight. We compare our results to the baseline set of variables that are currently used for benchmarking purposes at the FAA. The current benchmarking method assumes that the four characteristics are enough to predict the travel time and calculate efficiency. Our study provided a methodology that predicts travel times within the last 100 nmi with greater accuracy.

The second study evaluates potential efficiency improvements to the deicing operations using simulation. We focus on Chicago O'Hare International Airport (ORD) as a case study to show time-saving benefits of implementing different queueing strategies and resource distribution. We used real radar data from ORD to analyze aircraft movement and usage of the central deicing facility, then we developed the simulation model based on assumptions derived from data. This study provides a unique approach and findings that are not found in existing literature. The results of our analysis will help airport managers and stakeholders in evaluating the benefits of operational scenarios different than the baseline that is currently implemented. Our findings also provide deicing times per aircraft type, this information will be useful for other researchers in related studies and airport simulations.

The third study analyzes airports clustered by their response to the COVID-19 shock event influence on the aviation sector. Specifically, the clusters are generated on a 3-Dimensional analysis of percent change in peak throughput, peak capacity, and total operations. These results are then interpreted in conjunction with characteristics of the airports (e.g. split of cargo and commercial passenger flights, airport size). The pre-COVID-19 verses post-COVID-19 delay metrics are also evaluated. Our findings provide information on what airport characteristics support an airport's demand and operational resiliency to shock events.

CHAPTER 2

Modeling Arrival Flight Times within the Terminal Airspace

Alsalous, O., & Hotle, S. (2021). Modeling Arrival Flight Times within the Terminal Airspace. *Transportation Research Record*, 03611981211011487.

2020 Fred Burggraf Award: Outstanding Paper by a Young Author in Aviation

ABSTRACT

Air Traffic Management (ATM) efficiency in the descent phase of flight is a key area of interest in aviation research for the US, Europe, and recently other parts of the world. Efficiency of arrival travel times within the terminal airspace is one of nineteen key performance indicators defined by the Federal Aviation Administration (FAA) and the International Civil Aviation Organization (ICAO), typically within 100 nautical miles (nmi) of arrival airports. In this study, we model the relationship between travel time within the terminal airspace and contributing factors using a multivariate Log-Linear model to quantify the impact that these factors have on the total travel time within the last 100 nmi. We compare our results to the baseline set of variables that are currently used for benchmarking purposes at the FAA. The analyzed data included flight and weather data from January 1, 2018 to March 31, 2018 for 5 US airports: Hartsfield-Jackson Atlanta International (ATL), Chicago O'Hare (ORD), John F. Kennedy International Airport (JFK), LaGuardia Airport (LGA), and San Francisco International Airport (SFO). The modeling results show that there is a significant improvement in travel times prediction accuracy compared to the baseline methodology when additional factors are included such as wind, meteorological conditions, demand and capacity, Ground Delay Programs (GDPs), market distance, time of day and day of week. Root Mean Squared Error (RMSE) values from out-of-sample testing were used to measure the accuracy of the estimated models.

Keywords: Air Traffic Management, A100, Arrival Efficiency, Terminal Airspace

2.1 Introduction

Air Traffic Management (ATM) efficiency in the descent phase of flight is a key area of interest in aviation research in the US, Europe, and recently other parts of the world. It is one of nineteen defined Key Performance Indicators (KPIs) (1). The FAA developed the Terminal Airspace Performance Indicator to measure arrival horizontal inefficiency taking place in the terminal airspace, typically within 100 nmi around the arrival airport, sometimes referred to as the A100 measure *(2)*. This indicator is equivalent to KPI08 identified by the International Civil Aviation Organization (ICAO) *(1)* that is calculated by comparing actual travel times to derived unimpeded times within the last 100 nmi of the arrival airport. The additional time is usually a result of Air Traffic Control (ATC) procedures where ATC procedures may include a combination of speed assignment, altitude assignment into denser air or stronger headwinds, heading assignment ("vectoring"), clearance to fly towards a known fix, or airborne holding about a specific fix. In contrast, a theoretical Benchmark Flight requires the least amount of Air Traffic Control procedures with the best practical speed and flight distance to the arrival runway. This concept of a Benchmark Flight is illustrated as a red line in **Figure** 2.1.

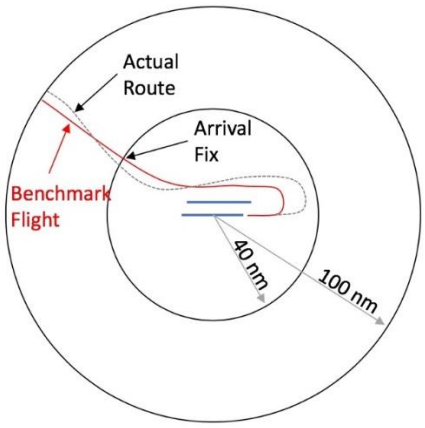

Figure 2.1 Illustration of additional time in terminal airspace

The A100 performance indicator is a standardized metric which can be used to analyze arrival performance at locations around the world. For example, the FAA/EUROCONTROL's joint reports include this metric to compare arrival performance at select U.S. and European airports. (2). In the current FAA methodology, the unimpeded time is often referred to as the benchmark time and is calculated from actual, historical flight data. To derive the benchmark time for a given airport, flights are first grouped by the following characteristics:

1. Arrival entry fix. The entry fix here refers to a point on the 40 nmi circle from the arrival airport that is an approach bearing cluster calculated from the data assuming that ATC merges arrival streams at the 40 nmi circle. Not to be confused for an arrival waypoint.
2. Arrival runway end.
3. Wake Turbulence Category (WTC). Aircraft types are categorized as Heavy, Medium, or Light.
4. Engine type. Aircraft are categorized as Jet, Turboprop, or Piston.

The travel time in the last 100 nmi of the airport is then calculated for each individual flight, being the time difference between entering the terminal airspace and wheels-on. For each of the above groups, the average of the flight travel times between the 5th to 15th percentiles is considered the benchmark time for that group. The minimum sample size necessary to compute an unimpeded time for each group is 50 flights, meaning the benchmark for any flight group is based on the average of at least 5 flights. Flight groups of less than 50 samples do not have a benchmark and, therefore, are not included in the A100 metric. Using the calculated unimpeded times per group, the unimpeded estimated time of arrival (ETA) for any flight can be calculated. Knowing the actual arrival time from data (Wheels On), the additional time in terminal airspace for any flight (i) is calculated by **Equation (2.1).**

$$Additional\ (A100)_i = Wheels\ On_i - ETA_i \quad (2.1)$$

It is evident that the groupings, although logical, are missing important factors for defining a benchmark. For example, within a single benchmark group, flights could experience good or bad weather and periods of high or low demand. This can lead the airport performance metrics to measuring more external factors (e.g. weather) instead of factors within the aviation system's control, such as how efficiently the pilot and air traffic controller handled the given situation. At an airport-level, the current benchmark would likely give high additional time metrics to airports frequently experiencing poor weather and, therefore, not directly measuring the airports' ability to operate in different scenarios. Also, this method removes the ability to interpret the magnitude of influence for each factor separately on the benchmark as the benchmarking groupings described above eliminates the effects of the categories by stratifying the data prior to calculating the metric.

Therefore, the purpose of this study is to quantify the impact that the benchmarking categories and additional other factors have on the total travel time within the last 100 nmi. These factors would also be important to include when estimating excess travel time. The total travel time within the last 100 nmi for any flight (i) is defined by **Equation (2.2)**.

$$Total\ A100\ Travel\ Time_i = Wheels\ On_i - \text{Time Entering the 100 nmi Circle}_i \quad (2.2)$$

The study's design of estimating the time within the 100 nmi instead of the benchmark-dependent excess time permits the benchmarking categories to be used as predictive variables, which will allow for interpretation of effects on the time within the 100 nmi circle. We model the relationship between travel time in the terminal airspace and contributing factors including airport-related factors, aircraft characteristics, weather, arrival time and day, and Traffic Management Initiatives (TMIs) among others. The analyzed data is for flights during the first

quarter of calendar year 2018 arriving into the following US airports: Hartsfield-Jackson Atlanta International (ATL), Chicago O'Hare (ORD), John F. Kennedy International Airport (JFK), LaGuardia Airport (LGA), and San Francisco International Airport (SFO). These airports were selected since they represent different conditions in terms of demand characteristics, weather patterns, fleet mix, wind variability, and runway configurations.

As far as flight time prediction, the FAA has the Traffic Flow Management System (TFMS) for supporting the management and monitoring of national air traffic flow. As part of TFMS, the Time Based Flow Management (TBFM) algorithms provide projections of flight times including estimated time of arrival (ETA). These projections are important for traffic managers and airlines, ETAs for example provide support for decision making such as gate crews scheduling and possible gate changes. However, our models achieve a different purpose which is understanding how omitted variables in the current A100 benchmarking process are affecting their robustness.

In this research, we evaluate the causal factors for terminal airspace inefficiency using a stepwise linear regression algorithm that uses historical data from the first quarter of calendar year 2018 to fit multiple models and select the best model based on the statistical analysis. All independent variables used are known when the flight reaches the 100 nmi point, allowing for the models to be used for predictive purposes. For example, distance flown inside the 100 nmi would impact the A100 time, but is not known at the 100 nmi point and, therefore was not used in the models.

The results of this work will help develop a better understanding of the impacts of different factors related to flight, airport, and weather conditions on terminal airspace operations and system predictability when it comes to planning. In addition to providing guidance on how to improve A100 benchmark times, this study can help predict future terminal airspace efficiency which potentially has great benefits to stakeholders including air traffic managers. For example, it can provide information to support decision making when it comes to balancing capacity utilization and inefficiencies during bad weather. Additionally, airlines could utilize this work for fuel efficiency purposes within the terminal airspace.

2.2 Literature Review

This study builds off existing literature measuring the efficiency of terminal airspace, where many of the studies expand on the FAA's benchmarking method that was previously described. For example, Alsalous and Galaviz-Schomisch expanded on the FAA's method measuring efficiency by calculating additional time in terminal airspace compared to benchmark times *(3)*. They analyzed the A100 metric for all flights during calendar year 2016 arriving into the US Core 30 airports with a detailed example at Los Angeles International Airport (LAX). In their study, they recommended using a time-based method instead of a distance-based method of calculation. They also found that Meteorological Conditions (MC) increased the time spent within the terminal airspace and recommended including MC in the benchmarking methodology based on their findings. In contrast to the Alsalous and Galaviz-Schomisch study, this paper

analyzes factors impacting the variation in the total amount of time spent in terminal airspace such as weather, capacity, and demand.

A study by Choi et al. included weather as an impacting factor on arrival delays by implementing machine learning algorithms to predict delays *(4)*. They used US domestic airline traffic data from the Airline On-time Performance dataset provided by the Bureau of Transportation Statistics (BTS), which contained traffic from the 2005 to 2015 for 45 major airports. The study used weather data from the National Oceanic and Atmospheric Administration (NOAA)'s Integrated Surface Database. Their model's binary output only classified whether a flight is delayed or not. The study considered a flight to be delayed if it arrived at the gate 15 minutes or more after the scheduled arrival time consistent with BTS's definition. Choi et al. considered various weather-related factors in depth. However, they did not include other non-weather factors such as aircraft characteristics and runway configuration. The power of their model's output remained limited due to not considering additional factors and not quantifying the magnitude of delay, which would be useful to capture the variability in travel times.

Other studies also estimated terminal airspace efficiency, evaluating improvement techniques from time and fuel saving perspectives. Knorr et al. analyzed efficiency in various flight phases including the decent phase using benchmarks similar to the derivation of the FAA benchmarking methodology *(5)*. It estimated the potential savings in time and fuel from ATM improvements for airports in the US and Europe. They used flight track radar data to develop a methodology that measures horizontal and vertical inefficiencies by analyzing flight trajectories. Their estimates indicate that more than 30% of the excess fuel burn on descent could be reduced through "better combined ATM and airline procedures" *(5)*. The highest potential for time and fuel savings in the US was for the New York area airports and for Europe was for London Heathrow (LHR). Cao et al.'s study on the Continuous Descent Approach (CDA) conducted a simulation-based assessment of the benefits of CDA. They used simulated flight trajectories for a full day at the four New York Center Metroplex airports using schedule data from March 1, 2005. Results showed that CDA can reduce noise, fuel consumption, and travel times within New York's terminal airspace *(6)*. Simić and Babić evaluated the impact of different ATC techniques on delays *(7)*. They included an analysis of air traffic complexity and how it relates to fuel, emissions, and time inefficiencies. However, they did not include how weather conditions can impact ATC efficiency within the terminal airspace.

Overall, existing literature evaluates important factors relevant to terminal airspace efficiency, but they were considered in isolation. Modeling the amount of time spent in terminal airspace while accounting for a wide range of factors is a rather unexplored research area which motivated this research. This gap in literature could be attributed to a lack of available data that covers a wide range of variables.

2.3 Data and Variables

In order to identify significant factors contributing to terminal airspace efficiency, it is important to have data related to each flight, including weather conditions both at the airport and at a higher altitude within the 100 nmi circle, aircraft characterisitcs, and airport operational variables. All factors considered in this study are listed in **Table** 2.1. This research was carried out using arrival flight data from January 1, 2018 through March 31, 2018 for 5 U.S. airports: Chicago O'Hare (ORD), San Francisco (SFO), Atlanta (ATL), John F. Kennedy airport (JFK), and LaGuardia (LGA). In addition, ORD was chosen to analyze the variability of A100 times at a single airport using runway-level models. These airports were selected due to their diversity of approach paths, runway configurations, historical weather patterns, fleet mix, and fluctuation of the demand-capacity ratio across a typical day.

The master data source in this analysis is FAA's Aviation System Performance Metrics (ASPM). It includes arrival efficiency data for flights traveling to and from the ASPM airports (currently 77) and operated by ASPM carriers *(8).* ASPM contains information on both the airport and individual flight. Flight level data includes variables that are useful to modeling flights based on characteristics such as origin airport, aircraft type, arrival fix, time and bearing at 100 nmi from arrival airport, arrival time, arrival runway. On the other hand, airport data is reported in 15-minute intervals including weather, active runway configuration, airport arrival and departure rates, and airport arrival and departure throughput.

Weather data were obtained from the Aviation Routine Weather Report (METAR) database that contains data reported from weather stations at each airport. This includes weather conditions such as ceiling, visibility, and Meteorological Conditions as one of two categories: Instrument Meteorological Conditions (IMC) or Visual Meteorological Conditions (VMC) as defined by the FAA *(2).* Clear skies were assigned a ceiling value of 30,000 ft above ground level, and a visibility of 10 miles. Additionally, METAR contains wind speed and direction at the airport surface.

Winds data at higher altitudes were obtained from the Rapid Refresh (RAP) numerical weather model provided by the National Oceanic and Atmospheric Administration (NOAA)*(9).* RAP generates weather data on a 13-km (8-mile) resolution horizontal grid every hour that contains variables such as wind, temperature, and humidity at 51 vertical levels. For this study, we averaged wind vector components within the 100 nmi circle for each airport at the 700 millibar (mb) isobaric (constant pressure) surface which is approximately at 10,000 foot altitude above sea level. This altitude was chosen to represent wind conditions at an intermediate point on the descent path between the last 100 and 40 nmi of the flight.

The Operations Network (OPSNET) is the official source of air traffic operations and delay data which is used to analyze the performance of the FAA's air traffic control facilities *(10)*. In this model we focused on one type of reported Traffic Management Initiatives (TMIs) which is Ground Delay Programs (GDPs). The FAA implements a GDP by holding flights at their departure airports due to conditions at their destination such as weather or congestion. The obtained data included the cause of delay, start time, and end time of each applied GDP which

we used to identify arrivals that took place during a GDP that was caused by conditions at the arrival airport being modeled.

The Base of Aircraft Data (BADA) *(11),* which is an aircraft performance model developed and maintained by EUROCONTROL, was used to map aircraft types to their weights to categorize them into 3 Wake Turbulence Categories (WTC) Heavy, Medium, or Light. The weight thresholds used for each group were the ICAO standard weights *(12)*. The WTC variable was created to capture the effect of the aircraft size on the time it spends in terminal airspace since aircraft of different sizes behave differently during sequencing and maneuvering.

All of the above data sources were merged to build a database of explanatory variables for each individual flight in the dataset. Some of the variables were directly used from the data source, while others were calculated by interacting variables together to create stronger predictors. **Table** 2.1 shows the explanation of variables derived from various data sources.

In addition to flight-level information, characteristics of the airport were also included that were hypothesized to affect ATC efficiency. Congestion levels at the airport were considered as air traffic demand is known to be cyclical depending on the time of day and day of week. Two main components of time of day variability can impact efficiency from an airport's perspective. The first is that demand changes over the course of the day and the other being possible ATC staffing variability. The FAA dynamically changes staffing levels including number of controllers by level of experience based on the demand at any given airport *(13)*. Therefore, the Shift variable was created in alignment with the ATC typical 8-hour shift change times in order to include this variability into the analysis. Additionally, a day of week variable was added to capture any weekly patterns.

In order to capture the effect of flight trajectory inefficiencies within the terminal airspace, we derived two variables from the ASPM data by calculating the difference between the flight bearing at 100 nmi and the flight bearing at 40 nmi, and the difference between the flight bearing at 40 nmi and the direction of the arrival runway. Bearings in the data are measured relative to the airport's center clockwise from the magnetic north, ASPM reports these values from radar data for each flight in the data set where applicable. These two variables were created as measures of the amount of maneuvering needed by the aircraft inside the terminal airspace. For example, a flight entering the 40 nmi circle from bearing 180 (from the south) needs more maneuvering to land on Runway 9 compared to Runway 36 due to the turns needed to line up with the runway for final approach.

Table 2.1 Definition of Variables

Variable Name	Definition	Source
Dependent Variable		
$T_{100\ to\ ON}$	Actual time spent in last 100 nmi (seconds)	ASPM
Independent Variables		
IMC	Meteorological Condition at the airport during the 15 minute interval of arrival • Instrument Meteorological Condition (IMC) • Visual Meteorological Condition (VMC)	METAR
Wind Speed	Wind speed at arrival airport (knots)	METAR & ASPM
Wind Direction	Wind direction arrival airport (degrees)	METAR & ASPM
Physical Class	Aircraft engine type (Piston, Jet, Turboprop)	ASPM
WTC	Wake Turbulence Category (Heavy, Medium, Light)	BADA & ICAO standards
Ceiling	Ceiling height at airport during the 15 minute interval (10,000 ft)	METAR
Visibility	Visibility distance at airport during the 15 minute interval (miles)	METAR
Arrival Runway	Arrival runway used by flight	ASPM
Arrival Fix	Arrival bearing cluster at 40 nmi form arrival airport	ASPM
Traffic Demand Ratio	15 minute arrival throughput/declared capacity (ratio in decimal form)	ASPM
GDP	The presence of a Ground Delay Program impacting the arrival airport (binary)	OPSNET
Shift	Working Shift • Day starting at 6AM • Evening starting at 2PM • Night starting at 10 PM	Based on 8-hour daily shifts assumption at ATC facilities
Day Of Week	The day of week on arrival	ASPM
Distance	Distance between origin and destination airports (nmi)	Calculated Great Circle Distance
Short	Indicate if the distance from origin airport is less than 200 nmi (binary)	Calculated Great Circle Distance
Difference in bearing from 100 to 40 nmi	Absolute difference between flight direction at 100 nmi and 40 nmi (degrees)	ASPM
Difference in bearing from 40 nmi to runway	Absolute difference between flight direction at 40 nmi and the runway direction (degrees)	ASPM
Mixed use	Indicate if the flight's arrival runway is in mixed use mode (binary)	ASPM
Crosswind	Calculated from wind speed and direction at the airport surface vs runway direction (knots)	METAR & ASPM
Tailwind	Calculated from wind speed and direction at the airport surface vs runway direction (knots), negative values indicate headwind	METAR & ASPM
Crosswind@10,000 ft.	Crosswind relative to the flight path from 100nm to 40 nmi, calculated from wind components (knots)	RAP (NOAA)

It must be noted that situations such as GDP-exempt international flights, runway closures due to construction and snow, and Time-Based Flow Management (TBFM) implementation are not directly incorporated into the models due to data availability limitations.

Our models include these indirectly. For example, runway closures due to snow or construction would impact the capacity (denominator) of the Traffic Demand Ratio. These capacities are the called rates set by the air traffic manager every 15 minutes. If there is a runway closure due to snow or construction, it would show up in this metric. Future studies on predicting the A100 time could expand on including these factors in a more direct manner. Also, future research could expand the data evaluated to include a full year instead of the three months that we used to develop the model. This is especially important to include more variability in weather conditions and traffic patterns, which is indirectly accounted for in this study's models.

The Traffic Demand Ratio variable can be impacted by other variables in the study such as the day of week and time of day since there are weekly and intraday traffic patterns that impact the numerator of the ratio. However, that impact does not cause multicollinearity, that is because demand ratio is calculated at the quarter hour level whereas the time of day is calculated as 8-hour shifts so Demand Ratio is intended to capture variations at a more aggregate level compared to the day of week and Shift. Therefore, we found that it was beneficial to keep all three variables in the models. Similarly, weather conditions can impact the denominator of the ratio, i.e., capacity of the airport, but not to the extent of causing collinearity

Table 2.2 provides the summary statistics for the numerical variables in the final combined dataset after removing records with missing data. There were 295,792 arrival flights in total for all 5 airports, the split was 30%, 32%, 15%, 10%, 13% for ORD, ATL, SFO, JFK, and LGA, respectively. Interpreting **Table** 2.2 shows that the average A100 time is 25.5 minutes. Also, it is possible for the demand ratio to exceed 1, where 7% of observations in the data it was greater than 1.

Table 2.2 Summary Statistics of Numerical Variables

Variable	Unit	Observations	Min	Mean	Max	Std Dev
Time_100_ON	Seconds	295792	773	1530	6937	286.46
Diff_100_40	Degrees	295792	0.00	17.89	343.64	43.73
Diff_40_RWY	Degrees	295792	2	106.1	313	78.44
DEMAND_RATIO	Dimensionless	295792	0.030	0.696	2	0.252
Dist	nmi	295792	100.73	717.9	9806.6	588.18
TAILWIND	Knots	295792	-15.588	5.915	39.47	5.72
CROSSWIND	Knots	295792	0.00	5.395	28.19	4.40
VISIBILITY	Miles	295792	0.00	9.036	10	2.41
CEILING	10,000 ft	295792	0.003	1.541	3	1.20
Crosswind@10,000	Knots	295792	0.00	18.54	80.88	12.9

2.4 Methodology

We developed statistical models to investigate how various factors from the data impact flight time within the terminal airspace. These statistical models cover operations at both the airport and runway levels. Log-linear regression (to the base *e*) was the primary method, where the statistical model in its general form is presented in **Equation (2.3)**

$$\log(Time_{100-on}) = X\beta + \varepsilon \quad (2.3)$$

$Time_{100-on}$: Vector of travel times of arrival flights from 100 nmi to wheels-on (seconds)

X: The design matrix containing explanatory variables

β: Parameters vector

ε : Random error vector

All predictor variables used in this study are known at the time of entering the 100 nmi circle. Upon examining these variables, some needed transformations in order to adhere to the linear model assumption of normality before estimating the model. Log(Time) was used as the dependent variable instead of the actual time. Using the natural log produced an approximately normal distribution of the dependent variable (log(Time)) compared to the skewed distribution of the original data (Time). Similarly, the market distance was transformed using log which reduced the effect of a few long-distance observations in the data, where 99.6% of studied flights were 2500 nmi or shorter. Arrival runway end was used as a variable in the airport-level models. Some runways were rarely used for arrival operations. Therefore, they were grouped under a new arrival runway category "Others" because they were not statistically significant.

This study contains two sets of models. The first are airport-level models that include the grouping characteristics currently used by the FAA in benchmarking flight times within the last 100 nmi, namely arrival fix, runway end, physical class, and WTC *(2)*. The model estimation process began with an initial model using those main four predictors as categorical variables, then other external variables were used additively, i.e. by adding one variable at a time in a step-wise approach. This approach was adopted to keep the most significant variables in the model with the goal of balancing the model's complexity and accuracy.

The second are runway-level models, which stratified the data by the arrival runway end then multiple linear regression models were estimated for each individual runway end. Therefore, the starting model for each runway end consisted of the same main four predictors excluding the runway end since it would be constant within each model. The purpose of creating runway-level models was to explore how the impact of variables changes based on the layout and orientation of the runways at the airport. A motivating example is the uncertainty in the route assigned to arrivals within the 40 nmi circle even when using the same arrival fix and runway. **Figure** 2.2 shows different flight paths taken by select arrivals to ORD Runway 28C from its 309° arrival entry at the 40 nmi. Even when the arrival flights use the same fix and runway, the flights can be assigned to different tracks. As in this case, these tracks are significantly different in terms of distance, which impacts the A100 time. This variation led to including not only airport-level, but also runway-level models.

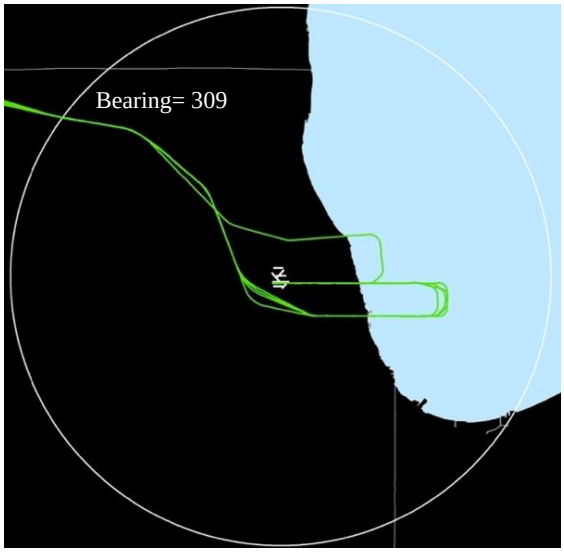

Figure 2.2 Flights with Different Trajectories Using the Same Arrival Fix and Runway

The data used in estimating each model is first split into two subsets, 80% as a training subset and the remaining 20% for testing. A bootstrap technique is implemented in order to evaluate model prediction accuracy. That is, random resampling without replacement is employed to fit each model over a large number of iterations. Then the selected model is validated using the 20% testing subset as out-of-sample comparison and the resulting Root Mean Squared Errors (RMSE) are reported. Finally, the distributions of RMSE resulting from the bootstrap are used to evaluate each model's accuracy and generalization ability.

2.5 Results and Discussion

This section covers the results of identifying how external factors outside the control of ATC impact the A100 travel times. The results of this study are contained in two sections. The first section evaluates the A100 travel times across the 5 select large hub airports in the United States being ORD, ATL, SFO, JFK, and LGA. The second section hones in on a single airport, ORD, to analyze the runways on an individual basis.

2.5.1 Comparison across Airports

In the models, the same variables were included across all airports to make a meaningful comparison. **Figure** 2.3 shows the correlations between all numerical continuous variables in the models. The difference in bearing from 40 nmi to the runway (Diff_40_RWY) had a positive correlation with the dependent variable, an increase in the difference is correlated with an increase in the A100 travel time. More maneuvering makes maintaining separation more challenging and might impact the ATC efficiency during busy periods. Ceiling and visibility had negative correlation coefficients, as expected, as a reduction in ceiling or visibility impacts the runway capacity negatively which can contribute to increased A100 travel times.

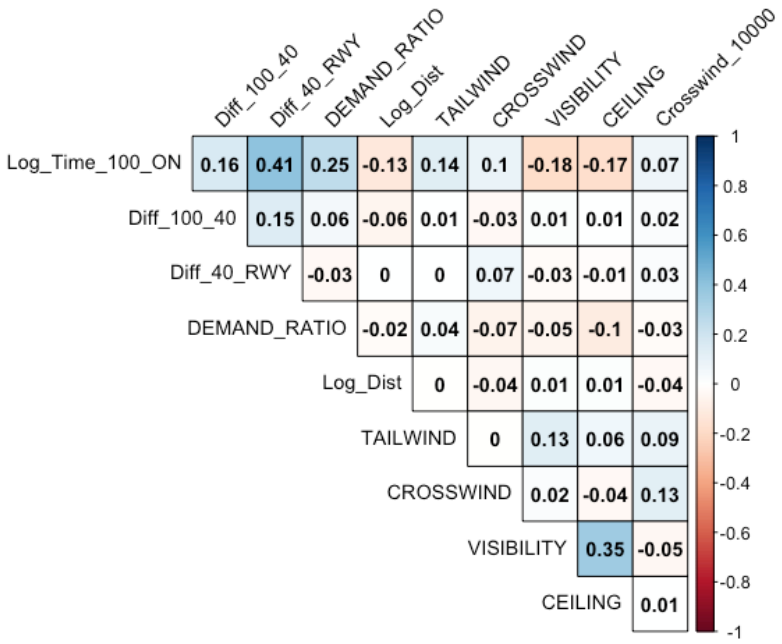

Figure 2.3 Correlation of Continuous Variables

Table 2.3 outlines all variables and their estimates, standard errors, and significance levels, where the number of observations reflect the 80% of the data that was used to estimate each model. There are multiple wind related variables that have significant impacts on the travel times within the last 100 nmi, these include crosswind and tailwind at the runway level in addition to crosswind at 10,000 foot altitude. The positive estimates for both crosswind variables indicate that an increase in crosswind results in increasing the A100 travel time, this is intuitive since these conditions can lead ATC to increase spacing or even put the aircraft in a holding pattern in extreme cases for safety considerations. ATL was the only airport that resulted in a negative estimate for surface crosswind, where wind is not typically an issue at that airport. This explanation was supported by exploring the data which shows that the crosswind rarely reached maximum allowable values at ATL. The model results show that increased tailwinds contribute to increasing the travel times, this is because increased tailwinds usually result in unfavorable conditions for landing due to possible controllability issues or increased approach ground speeds that can increase the risk of runway overruns.

Meteorological conditions in the estimated model show that A100 travel times are expected to increase during IMC periods, this is due to increased separation criteria compared to operations under VMC. For example, interpretation of the log-linear model shows that A100 travel times for arrivals at ATL are expected to increase by 6.5% on average for flights under IMC compared to flights under VMC, all else equal. Similarly, an increase in ceiling or visibility results in a reduction in A100 travel times across all airports.

The demand to capacity ratio at the time of arrival has a significant impact on the amount of time spent in the terminal airspace across all 5 studied airports. This variable captures the effect of ATM tactical measures to manage capacity constraints during periods of increased

demand. Our estimated model predicts less efficient flight trajectories in terms of time when the demand ratio is increased. All airports also increased the travel times when GDPs were implemented, the impact was greater at JFK as our model estimates a 15% increase in A100 travel times on average compared to non-GDP time periods. The impact on JFK could be attributed to its busy and complicated airspace which makes recovery more challenging. Additionally, the time of day had a significant but yet very different impact across all airports.

It was found that flights arriving from airports less than 200 nmi have an increase in their travel time within the terminal airspace. This could be a result of their different vertical profiles since they do not reach the same cruise altitude as other flights prior to entering the 100 nmi circle. Our model estimates a 4.5% increase of travel times within the terminal airspace for flights arriving from airports within 200 nmi from SFO. The market distance variable was also significant, the model shows that the further the departure airport the less time flights spend in the last 100 nmi. This could be as a result of being in communication with ATC during various stages of flight for a longer time which allows for smoother operation.

The impact of the day of week varied within each airport as well as across airports when compared to Fridays, the base category. An interesting observation at SFO and JFK is that all days are expected to result in shorter A100 travel times compared to Fridays on average. The shift variables were significant with the exceptions of evenings at LGA and nights at ATL. Although some categories were not statistically significant, we kept both the day of week and shift variables in the model as they are considered significant as a whole.

The physical class of the aircraft shows that both piston and turboprop aircraft are expected to have an increase in A100 travel times compared to jets as expected due to the differences in their speed profiles. Medium aircraft out of the WTC variable are expected to spend less time than light and heavy aircraft. Omitted physical class coefficients at SFO and JFK are due to an insignificant number of arrivals in those physical classes.

The increase in the difference in bearing between crossing the 100 nmi circle and 40 nmi resulted in longer travel times across all airports. This is an intuitive result since it indicates that when the aircraft needs to make turns to change direction, it increases the travel time due to the longer path and potential speed reduction while maneuvering. Similarly, an increase in the difference in bearing between crossing the 40 nmi circle and the runway direction is expected to increase the travel time.

The model predicts that when the arrival runway is being used in mixed mode (arrivals and departures), it increases the A100 travel time. This variable was intended to capture the decrease in arrival capacity resulting from implementing the arrival-departure separation by ATC when operating in mixed mode, therefore, it is expected to have a higher impact at airports with fewer runways. Our model estimates that mixed mode operations will increase travel times by 4% at LGA which has only two runways, fewer than all other four airports in this study. Additionally, the impact of arrival fix and runway used was significant but varied for each airport since they all have a different number of runways and their unique runway-fix combinations. The next section looks into the effects on each individual runway.

Table 2.3 Airport-Level Models Specification

	ORD			ATL			SFO			JFK			LGA		
Variable	Estimate	Std. Error		Estimate	Std. Error		Estimate	Std. Error		Estimate	Std. Error		Estimate	Std. Error	
(Intercept)	7.2911	0.01	***	7.2426	0.01	***	7.0942	0.01	***	7.2921	0.01	***	7.4096	0.10	***
CROSSWIND	0.0012	0.00	***	-0.0011	0.00	***	0.0023	0.00	***	0.0026	0.00	***	0.0013	0.00	***
TAILWIND	0.0027	0.00	***	0.0023	0.00	***	0.0032	0.00	***	0.0045	0.00	***	0.0035	0.00	***
IMC	0.0217	0.00	***	0.0635	0.00	***	0.0285	0.00	***	-0.0123	0.00	**	0.0598	0.00	***
DEMAND RATIO	0.1343	0.00	***	0.1772	0.00	***	0.1120	0.00	***	0.1844	0.00	***	0.1631	0.00	***
GDP	0.0975	0.00	***	0.1032	0.00	***	0.0770	0.00	***	0.1405	0.00	***	0.0695	0.00	***
DAY:MON	-0.0030	0.00	*	-0.0104	0.00	***	-0.0208	0.00	***	-0.0243	0.00	***	-0.0181	0.00	***
DAY:SAT	0.0022	0.00		-0.0027	0.00	.	-0.0044	0.00	.	-0.0535	0.00	***	-0.0250	0.00	***
DAY:SUN	-0.0045	0.00	**	-0.0060	0.00	***	-0.0054	0.00	*	-0.0346	0.00	***	-0.0230	0.00	***
DAY:THU	0.0033	0.00	*	0.0030	0.00	*	-0.0081	0.00	***	-0.0244	0.00	***	-0.0220	0.00	***
DAY:TUE	-0.0023	0.00		0.0048	0.00	***	-0.0219	0.00	***	-0.0218	0.00	***	-0.0100	0.00	***
DAY:WED	0.0001	0.00		0.0018	0.00		-0.0131	0.00	***	-0.0368	0.00	***	0.0045	0.00	
SHIFT:EVE	0.0023	0.00	**	-0.0022	0.00	**	-0.0194	0.00	***	0.0269	0.00	***	0.0010	0.00	
SHIFT:NIGHT	-0.0089	0.00	***	0.0012	0.00		0.0134	0.00	***	0.0088	0.00	***	0.0090	0.00	**
PHYS_CLASS:P	0.4051	0.02	***	0.4721	0.02	***	-	-	-	0.4566	0.06	***	0.4772	0.12	***
PHYS_CLASS:T	0.3780	0.02	***	0.2113	0.01	***	0.2147	0.01	***	-	-	-	0.1356	0.10	
Diff_100_40	0.0005	0.00	***	0.0026	0.00	***	0.0004	0.00	***	0.0001	0.00	***	0.0001	0.00	***
Diff_40_RWY	0.0011	0.00	***	0.0008	0.00	***	0.0093	0.00	***	0.0003	0.00	***	0.0011	0.00	***
CEILING	-0.0094	0.00	***	-0.0116	0.00	***	-0.0414	0.00	***	0.0036	0.00	***	-0.0067	0.00	***
VISIBILITY	-0.0044	0.00	***	-0.0045	0.00	***	0.0019	0.00	***	-0.0110	0.00	***	-0.0026	0.00	***
WTC:L	0.0905	0.02	***	0.0908	0.01	***	0.0218	0.01	.	-0.0290	0.03		-	-	-
WTC:M	-0.0083	0.00	***	-0.0158	0.00	***	-0.0038	0.00		-0.0139	0.00	***	-0.0355	0.10	
Mixed:YES	-0.0211	0.01	***	0.0560	0.00	***	0.0145	0.00	***	0.0382	0.01	**	0.0407	0.00	***
Crosswind_10000	0.0002	0.00	***	0.0003	0.00	***	0.0001	0.00		0.0005	0.00	***	0.0001	0.00	*
Log(Dist)	-0.0130	0.00	***	-0.0183	0.00	***	-0.0071	0.00	***	-0.0040	0.00	**	-0.0076	0.00	***
Dist_Short:YES	0.0249	0.00	***	0.0069	0.00	***	0.0441	0.00	***	0.0351	0.00	***	0.0130	0.00	***
Runway and Fix Indicators Included															
Adjusted R-squared	0.541			0.536			0.537			0.587			0.437		
Observations	70440			93277			45129			30975			38323		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Runway and Fix indicator variables included in all models

In addition to building our models described in **Table** 2.3, we used the same dataset to fit baseline models for comparison purposes. These baseline models used only the four basic variables from the existing benchmarking method mentioned in the introduction: Arrival Fix, Runway, WTC, Engine Type.

Figure 2.4 shows the resulting RMSE distributions from fitting and testing our New Models from **Table** 2.3 and the baseline models for 100 bootstrap iterations. It shows that our new models generally have an improved prediction accuracy compared to the baseline method seen through the smaller RMSE values. Comparison of the RMSE medians shows the improvement ranged from 39 to 50 seconds across the airports. When comparing the new models' results across airports, accuracy was higher for ORD and ATL through their lower errors and less variability in the results. This could be a result of the smaller number of observations at the remaining three airports, additional explanatory variables could also be needed to capture other impacting factors to improve the performance of the models.

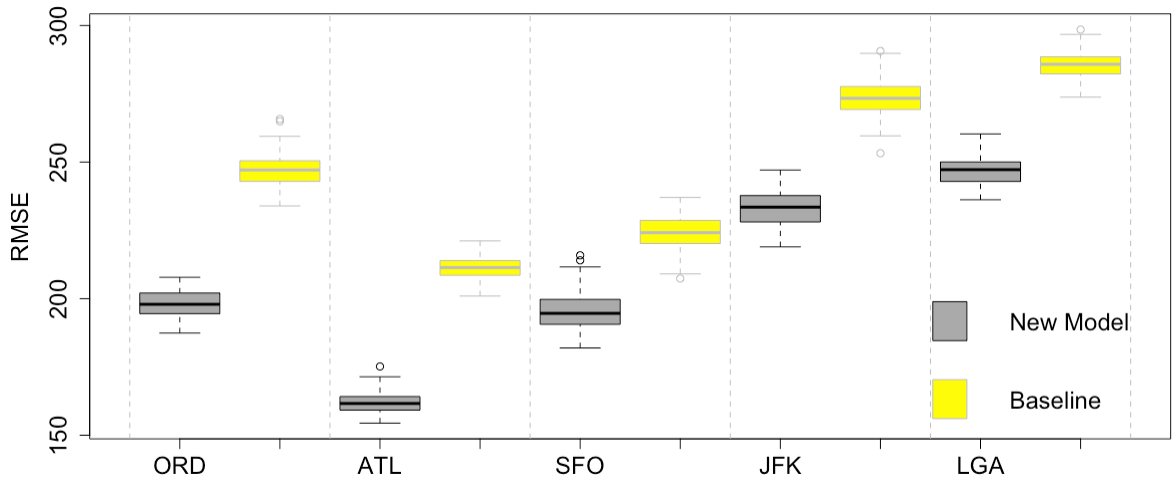

Figure 2.4 RMSE distributions for all Airports

2.5.2 Comparison across Runways

The runway-level models included less variables compared to the airport-wide models because breaking down the data by individual runways resulted in multi-collinearity between some of the variables. WTC and engine type variables were highly correlated for individual runways due to common operational practices at the airport such as routing medium jets to a certain runway and routing piston aircraft to another. Similarly, the "difference in bearing from 40 nmi to runway" variable was removed since it was highly correlated with the arrival fix, this is because the runway is fixed for each subset so the arrival fix was enough to cover that effect. The resulting models show that each arrival fix has a very different impact on the A100 travel times due to the difference in required maneuvering to sequence arrival flights approaching from different directions into the same runway.

Similar to airport-wide models, runway-level models show that low winds, high ceilings, and high visibility are significant in decreasing the travel times within the terminal airspace.

They also predict an increase in travel times during IMC weather, GDPs, and periods of high demand. The results show expected a decrease in A100 travel times when the departure airport is further away. Additionally, Medium aircraft are expected to spend less time compared to heavy. **Table** 2.4 outlines the estimated models including all variables and their estimates, standard errors, and significance levels.

Table 2.4 Runway-Level Models Specification (ORD)

	10C			9L			27L			27R			28C		
Variable	Estimate	Std. Error		Estimate	Std. Error		Estimate	Std. Error		Estimate	Std. Error		Estimate	Std. Error	
(Intercept)	7.5108	0.0158	***	7.4090	0.0486	***	7.3748	0.0129	***	7.3893	0.0428	***	7.2928	0.0144	***
CROSSWIND	0.0024	0.0002	***	-0.0008	0.0003	**	0.0014	0.0001	***	0.0025	0.0002	***	0.0004	0.0002	.
TAILWIND	0.0032	0.0002	***	0.0028	0.0002	***	0.0032	0.0001	***	0.0024	0.0002	***	0.0018	0.0002	***
IMC	0.0196	0.0032	***	0.0527	0.0034	***	0.0023	0.0020		0.0316	0.0027	***	0.0227	0.0030	***
DEMAND_RATIO	0.1932	0.0050	***	0.1377	0.0057	***	0.1017	0.0032	***	0.1395	0.0047	***	0.1801	0.0049	***
GDP	0.0438	0.0065	***	0.0309	0.0068	***	0.1001	0.0035	***	0.0611	0.0044	***	0.0949	0.0050	***
DAYMON	-0.0232	0.0040	***	0.0172	0.0042	***	-0.0098	0.0023	***	0.0102	0.0031	**	0.0052	0.0035	
DAYSAT	-0.0065	0.0037	.	0.0187	0.0040	***	-0.0011	0.0025		-0.0062	0.0037	.	-0.0121	0.0039	**
DAYSUN	-0.0197	0.0038	***	0.0121	0.0041	**	-0.0074	0.0025	**	0.0113	0.0035	**	-0.0218	0.0037	***
DAYTHU	-0.0139	0.0037	***	0.0146	0.0039	***	0.0040	0.0023	.	0.0029	0.0032		0.0025	0.0034	
DAYTUE	-0.0057	0.0043		0.0194	0.0046	***	-0.0153	0.0022	***	0.0082	0.0030	**	0.0047	0.0033	
DAYWED	-0.0275	0.0040	***	0.0110	0.0043	*	0.0004	0.0023		-0.0046	0.0031		0.0214	0.0034	***
SHIFTEVE	-0.0101	0.0022	***	-0.0033	0.0022		0.0042	0.0013	**	0.0045	0.0017	**	0.0261	0.0020	***
SHIFTNIGHT	-0.0108	0.0042	*	-0.0183	0.0460		-0.0157	0.0027	***	0.0742	0.0195	***	-0.0194	0.0046	***
Diff_100_40	0.0010	0.0001	***	0.0009	0.0000	***	0.0003	0.0000	***	0.0003	0.0000	***	0.0005	0.0001	***
CEILING	-0.0038	0.0016	*	-0.0179	0.0017	***	-0.0076	0.0009	***	-0.0098	0.0012	***	-0.0104	0.0012	***
VISIBILITY	-0.0014	0.0005	*	-0.0024	0.0006	***	-0.0054	0.0003	***	-0.0070	0.0004	***	-0.0026	0.0005	***
WTCL	0.4881	0.0204	***	0.4599	0.0468	***	0.3878	0.0159	***	0.3830	0.0418	***	0.4896	0.0088	***
WTCM	-0.0340	0.0045	***	0.0371	0.0460		-0.0190	0.0088	*	-0.0758	0.0414	.	0.0029	0.0037	
Crosswind_10000	0.0006	0.0001	***	-0.0005	0.0001	***	0.0002	0.0000	***	0.0001	0.0001		0.0005	0.0001	***
Log_Dist	-0.0219	0.0019	***	-0.0136	0.0020	***	-0.0125	0.0013	***	-0.0125	0.0016	***	-0.0105	0.0018	***
Dist_Short:YES	0.0203	0.0045	***	0.0139	0.0038	***	0.0142	0.0027	***	0.0182	0.0030	***	0.0309	0.0039	***
ARR_FIX130	-0.0034	0.0052		0.0341	0.0037	***	-0.0569	0.0018	***	0.1617	0.0099	***	-0.0408	0.0056	***
ARR_FIX230	-0.2677	0.0052	***	-0.1785	0.0105	***	0.0707	0.0023	***	0.1273	0.0041	***	0.0491	0.0055	***
ARR_FIX309	-0.2580	0.0060	***	-0.2493	0.0024	***	0.0482	0.0028	***	0.0672	0.0019	***	0.0885	0.0067	***
Adjusted R-squared:	0.679			0.697			0.345			0.541			0.520		
Number of Observations	10732			7830			22842			11994			11859		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

We compared the prediction performance among runway models as well as the airport-wide model, we also included the baseline models for reference. **Figure** 2.5 shows the resulting RMSE distributions from running 100 iterations of the bootstrap algorithm. Comparison of the RMSE medians shows the improvement ranged from 20 to 39 seconds across the runways. The new runway models at ORD had slightly lower errors compared to the airport-wide model, however, there is more uncertainty in the output seen through the wider spread of their corresponding boxplots. Also, the effects of the low adjusted R^2 value for runway 27L in the new model means there is reduced improvement over the four variable baseline model.

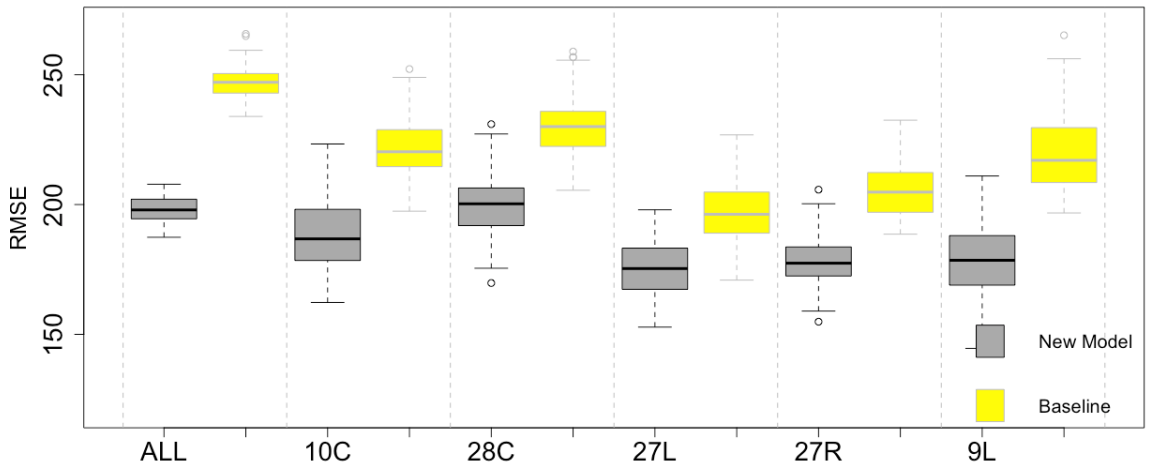

Figure 2.5 RMSE Distribution Comparison for Runway Models at ORD

2.6 Conclusions

Our study found that while the four characteristics used to benchmark flight A100 times are significant, many other variables outside the control of ATC are also significant. This means that the A100 delay metrics combine the results of external factors and ATC performance. Any change over time in an airport's A100 performance using the FAA methodology requires additional analyses as to the cause. Second, our models using only independent variables that are known when the flight is at the 100 nmi crossing are able to predict flight times within the terminal area with median RMSE improvements ranging from 39 to 50 seconds in the airport-level models.

Given the results can differentiate between the factors contributing to A100 performance, future research could focus more on the factors in control of the ATC. For example, our models contain the three shifts for ATC. It would be interesting to include staffing levels when each flight arrives to see how ATC experience impacts ATC performance. This would quantify how the performance of a Certified Professional Controller (CPC) on average compares with a Developmental or a Certified Professional Controller-In Training (CPC-IT).

2.7 Acknowledgments

We would like to thank Joe Napoli of the Federal Aviation Administration's Office of Performance Analysis for his feedback on results of the study.

2.8 Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Osama Alsalous, Dr. Susan Hotle; data collection: Osama Alsalous; analysis and interpretation of results: Osama Alsalous. Author; draft manuscript preparation: Osama Alsalous, Dr. Susan Hotle. All authors reviewed the results and approved the final version of the manuscript.

References

1. ICAO. KPI Overview 2019 [Available from: https://www4.icao.int/ganpportal/ASBU/KPI.
2. FAA/EUROCONTROL. Comparison of Air Traffic Management-Related Operational Performance: U.S./Europe. 2019.
3. Alsalous O, Galaviz-Schomisch R, editors. An Assessment of the Terminal Airspace Performance Indicator. 2018 Aviation Technology, Integration, and Operations Conference; 2018.
4. Choi S, Kim YJ, Briceno S, Mavris D, editors. Prediction of weather-induced airline delays based on machine learning algorithms. 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC); 2016: IEEE.
5. Knorr D, Chen X, Rose M, Gulding J, Enaud P, Hegendoerfer H, editors. Estimating ATM efficiency pools in the descent phase of flight. 9th USA/Europe air traffic management research and development seminar (ATM2011); 2011.
6. Cao Y, Kotegawa T, Post J, editors. Evaluation of continuous descent approach as a standard terminal airspace operation. 9th USA/Europe Air Traffic Management R&D Seminar; 2011: Federal Aviation Administration and EUROCONTROL Berlin.
7. Simić TK, Babić O. Airport traffic complexity and environment efficiency metrics for evaluation of ATM measures. Journal of Air Transport Management. 2015;42:260-71.
8. FAA. ASPM [Available from: https://aspmhelp.faa.gov/index/Aviation_System_Performance_Metrics_(ASPM).html.
9. Benjamin SG, Weygandt SS, Brown JM, Hu M, Alexander CR, Smirnova TG, et al. A North American hourly assimilation and model forecast cycle: The Rapid Refresh. Monthly Weather Review. 2016;144(4):1669-94.
10. FAA. OPSNET [Available from: https://aspmhelp.faa.gov/index.php/index/Operations_Network_(OPSNET).html.
11. Nuic A. User manual for the Base of Aircraft Data (BADA) revision 3.10. Atmosphere. 2010;2010:001.
12. ICAO. ICAO Aircraft Type Designators 2020 [Available from: https://www.icao.int/publications/DOC8643/Pages/default.aspx.
13. FAA. Air Traffic Controller Workforce Plan. 2019.

CHAPTER 3

Analysis of Deicing Capacity and Operations Using Simulation of ASDE-X Events

Abstract

Aircraft deicing at centralized deicing facilities during winter weather contributes to delays due to increased taxi out times. The main 34 airports in the United States reported 1,792,877 minutes of air traffic management delays related to ice and snow in 2019. In this research study, we use Airport Surface Detection Equipment, Model X (ASDE-X) (1) data at Chicago O'Hare International Airport (ORD) to analyze deicing operations. We identify deicing operations and extract deicing times for each aircraft entering the centralized deicing facility. The extracted deicing times per individual aircraft types are used to develop a simulation model to evaluate the efficiency of the centralized deicing facility. Our simulation model achieves this by tracking the number of minutes spent by aircraft in the system in addition to resource utilization during deicing operations. Currently at the study airport, individual lanes of the deicing facility are controlled by the airlines, determining which lanes a flight can use. We run multiple simulation scenarios by changing the queueing strategy and resource allocation at the airport and compare them to the current day base scenario. Our results show that if all deicing pads were open to all flights, removing airline control, a first come first served strategy with the pads having individual queues can save 25.1% of the aircraft minutes spent in the deicing system. Removing airline control using a first come first served with a combined queue approach reduced aircraft minutes by 17.8%. Assigning flights to lanes based on their mean deicing time (slow, medium, and fast groups) increased aircraft minutes by 1.3% on average.

3.1 Introduction

Deicing is the process of removing ice, snow, or slush from the aircraft surface after ice has already accumulated on the surface. Anti-icing is the prevention of the ice formation on the surface, where it entails spraying the surface with fluids that will delay the ice buildup for a period of time. The need for anti-icing depends on weather conditions such as temperature and precipitation. Sometimes the term deicing is used to describe the two steps combined, that is deicing followed by anti-icing. For the remainder of this paper, we use the term deicing to include deicing and anti-icing. Deicing is an important step in departures planning to ensure the safety of aircraft prior to takeoff because layers of snow or ice can have a negative impact on the aircraft performance. Ice contamination can alter the aerodynamic properties of the aircraft causing loss of control or reduction of generated lift from wings. Additionally, pieces of ice that break from the surface, during takeoff roll for example, can cause engine failure if ingested into jet engines (2, 3).

The need for deicing at an airport contributes to delays due to increased taxi out times. This delay is increased when aircraft have to taxi to centralized deicing facilities that are in remote locations away from the shortest path to their departure runway. The reported departure delays at the main 34 airports in the US showed that the primary causal factor is weather. In 2020, the Operations Network (OPSNET) delay data show that 76% of Traffic Management Initiatives (TMIs) delays were attributed to weather. Ice and snow conditions were responsible for 14% of weather delays in 2020, which is equal to 252,288 minutes of reported delays. In 2019, delays due to ice and snow were equal to 1,792,877 minutes (4). The 2019 delays are much higher compared to 2020 and considered more representative of the magnitude since air travel demand was at its peak before the COVID-19 pandemic caused it to drop significantly due to health restrictions in 2020.

One of the challenges to improving deicing efficiency is that current datasets do not report deicing delays consistently, most of the delays are attributed to the broader category of ice/snow. To put these delays in perspective, the aforementioned ice and snow delays in 2019 costed airlines approximately $133.1 million in direct costs. This was calculated using the US passenger carrier delay costs estimates published by Airlines for America. The total cost of delays is estimated to be higher when other components are included such as the value of a passengers' time and missed demand due to cancellations (5).

Deicing is important in airports during winter months, this operation takes place either at the gate or in a dedicated area known as centralized deicing facility. Centralized deicing facilities include deicing pads where aircraft can park to receive deicing treatment (6). The deicing pads are also referred to as bays or lanes. The process is performed using special deicing trucks that approach the aircraft from both sides, usually two to four trucks are deployed per deicing lane according to conversations with airline industry experts. The decision whether to deice at the gate or pad is a strategic choice by airlines, it depends on the size of the aircraft and the availability of trucks.

Since anti-icing fluids prevent ice formation on the aircraft surface for a limited time, the deicing process needs to be efficient to take place and send aircraft to their assigned departure

runway within that window of time. This time constraint is referred to as holdover time, the Federal Aviation Administration (FAA) defines it as the time starting at the beginning of deicing, or the beginning of the anti-icing step in a two-step process (6). Guidelines for holdover times are published by the FAA every year for planning purposes, the times depend on several factors including the outside air temperature, fluid type, and aircraft surface material. For example, the commonly used Type I fluid has a holdover time of up to 45 minutes (7). Managing the deicing process at the airport requires coordination of aircraft movement and scheduling their use of the limited resources. Minimizing the delays that happen as a result of the deicing process is of interest to airlines and airport management.

3.2 Literature Review

In review of existing literature, we consider studies that modeled deicing operations and the efficiency of existing queuing methods. Currently, there are a limited number of studies on this topic. Also, there are relatively few studies that derived deicing times from real data and there was a lack especially in having distributions of deicing times for individual aircraft types.

Norin et al. (8) studied the optimization of scheduling deicing vehicles with the objective of minimizing the delays due to deicing. They used flight schedule data from one day of February 2007 at Stockholm Arlanda Airport. The observed data showed that deicing times range from 1 to 18 minutes, while the setup time for each operation is 20 minutes. Their research evaluated optimization algorithms for scheduling and assigning deicing trucks. They simulated the use of Greedy Randomized Adaptive Search Procedures (GRASP) algorithm and compared it to a simple first come first served (FCFS) strategy for reference. They showed that using the optimized deicing vehicles schedules reduced flight deicing waiting times compared to the FCFS method currently used in practice. The simulation results showed that the GRASP algorithm reduced flight waiting times from 14 to 10 minutes on average.

A research study by Mao et al. (3) approached the deicing process as a planning and scheduling problem. They recognized that deicing involves resource management and making choices within resources capacity constraints. They developed an agent-based model with the objective to find a schedule where the total delay cost of all deicing aircraft is minimized, they calculated the aircraft delay as the delay in takeoff time. Their modeling scenario included one deicing station with three deicing bays. They used a random distribution of start of deicing times that is not derived from an actual deicing demand scenario. In their scheduling approach, a fixed deicing time of 20 minutes was used regardless of aircraft type. This last assumption is an approximation, while it is unclear if it was based on data, it excludes the variation in deicing times within the same aircraft type and from one type to another. Such variability impacts the service rate of deicing facilities in real life. They implemented a multi-agent model that is based on a first come first served (FCFS) queueing strategy, however, they added a de-commitment penalty. The de-commitment penalty is introduced to improve scheduling since individual agents will reserve deicing time slots when they have high confidence that they will make it to the deicing station on time. They simulated random deicing demand over a period of six hours, the simulation was repeated for different number of aircraft requesting deicing services during the six-hour period. Their experimental results show that FCFS with de-commitment penalty

improved the efficiency of the deicing process compared to a simple first come first served approach regardless of the number of aircraft. For example, delays decreased from 546 to 208 minutes in the case of 40 aircraft in the system.

Xing and Li (9) researched algorithms for scheduling aircraft deicing. They used a multi-agent model to improve the process of scheduling, they experimented with two algorithms. In their model, they assumed a centralized deicing facility with 6 bays. The model included variable deicing times based on three aircraft types, they assumed 4, 5, and 6 minutes. However, it was not clear where the times were derived from and what each aircraft type represents. They simulated the process of assigning aircraft to deicing bays by creating a schedule that minimizes the total departure delays. They used actual deicing demand during a two-hour period at Beijing Capital International airport. The output of their model shows an improvement when their decision-making algorithms are implemented compared to a simple first come first served method. They demonstrated the benefits of having a multi-agent model for deicing scheduling. .

Xing and Li (10) built on previous deicing scheduling research. They also ran a multi-agent model using data of domestic departures at Beijing Capital International Airport. In this paper they introduced an algorithm that varies priority of deicing requests. This priority algorithm is used to maximize resources utilization and reduce the overall delays. Their research presented the High Priority First Served (HPFS) algorithm, they also presented an algorithm that incorporates HPFS with a de-commitment penalty and airlines negotiation. The de-commitment penalty is implemented to reduce cases where airlines schedule a deicing time slot but do not use it, the negotiation part allows airlines to transfer time slots to others to avoid de-commitment. This algorithm is designed to promote airline equity and resource utilization in the distribution of deicing time slots between different airlines at the airport. Their simulation of 553 aircraft requiring deicing over a period of one day showed improvement in delays compared to a first come first served approach. Their simulation output showed that cumulative improvements increase as demand on deicing builds up, for example, there was a total of 225 minutes in delay savings for 350 aircraft in the system.

Xing and Lian (11) used a similar model also based on Beijing Capital International Airport departures data. That research focused on maximizing the resources utilization as a cooperative objective across airlines at the airport. They simulated two hours of departures that included 80 flights uniformly distributed over that time period. The simulation results showed that delays per airline was relatively evenly distributed. They concluded that a cooperative approach between airlines is an effective strategy to maximize the utilization of deicing resources at the airport, however they did not provide baseline results to compare with. Related research by Wu et al. (12) introduced a model to optimize the deicing process using data of domestic departures at Beijing Capital International Airport. Their model executes a priority comparison between flights in the deicing queue and schedules flights for deicing based on their scheduled departure times and airline. The results showed an improvement to the existing first come first served (FCFS) system, a simulation of 30 flights showed 308 minutes of delays in the case of FCFS, while the priority comparison method resulted in 242 minutes.

Zhu et al. modeled centralized deicing facilities operations at Incheon International Airport (ICN) in Korea (13). Their paper is the most relevant to our study within the reviewed literature. They used actual deicing data to estimate the time it takes aircraft to deice. In their model, they assumed that deicing times follow a normal distribution. They created distributions per aircraft categories based on their wingspan according to ICAO Aerodrome Reference Code. This categorization is helpful in calculating the capacity of each deicing pad. Also, their analysis showed that larger aircraft take longer to deice compared to small aircraft.

While our paper has similarities in methodology to Zhu et al. (13), there are noticeable differences. We create deicing time distributions for each individual aircraft type instead of grouped together by ICAO Aerodrome Reference Code. Also, deicing times do not follow a normal distribution and we find that small aircraft tend to take longer at our study airport. These differences are mostly driven by the way deicing operations are conducted in different airports, such as how many trucks are assigned to a small aircraft verses a large aircraft.

3.3 Research Motivation

From the reviewed research, there is a relatively small number of papers that study the efficiency of centralized deicing facilities. There is a lack of studies that use actual deicing times derived from ASDE-X airport radar data and disaggregated by aircraft type. Also, as shown by the literature review, limited knowledge is known about deicing processes and logistics at airports in the United States.

This paper examines deicing operations at Chicago O'Hare International Airport (ORD). We use airport surface radar data to analyze the flow of deicing operations taking place in the centralized deicing facility at ORD. We present deicing times distributions by individual aircraft type extracted using our deicing time extraction algorithm. We also simulate the operations of the centralized deicing facility at ORD considering the real-life airport strategy as our baseline scenario. Additionally, we simulate three additional scenarios with different queueing styles to investigate how deicing efficiency can be improved. In our simulation runs, we pull from the extracted deicing times distributions for each aircraft.

There are advanced simulation tools available for the industry such as Total Airspace and Airport Modeler (TAAM), Simmod PRO, and AirTOP that are capable of modeling and optimizing various airport operations including deicing. The deicing times used in these models draw from user-defined distributions, which would lead to using unrealistic default values if the user does not have access to data. One application of our analysis is to use our deicing times distributions with such simulation models as inputs for airports with similar characteristics. Our work is also useful for airport planners and for developing new simulation models.

3.4 Data

This research used radar data from the Airport Surface Detection-Model X (ASDE-X) system. ASDE-X is a surveillance system that uses radar, multilateration, and satellite technology to track the movement of aircraft and vehicles on an airport's surface. It was developed to improve the situational awareness of air traffic controllers to help reduce incidents. Currently, there are 35 airports in the United States equipped with ASDE-X systems (1). We

built our model using the airport layout data provided by the FAA. The layout data was in a shapefile format, the combination of layout and ASDE-X data was used to visualize the movement of aircraft on the airport surface and extract key deicing event information from the ASDE-X data. Section 4.1 discusses in detail the airport layout of ORD airport and Section 4.2 covers information on the ASDE-X data.

3.4.1 ORD Airport Layout

The layout data identified the location of the centralized deicing facility and the movement of aircraft inside and outside the deicing area. Our analysis included the following elements found in the shapefile data: runways, aprons and hardstands, and taxiways. The central deicing facility in ORD was included within the list of apron areas and hardstands. Currently, the central deicing facility has a total capacity of 10 lanes that can each serve a narrow body aircraft. The facility is divided into three areas or pads; 4 lanes in the south area, 2 lanes in the middle, and 4 in the north area. The south area is for the use of American Airlines and affiliates, the north area is for United Airlines and affiliates, while the remaining 2 lanes are for all other airlines who choose to not deice at the gate. Figure 3.1 shows the location of the deicing facility in the airport and the distribution of deicing areas.

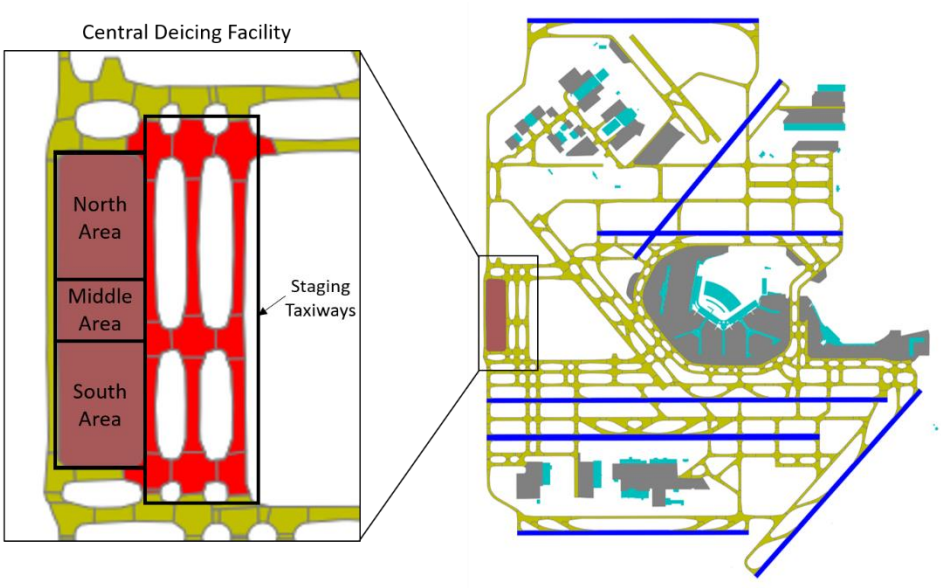

Figure 3.1 ORD Airport Layout showing the Central Deicing Facility

In our analysis, the number of lanes in each area represents the capacity based on the number of narrow body aircraft or small regional jets it can serve simultaneously. That is the case for current operations at ORD based on ASDE-X data. This assumption is essential for our simulation since the capacity of a deicing pad can change based on the size of aircraft using it. For example, the space required to deice two narrow body jets can be enough for only one wide body jet. Figure 3.2 illustrates utilizing a deicing pad by two Boeing 737 aircraft vs one Boeing 777, the deicing pad layout shown was adapted from the FAA design guidelines (6).

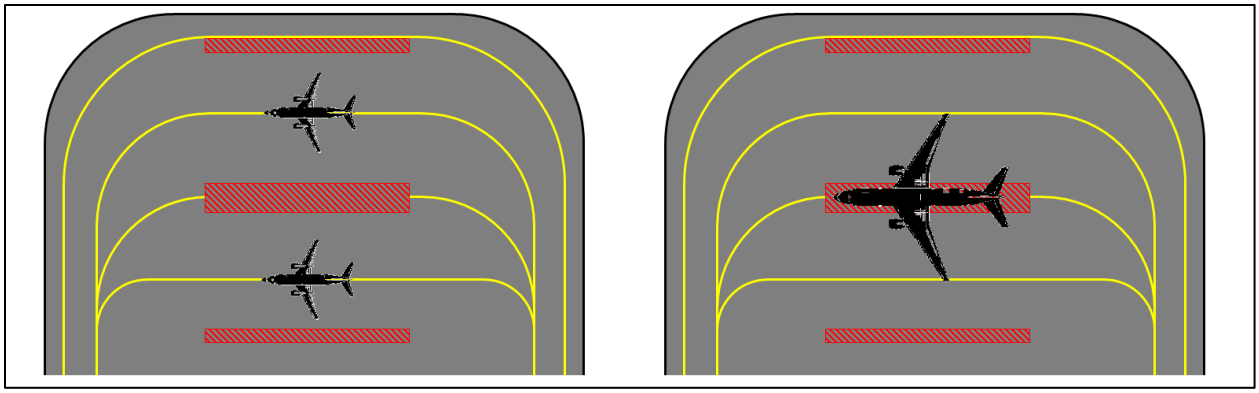

Figure 3.2 Space Required by two Narrow Body Jets Compared to a Wide Body Jet

3.4.2 ASDE-X Data Analysis

This research used data for the month of February 2020 which included 22 days that required aircraft to deice. The ASDE-X data contained information such as vehicle identification, vehicle position in latitude and longitude, and timestamp. The vehicle identification included the aircraft type and the aircraft identifier which is made up of the airline code and flight number. We used SeatGuru (14) to add the number of seats by aircraft type and airline combination in the data. There was a total of 41 unique combinations of airline and aircraft types that used the deicing facility in our ASDE-X data.

The ASDE-X system has an update rate of one second, we used this high-resolution data to calculate the time in seconds that each aircraft spent inside the deicing facility. During the analysis of aircraft movement, it was noticed that some aircraft go through stop-and-go movement patterns which could indicate queueing inside the deicing facility. This was validated by creating video animations that showed the movement of all aircraft at a given time period. We created individual speed profiles for each aircraft that entered the deicing facility to study their movement. We developed a deicing detection algorithm that analyzed each speed profile and identified time segments spent traveling, deicing, or waiting within the deicing facility. Examples of the visual output of the algorithm are shown in Figure 3.3. The x-axis shows the time as reported in the data in Unix Epoch, i.e. seconds since 1/1/1970 00:00:00. Note that once deicing and waiting segments are identified, the remaining time is automatically recorded as traveling time.

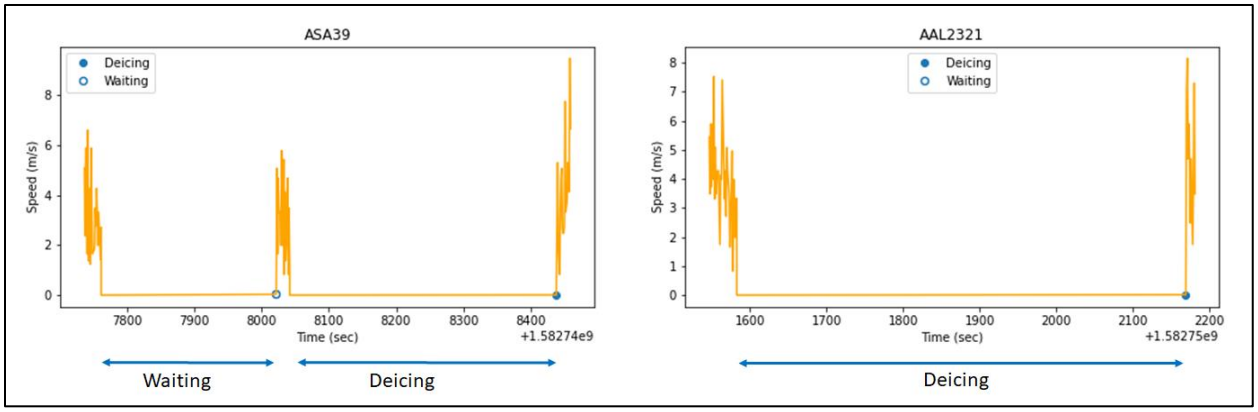

Figure 3.3 Speed Profiles Analysis Examples

Our deicing detection algorithm identifies all flights that enter the deicing facility. It returns a valid flight deicing entry based on the assumption that an aircraft completely stops and spends at least four minutes. This removes the few cases of aircraft using the deicing facility as a taxiway. The algorithm also detects when multiple valid entries are created by the same aircraft. In this case, the algorithm returns the most recent one as the flight's deicing event. Once a valid entry is recorded, the algorithm analyzes the speed profile and returns the time of entry, time of exit, deicing time, travel time, and waiting time if applicable (i.e. some flights do not have a waiting time). We included criteria in the algorithm to overcome noise in the radar data as there were occasions of unsteady position data in some of the flight tracks. In the ASDE-X data during February 2020, our algorithm detected 22 days where the deicing facility at ORD was used by a total of 1,504 aircraft. Figure 3.4 shows the distribution of the extracted times for all 1,504 deicing records.

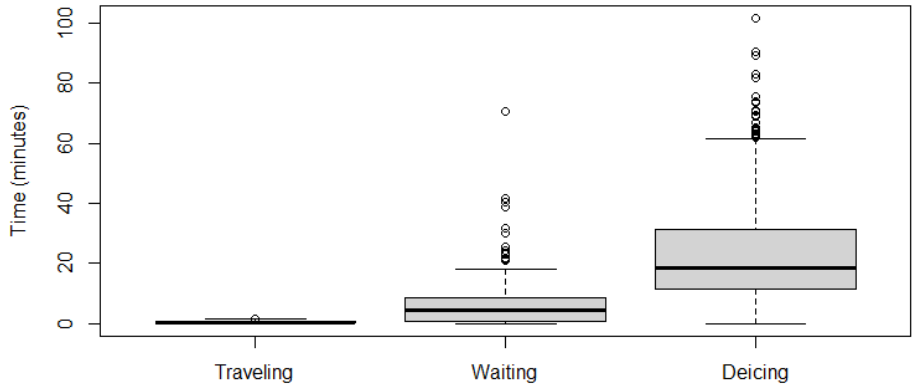

Figure 3.4 Box Plots of Times Spent Inside the Deicing Facility

The distributions in Figure 3.4 show a reasonable trend of how aircraft spend the time inside the deicing facility. Aircraft spent more time deicing compared to waiting, and traveling inside the deicing facility was relatively short as expected. The deicing times are considered closer to reality after separating traveling and waiting times. We acknowledge that some other inefficiencies may be included in the deicing time because they are difficult to detect from the data. For example, the deicing time may include the waiting time for communication with the

tower and the deicing crew, or waiting for deicing trucks to get in position. The results of our algorithm had a mean deicing time of 23 minutes and a median of 18.7 minutes for all 1,504 observations.

We further analyzed the deicing times by aircraft type groupings, where the data shows 19 unique aircraft types that used the deicing facility during the analysis time period. Figure 3.5 shows distributions of deicing times for the 10 aircraft types that had 30 or more records in the data. The minimum number of observations was set to 30 for statistical purposes in accordance with the Central Limit Theorem (15).

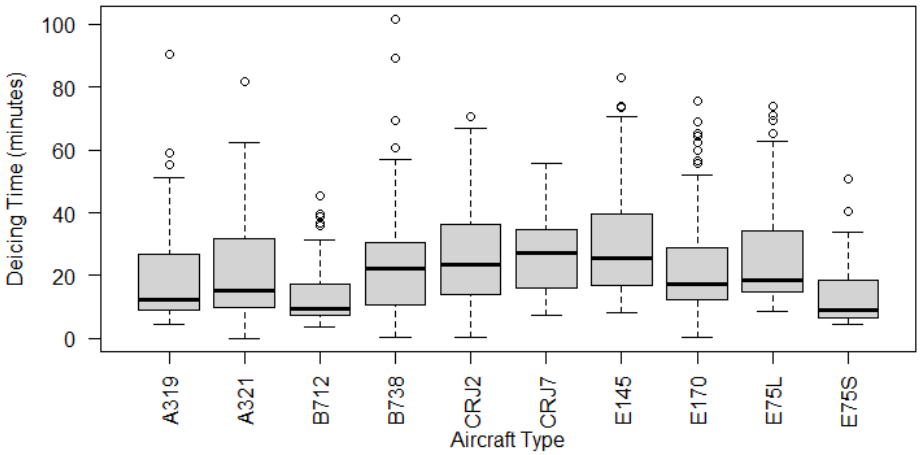

Figure 3.5 Distribution of Deicing Times by Aircraft Type

The data showed that larger aircraft tend to take less time deicing compared to small ones. For example, Figure 3.5 shows that the median deicing times for a narrow body jet with over 100 seats like an Airbus A319 or a Boeing 717-200 were 12.5 and 9.5 minutes, respectively. On the other hand, 50-seat regional jets such as Bombardier CRJ-200 or Embraer E-145 had a median of 23.4 and 25.5 minutes, respectively. These findings were explained by airline industry experts that airlines typically deploy two deicing trucks for small aircraft, while they tend to deploy three or more trucks for larger aircraft. Additionally, waiting for tower clearance after deicing is complete can be a source of longer times for regional jets. However, there is no indication in the data that can separate those events from the deicing times.

Additionally, our algorithm detects the first appearance of the aircraft into the adjacent taxiways, indicated as staging taxiways in Figure 3.1. Entry to the staging area is considered the beginning of demand that the aircraft is placing on the deicing facility. According to conversations with airline industry experts at ORD, the operations at the airport are organized in a way that queueing for the deicing facility starts in these taxiways.

3.5 Methodology

We developed a process-based discrete-event simulation model to analyze the deicing operations. The model simulates a queueing system and calculates the time spent in the system by each aircraft by drawing from the extracted deicing times distributions. We developed the

following simulation assumptions using observations from visualizing the data and conversations with airline industry experts at ORD:

1) The total capacity of the deicing facility is 10 lanes. The default capacity distribution is according to the three areas shown in Figure 3.1.
2) Aircraft arrival into the queueing system is the time an aircraft enters the staging taxiways shown in Figure 3.1, this time indicates the start of deicing demand.
3) The average speed of taxiing inside the staging area is 12 miles per hour.
4) When an aircraft is finished with deicing, the next aircraft in the queue starts moving and enters the deicing facility at the same time that the aircraft in front of it leaves. If the next aircraft in the queue has not traveled across the staging area, taxi time is included before it starts deicing.
5) The deicing time needed for each aircraft type will be drawn from a distribution derived from data. In the case of aircraft types that did not have a sample size of at least 30, the model opts for the average observed deicing time for that aircraft type.

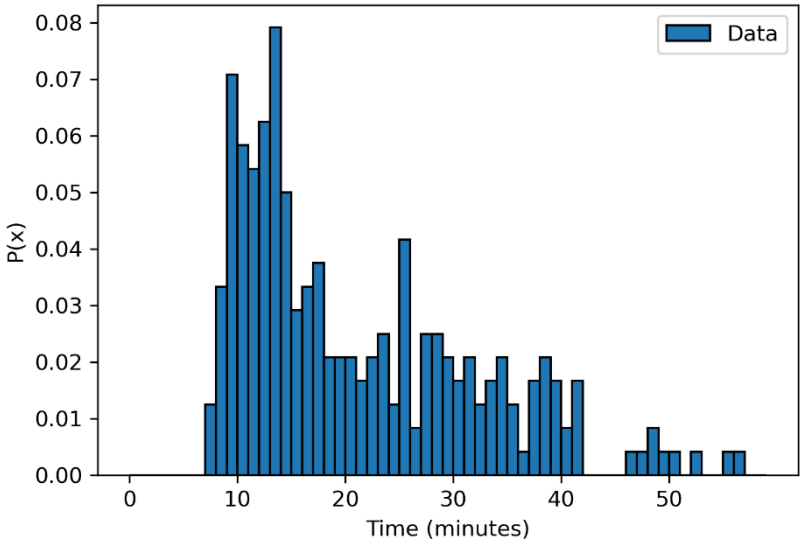

Figure 3.6 Deicing Times Distribution for Embraer E170 Data

Figure 3.6 shows the distribution of deicing times data for a single aircraft type. The observed deicing times are on the x-axis divided into one-minute bins while the probability of each bin is on the y-axis. It must be noted that each distribution was created for the data between the 2nd and 98th percentiles to eliminate outliers.

The deicing simulation model was built using SimPy package in Python. In the model, the airport is the simulation environment that contains deicing pads as capacity constrained resources. Aircraft that want to deice arrive into a queueing system to be processed. The processing function first generates the required deicing time for each aircraft by a random draw from the distribution for the aircraft type, if available. Otherwise an average deicing time is used. The next step in the function is to decide which deicing pad the aircraft is going to put a request on, this decision is made based on a user-defined resource distribution strategy. Finally, the

function puts each aircraft in the proper queue based on the implemented strategy. In this process, the simulation considers the time it takes to taxi from entering the staging taxiways to reach the deicing facility.

Sections 5.1 to 5.4 describe the four simulation scenarios that were tested in this study to see how different queuing techniques impact the total and standard deviation of aircraft minutes spent in the deicing facility. They use the demand evaluated from the ASDE-X data during February 26, 2020, which was a day where departures used the deicing facility at ORD. We selected a busy two-hour period from 8:00 to 10:00 AM local time for the simulation, when there were 38 aircraft placing demand on the deicing facility within that time window.

We developed the capability of generating random demand that has similar characteristics to traffic at the study airport. The demand is generated by two steps. The first step is generating a random time for the start of demand, this time is output as the headway in seconds between an aircraft and the next one entering the deicing system. Figure 3.7 shows the probabilities of headway times from the data and the distribution that our model uses to generate random times.

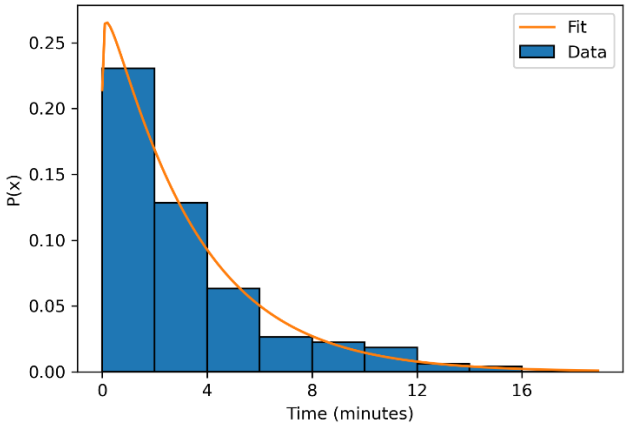

Figure 3.7 Probabilities of Start of Demand Headway

The next step is selecting a random aircraft type-airline combination for the generated start of demand time. This is achieved by drawing random pairs of aircraft types and airlines that are based on their probabilities in the deicing facility usage data. For example, there is a 16.4% chance of having a Boeing 737-800 operated by American Airlines placing demand on the deicing facility based on the data. Although we use actual demand in our study, this capability is useful to simulate what-if scenarios based on user-defined conditions such as increased traffic or changes in fleet mix.

3.5.1 Simulation Scenarios

The first scenario was the Baseline, it used all of the default assumptions to simulate the current operations of the central deicing facility at ORD. The purpose of this scenario is to validate our model's output when drawing from the deicing time distributions in comparison to the actual data. In this scenario, the two major airlines operate the deicing facility with dedicated areas for them and their affiliates. The total capacity of 10 narrow body aircraft is divided into

three areas as shown in Figure 3.8. American Airlines control four pads, United Airlines control four pads, and other airlines have access to the remaining two pads. Aircraft are assigned to one of the three deicing areas based on the operating airline, each deicing area has its own queue that takes place on the adjacent staging taxiways. Figure 3.8 shows the key points that aircraft pass through while in the deicing system. Aircraft enter the staging area from one of the two points marked with stars. Aircraft line up in their queue and enter the deicing area from points A, B, or C depending on the deicing area assigned to them.

In all of our simulation scenarios, aircraft are assumed to taxi at an average speed of 12 mph inside the staging area shown in Figure 3.8. Additionally, when an aircraft is done with deicing, the deicing facility operators coordinate with the next aircraft in the queue so that it starts moving and enters the deicing facility at the same time the aircraft that just finished exits. If the next aircraft was still far away from the entry point, taxi time is included in the process.

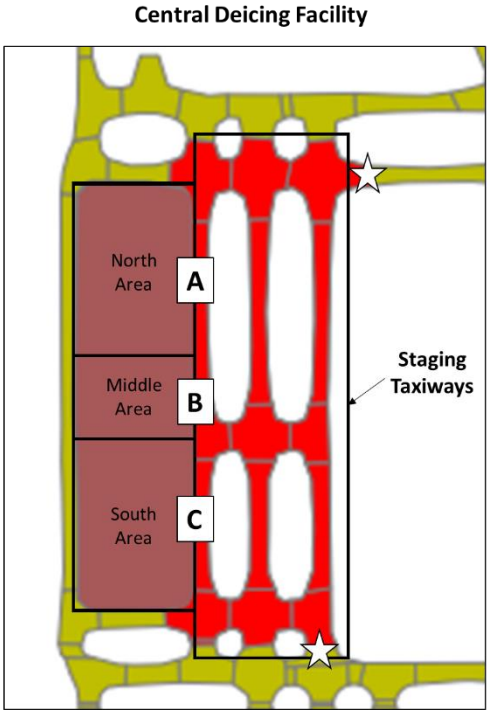

Figure 3.8 Key Points for Queueing at the Central Deicing Facility

The second scenario was the First Come First Served with individual queues (FCFSi). This scenario assumes that all three deicing areas shown in Figure 3.8 are available to all airlines equally. In this scenario aircraft also enter the staging area from one of the points marked by stars in Figure 3.8. Each one of the three deicing areas has its independent queue, aircraft line up and enter their designated deicing area from point A, B, or C. Each aircraft is put in the shortest queue at the time of start of demand, this means that aircraft are assigned to an area's queue when they enter the staging taxiways. The purpose of this scenario is to simulate operations if the FAA controlled the queueing process without changing the capacity configuration, hence we kept the individual queues.

The third scenario was dividing the deicing facility into slow, medium, and fast areas based on aircraft deicing times. This is similar to the Transportation Security Administration (TSA) type queuing with expedited lanes. The existing capacity configuration in Figure 3.8 was kept at four, two, four lanes for the south, middle, north areas, respectively. Keeping the configuration allowed for a comparison with the baseline scenario to evaluate the impact of changing the queueing strategy only. Therefore, the medium category was assigned to the south area, the fast to the middle, and the slow to the north area. Aircraft were classified into each category based on the average deicing times observed in the data. The classification of aircraft types can be seen in Figure 3.9.

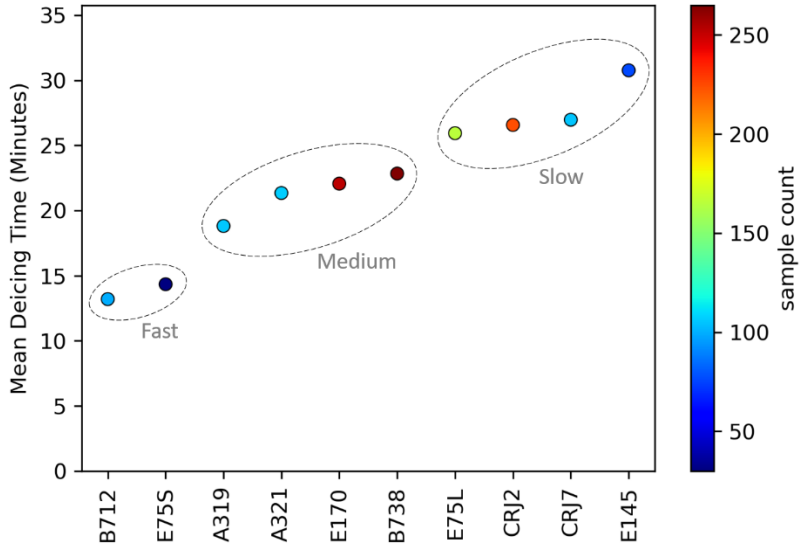

Figure 3.9 Mean deicing times per aircraft types with service speed categories

The classification of aircraft was done by a K-means clustering algorithm. Only the 10 aircraft types with at least 30 samples were used, the other nine types are assigned to the medium category by default. Figure 3.9 shows that aircraft with mean deicing time less than 15 minutes were considered fast, above 25 minutes were considered slow.

In this scenario, aircraft use the same points shown in Figure 3.8, they enter the staging area from the points marked with stars and line up in individual queues for each area. Their entry point to each area is one of the three points A, B, or C depending on the area they were assigned to. This scenario is intended to evaluate the benefits of reducing the wait time for aircraft in the fast and medium category since they will not have to wait behind an aircraft of the slow category. Additionally, to increase the efficiency of this scenario, aircraft are allowed to use other areas if their respective area was full, and if there is at least one empty lane in the other area.

The fourth scenario was the First Come First Served with a combined queue (FCFSc). In this scenario, we simulate operations if the FAA controlled the queueing for the entire deicing facility without dividing it to the three areas. In this case all 10 lanes are available for all aircraft in the queue equally, the next aircraft in line is assigned to the first available lane. As shown in Figure 3.10, aircraft enter from one of the points marked by a star and the queue begins at point

D. The simulation model assigns aircraft to open lanes as soon as they enter the staging area. If all lanes were busy at the start of demand, the assignment happens at the moment a lane is freed up and movement of aircraft is coordinated to enter the deicing facility at the same time the previous aircraft leaves.

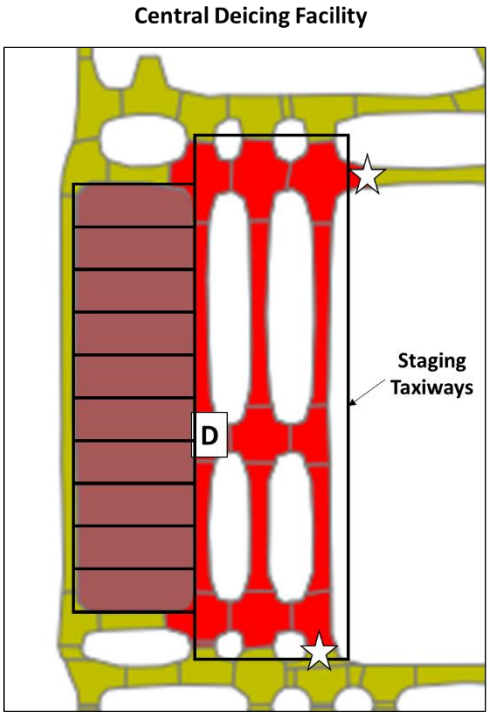

Figure 3.10 Key Points for Queueing at the Central Deicing Facility (FCFSc)

3.6 Results

Our analysis shows that during the time periods the deicing facility was active, 15.5% of all departures deiced at the central deicing facility. The remaining departing aircraft were assumed to have deiced at their gates. The breakdown of the number of records per aircraft type is shown in Table 3.1. The table shows the 19 types that were observed at least once in the deicing facility during our analysis time frame, i.e. February 2020.

Table 3.1 Deicing Facility Usage by Aircraft Type

Aircraft Type	Did not use facility (%)		Used facility (%)	
B712	124	(55.61%)	99	(44.39%)
MD90	4	(57.14%)	3	(42.86%)
MD88	14	(58.33%)	10	(41.67%)
CRJ9	36	(63.16%)	21	(36.84%)
E170	570	(69.26%)	253	(30.74%)
E75L	505	(75.60%)	163	(24.40%)
A321	391	(78.67%)	106	(21.33%)
E45X	41	(78.85%)	11	(21.15%)
E75S	127	(80.89%)	30	(19.11%)
B738	1135	(81.19%)	263	(18.81%)
CRJ2	1092	(82.79%)	227	(17.21%)
A319	569	(84.30%)	106	(15.70%)
CRJ7	1248	(92.24%)	105	(7.76%)
E145	1225	(94.16%)	76	(5.84%)
A320	564	(97.58%)	14	(2.42%)
B789	41	(97.62%)	1	(2.38%)
B739	581	(97.65%)	14	(2.35%)
B772	74	(98.67%)	1	(1.33%)
B77W	95	(98.96%)	1	(1.04%)

Wide body aircraft such as Boeing 777 rarely deice at the central facility as shown by Table 3.1. Airlines most likely prefer that since wide body aircraft take twice the space required by a narrow body aircraft, refer to Figure 3.2 for an illustration. This strategy makes the operations more efficient, when narrow body aircraft are sent to the deicing facility, airlines are able to free two gates. Making the gates available helps airlines avoid arrival flights delays and taxiway congestion. Additionally, this practice increases system predictability since the number of aircraft that can be accommodated by the deicing pads does not change.

3.6.1 Simulation Results

In this section, we present the results of our simulation model. The demand for deicing was created using ASDE-X data from the 26th of February, 2020 which was a busy day with many departures using the deicing facility at ORD. We selected a busy two-hour period from 8:00 to 10:00 AM local time for the simulation, there were 38 aircraft placing demand on the deicing facility within that time window. The same demand was used for all scenarios, random deicing times for each aircraft in the system were drawn from the distributions created by our ASDE-X data analysis algorithm. Table 3.2 shows the statistics of all simulation scenarios in addition to the actual operations, all simulation numbers are the average results of 100 simulation runs per scenario.

Table 3.2 Comparison of Simulation Results

Scenario	Number of aircraft	Average Aircraft minutes	Standard Deviation	Average Seat Minutes	Standard Deviation	Time last aircraft served (Hours)
Actual	38	34.2	18.1	3,549	2,275	2.49
Baseline	38	33.9	17.0	3,773	2,821	2.51
FCFSi	38	25.4	14.6	2,761	2,187	2.42
Slow, Medium, Fast	38	34.4	17.1	3,830	2,880	2.60
FCFSc	38	27.9	13.0	3,009	1,971	2.51

The FCFSi scenario simulated operations if Air Traffic Control (ATC) at the airport controlled all deicing areas instead of dividing them between airlines. The idea is to explore improvements in the capacity utilization since there are periods of times that one airline's departure schedule is busier than others. Table 3.2 shows that implementing FCFSi saves 8.5 minutes per aircraft on average compared to the baseline scenario. This saving is equivalent to 25.1% of the aircraft minutes spent in the deicing system and to 26.8% of seat minutes.

When each deicing area operates independently as in the baseline scenario, some might be under stress while other areas are underutilized. To help analyze capacity utilization, our model returns the number of aircraft in the queuing system at every step of the simulation. Figure 3.11 demonstrates resource utilization for all four simulation scenarios over the demand period of two hours. The capacity of each area is plotted as a horizontal line, periods of time when the number of aircraft line crosses above the capacity indicate the formation of a queue.

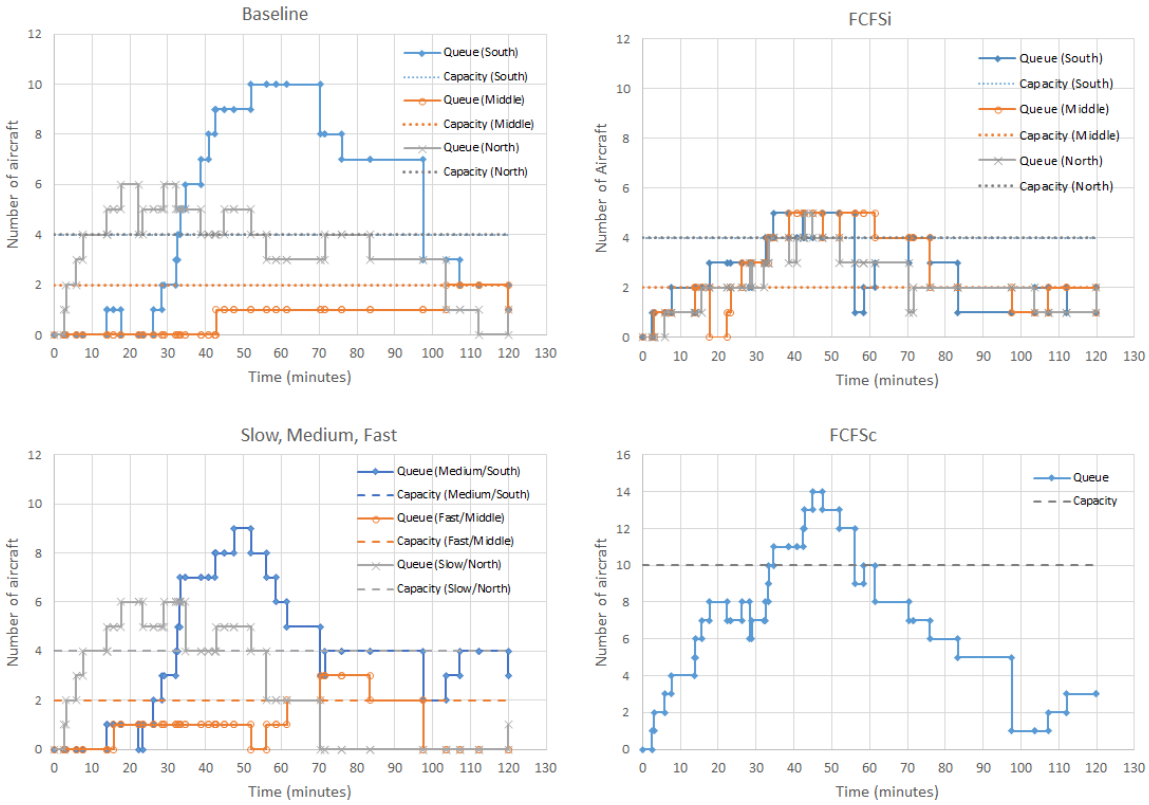

Figure 3.11 Number of Aircraft in the Queueing System for each Simulation Scenario

The comparison between the baseline and FCFSi portions of Figure 3.11 helps visualize the potential improvements by only changing the queue assignment criteria, while all other conditions remain equal. It can be seen that the FCFSi strategy leveled out the queues among deicing areas. The queue builds up quickly between 33 minutes and 52 minutes of simulation time during the baseline run. This buildup is due to the peaky departures schedule by airlines using that area. When FCFSi is implemented, the queue of the South pad is absorbed by other deicing areas reducing the imbalance that took place in the baseline scenario.

Additionally, the FCFSc showed improvement to the baseline scenario. Our simulations show that it is expected to save 6.0 minutes per aircraft on average, which is 17.8% when compared to the baseline scenario. This saving is also equivalent to 20.3% of the seat minutes spent in the deicing system. When FCFSc is compared to FCFSi, the improvement was smaller. This was because the FCFSi did not leave a lot of time periods of underutilized capacity, as shown in Figure 3.11. Underutilization of capacity in FCFSi could happen in the rare occasions when an aircraft is assigned a queue behind another aircraft that is at the tail end of the deicing distribution. Additionally, the taxi time inside the staging area was increased in the FCFSc scenario because they all lined up in a single queue then had to travel towards their assigned lane.

The results from Table 3.2 indicate that the deicing speed scenario (Slow, Medium, and Fast) returned similar numbers to the baseline. This was due to the two scenarios being analogous in the sense that the aircraft types with large sample sizes were dominated by a certain

airline. For example, 95% of all CRJ2 in the data was operated by airlines affiliated with United Airlines so they were directed to their deicing pad in the baseline and to the "Slow" pad in the other scenario. Similarly, 90% of B737-800 were operated by American Airlines and were directed to their deicing pad in the case of the baseline and to the "Medium" deicing pad in the deicing speed scenario. Due to this pattern in the data, the capacity was utilized similarly in both scenarios as Figure 3.11. Our results show that the deicing speed scenario was slightly less efficient as it increased the aircraft minutes in the system by 27 seconds on average. This increase is equivalent to 1.3% of mean aircraft minutes and 1.5% of mean seat minutes.

3.6.2 Deicing Rates per Aircraft Type

We calculated the deicing rates in aircraft per hour units using our simulation model. For the purpose of this calculation, we ran our model using continuous demand of a single aircraft type at a time. The model was run for deicing pad capacities ranging from one to four lanes per aircraft type. Each run simulated 10 hours of operations. The results are shown in Table 3.3.

Table 3.3 Deicing Throughput Rates (Aircraft/Hour) for Different Capacities

Aircraft Type	1 Lane		2 Lanes		3 Lanes		4 Lanes	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
A319	3.5	0.4	6.8	0.6	10.1	0.7	13.5	0.7
A321	2.9	0.3	5.9	0.5	8.8	0.6	11.7	0.6
B738	2.7	0.3	5.4	0.4	8.1	0.5	10.8	0.6
B712	4.6	0.4	9.3	0.6	14.0	0.9	18.5	0.8
CRJ2	2.3	0.2	4.6	0.4	6.7	0.5	9.1	0.5
CRJ7	2.3	0.2	4.5	0.3	6.8	0.4	9.0	0.5
E145	2.0	0.2	4.0	0.4	6.0	0.4	8.0	0.5
E170	2.8	0.3	5.5	0.4	8.4	0.5	11.2	0.6
E75L	2.4	0.3	4.7	0.4	7.1	0.4	9.4	0.6
E75S	4.8	0.4	9.7	0.7	14.6	0.8	19.3	1.0

Table 3.3 shows the averages and standard deviations from 100 simulation runs for each aircraft type and capacity. Only aircraft types with at least 30 observations in the data were used since a smaller sample may not result in proper approximation of reality. Additionally, the relationship between throughput and the number and capacity is nonlinear, meaning that the throughput of two lanes is not exactly double that of a single lane, and so on. That nonlinear relationship is a result of drawing random deicing times from distributions for each aircraft in the simulation. The throughput rates presented can be useful to inform future work in airport planning that involves deicing facilities.

3.7 Conclusions

We developed an algorithm that uses ASDE-X data to detect usage of deicing facilities at airports. Our algorithm analyzes speed profiles of aircraft while using the deicing facility to extract deicing, waiting, and taxiing times. Our analysis of 1,504 departures at Chicago O'Hare (ORD) showed that 15.5% of all departures used the centralized deicing facility. There were 19 unique aircraft types in our deicing data, we presented statistics of deicing times per aircraft type

including mean deicing times and deicing rates in aircraft/hour. Our work fills a gap in the currently available data by providing deicing time estimates that can be useful for other researchers in the area of airport planning and simulation.

This study also presented a discrete-event simulation model that uses the extracted deicing times to simulate the usage of a centralized deicing facility with multiple deicing pads. We ran the simulation model to explore the impacts of changing the capacity utilization strategy at ORD. Our results show that the amount of time spent inside the deicing system can be reduced the most if the airport implements a policy based on our FCFSi strategy. The improvement was a result of minimizing the time periods that deicing pads were underutilized by assigning aircraft to available capacity regardless of the operating airline.

In the baseline scenario, aircraft spent 33.9 minutes on average. The simulation results showed that implementing our FCFSi strategy saved 8.5 minutes per aircraft on average compared to the baseline, this is equivalent to 25.1% of aircraft minutes and 26.8% of seat minutes. Our FCFSc strategy saved 6.0 minutes per aircraft on average, which is 17.8% of aircraft minutes and 20.3% of seat minutes. The FCFSc improvement to the baseline was less compared to the FCFSi, the latter did not leave a lot of inefficiency in queueing and capacity utilization due to having independent queues per deicing pad. Finally, assigning aircraft to deicing pads based on their deicing speed (slow, medium, fast) showed similar results to the baseline. This was due to certain airlines dominating specific aircraft types in the data, both scenarios showed similar queueing behavior with a small decrease in efficiency in our deicing speed scenario. The deicing speed scenario increased aircraft minutes in the system by 1.3% and aircraft seat minutes by 1.5% on average compared to the baseline.

3.8 Limitations and Future Work

A limitation to our study is the sample size of deicing aircraft. Having more data can help in increasing the number of aircraft covered by the study, it can also improve the distributions that were used in the simulation for more reliable results. We also recommend including weather conditions at the airport for future work. Weather conditions such as temperature and precipitation play a role in the decision to require deicing, anti-icing, or both. We expect that including weather conditions can help in understanding how they impact the amount of time aircraft spend in deicing pads. Finally, future work may include scenarios of changes in fleet mix at the airport.

3.9 Acknowledgment

This work was funded by the Federal Aviation Administration's NEXTOR III grant.

References

1. Airport Surface Detection Equipment, Model X (ASDE-X) https://www.faa.gov/air_traffic/technology/asde-x/ accessed Oct 18, 2021
2. Thomas SK, Cassoni RP, MacArthur CD. Aircraft anti-icing and de-icing techniques and modeling. Journal of aircraft. 1996 Sep;33(5):841-54.
3. Mao X, ter Mors A, Roos N, Witteveen C. Agent-based scheduling for aircraft deicing. In Proceedings of the 18th Belgium-Netherlands Conference on Artificial Intelligence 2006 Oct 5 (pp. 229-236). BNVKI.
4. The Operations Network (OPSNET) data. https://aspm.faa.gov/opsnet/sys/main.asp Accessed Feb 24, 2022
5. Airlines for America, U.S. Passenger Carrier Delay Costs https://www.airlines.org/dataset/u-s-passenger-carrier-delay-costs/ Accessed Feb 18, 2022
6. FAA, AC 150/5300-14D - Design of Aircraft Deicing Facilities, https://www.faa.gov/airports/resources/advisory_circulars/index.cfm/go/document.current/documentNumber/150_5300-14 Accessed Oct 5, 2021
7. FAA Holdover Time Guidelines Winter 2021-2022, https://www.faa.gov/other_visit/aviation_industry/airline_operators/airline_safety/media/FAA_2021-22_HoldoverTables.pdf Accessed Oct 12, 2021
8. Norin A, Yuan D, Granberg TA, Värbrand P. Scheduling de-icing vehicles within airport logistics: a heuristic algorithm and performance evaluation. Journal of the Operational Research Society. 2012 Aug;63(8):1116-25.
9. Zhiwei X, Yi L. Research of algorithms for aircraft ground deicing operation scheduling model. In 2010 8th World Congress on Intelligent Control and Automation 2010 Jul 7 (pp. 5973-5977). IEEE.
10. Xing Z, Li J. Research of game-theoretic approach for aircraft ground deicing operation scheduling. In2011 Second International Conference on Mechanic Automation and Control Engineering 2011 Jul 15 (pp. 570-574). IEEE.
11. Xing Z, Lian G. Cooperative game theoretical research for aircraft deicing operation scheduling. InProceedings of the 10th World Congress on Intelligent Control and Automation 2012 Jul 6 (pp. 2407-2411). IEEE.
12. Wu ML, Li LY, Chen B. Optimization research of the aircraft deicing problem. In Advanced Materials Research 2013 (Vol. 616, pp. 1926-1929). Trans Tech Publications Ltd.
13. Zhu Z, Jung Y, Eun Y, Hosagrahara V, Lee H, Jeon D. Modeling Deicing Operations in Departure Scheduling using Fast Time Simulation. In 2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC) 2019 Sep 8 (pp. 1-10). IEEE.
14. SeatGuru Airline Seat Maps https://www.seatguru.com/ accessed Oct 22, 2021
15. Ross SM. Introductory statistics. Academic Press; 2017.

CHAPTER 4

Airport scheduling and operational performance: A clustering analysis of airport response to COVID-19

4.1 Introduction

Passenger air travel has been steadily growing in the past decade, both globally and in the United States (1 ,2 , 3). Throughput at airports has been increasing to levels that are getting closer to the declared capacity of airports. During peak hours, some airports are regularly operating close to their capacity, some of these airports are so busy to the point where the FAA need to implement slot control or schedule monitoring to prevent schedules of exceeding capacities (4). There have been trends to help alleviate stress on airports such as the increasing average seats per flight in domestic US flights, this helps the National Airspace System (NAS) accommodate more passengers using the same runway capacity available at airports. Also, the FAA implemented strategies to manage demand-capacity imbalances when capacity is reduced due to various reasons. For example, weather events cause airports to reduce capacity which results in time periods where demand is exceeding capacity, here is where the FAA implements strategies to manage the imbalances until the system recovers from the associated delays that occur because of that.

In early 2020, the Coronavirus disease 2019 (COVID-19) pandemic started and forced air travel demand to decrease sharply in most parts of the world due to travel restrictions that were put in place to limit the spread of the virus. The pandemic also impacted capacity due to reasons such as workforce social distancing, days when ATC facilities were shut down due to COVID cases, and financial challenges due to the decreased demand. Although the COVID-19 pandemic impacted both demand and capacity, the reduced demand created a new challenge in the system since capacity exceeded demand by very large margins in the NAS. This scenario is unique because the delays in the system did not fall to zero despite the sharp drop in demand.

The purpose of this study is to evaluate 77 United States (US) airports to compare their responses to the COVID-19 pandemic in terms of capacity, throughput, and the resulting operational performance. We evaluate the response of airports to the initial shock event during 2020 in addition to the recovery period that followed in 2021. The motivation for this study stems from the ability for this shock event to measure the degree to which airport operational performance is driven by the airport versus the airlines operating at the airport. For example, prior to COVID-19, congested airports would have peak times with increased number of scheduled operations. This increased demand on a set amount of resources (capacity) typically results in delays, impacting the airport's Key Performance Indicators (KPIs) for on-time performance and for capacity and throughput efficiency. However, COVID-19's reduced demand can show if airlines continue to peak their schedules resulting in delay even during a low-demand time period. Additionally, the recovery period provides insight on the resilience of the NAS from an operational performance standpoint when managing the increase of demand that followed the shock event caused by the COVID-19 pandemic.

4.2 Literature Review

Throughput and capacity have been extensively studied in literature because balancing the two variables is an important goal of air traffic management. Throughput at airports is limited

by the available capabilities in the system which include existing infrastructure and operational capacity. From a performance standpoint, airports set their objectives to maximize throughput and minimize delays without exceeding capacity. The estimation of airport capacity is a topic of interest since it is important for performance measurement. Di Mascio et al. compare different methods of capacity estimation that range from using tables and charts to sophisticated simulations (5). The use of simulation to estimate airport capacity is demonstrated in a study by Bubalo and Daduna (6). They used the modeling tool SIMMOD to simulate different scenarios at Berlin-Brandenburg International airport to evaluate practical capacity under realistic demand and possible capacity gains.

Over the past few decades, there have been several studies that proposed models to optimize the available capacity at airports. For example, Gilbo discusses the estimation and representation of airport capacity and presents an optimization model to help satisfy the demand and mitigate congestion during busy time periods (7). The objective of that optimization is to minimize the total delay in arrivals and departures by varying the capacity of each operation. Gilbo's model is based on the notion that arrivals and departures are interdependent since both types of operations share the available resources at any airport to varying degrees. Zografos et al. discuss how slot scheduling can be used as a technique to increase the utilization of airport capacity (8).

Jacquillat and Odoni discuss the dynamic relationship between airport capacity, demand, and on-time performance (9). They bring insights to the non-linear relationship between flight schedules and delays, this is why peak throughput is one of the key performance indicators and often compared with airport capacity to assess how stressed an airport can be during times of high demand. Odoni et al. compare between Newark International (EWR) and Frankfurt International (FRA) to demonstrate how differences in scheduling practices can lead to dramatic impacts on delays (10). In the case of EWR, peaking of demand during certain hours of the day lead to operating the airport at or above its declared capacity which resulted in significant increases in delays. In addition, research using simulation models showed that there are multiple benefits to capacity optimization, while it is aimed at mitigating delays, it can help in reducing fuel consumption and emissions that harm the environment (6, 11).

COVID-19 impacted the operational performance of the NAS significantly, the number of departures in the United States dropped by over 70% in May 2020 compared to May 2019 (12). An analysis by Monmousseau et al. shows that although demand dropped sharply, the portion of passengers who faced interruptions still suffered from high delays (13). Their study analyzed data on how airlines reacted differently to the COVID-19 challenges. Research by Guo et al. provided a qualitative assessment of airport network resilience during the COVID-19 pandemic (14). Their research assessed resilience in the aviation system from a network perspective in China and Europe. They included networks of 232 airports in China and 82 airports in Europe, their conclusions show that the two networks recovered differently due to implementing different policies in each region. While their assessment is qualitative, it shows the importance of policy and airport characteristics in the ability to recover when responding to disruptive events such as natural disasters, technological failures, or pandemics. In our research we investigate how airports in the US reacted to the COVID-19 pandemic, we identify the significant factors driving their behavior including airlines variables and other factors such as airport size and multi-airport city status.

4.3 Data

In this analysis we study the 77 airports of the Aviation System Performance Metrics (ASPM) database (15). These airports include the core 30 airports in the U.S. To analyze the responses of airports to the COVID-19 shock event, we use three sets of airport metrics. Pre-COVID metrics span May 1, 2019-September 30, 2019 and the two post-COVID metrics include the same dates, May through September, of 2020 and 2021 to account for "shock" and "recovery", respectively. Throughout our study, pre-COVID refers to the time period prior to COVID-19 and the term post-COVID refers to time periods affected by COVID-19. The same months of each year were used to account for any seasonality changes in schedules, where certain airports have high demand for certain months of the years. These months were chosen with respect to regulations in the Coronavirus Aid, Relief, and Economic Security (CARES) Act (16) and the Coronavirus Response and Relief Supplemental Appropriation (CRRSA) Act (17). The study time periods are shown in Figure 4.1 along with the evolution of air traffic at the ASPM 77 airports. The data showed that the number of arrivals declined by 67% during April 2020, the lowest point in the curve, compared to April 2019.

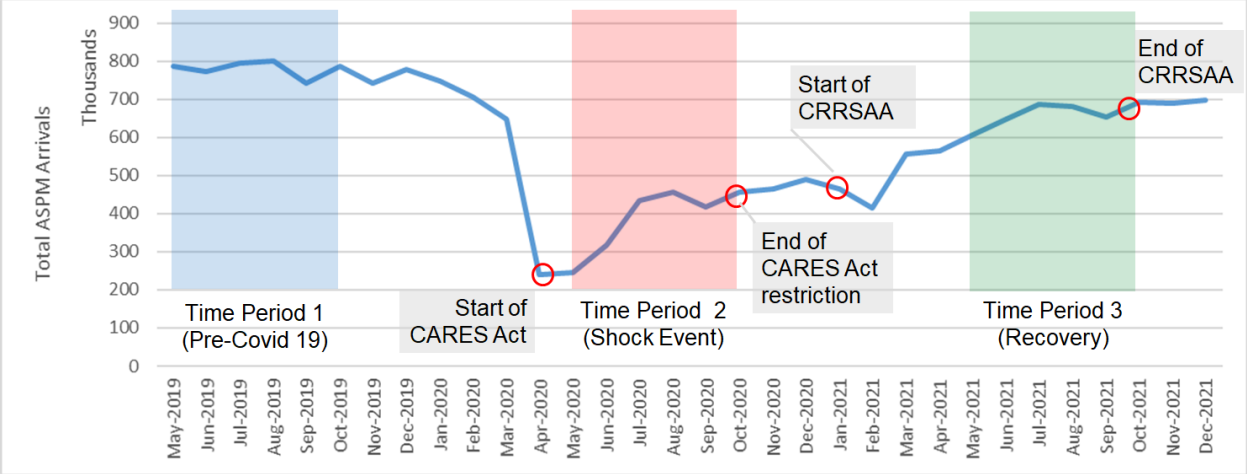

Figure 4.1 Evolution of air traffic over the study timeline (ASPM 77 airports)

The CARES Act provided stimulus funds to airlines in exchange for following imposed minimum service requirements along with restrictions of staffing or pay cuts. It was signed into law on March 27, 2020 and airline restrictions ended on September 30, 2020 (18). The choice of our study months gave a one-month buffer before the beginning of government financial assistance as it took time for airlines to stabilize their schedules and the logistics to follow the service regulations. Minimum Service Obligations (MSOs) required air carriers that accepted financial assistance to maintain minimum levels of scheduled air transportation service to points served by that air carrier before March 1, 2020, with some exceptions. Under the CARES Act, MSO levels depended on the size of the carrier and the level of service at those points in the air carrier's network. Carriers were allowed to consolidate service at a single airport in multi-airport metropolitan areas.

The CRRSA Act was signed into law on December 27, 2020 (19) providing extensions to the payroll assistance for air carriers. The CRRSA Act re-implemented the MSOs from the CARES Act from January 15, 2021 through March 31, 2021, and then a final order was issued on April 29, 2021 to extend the financial assistance regulations through September 30, 2021(20).

Since air travel demand was recovering during 2021, airports covered under the CRRSA Act were able to meet and exceed the required MSOs. Specifically, the U.S. Department of Transportation recognized that during the effective time period of CRRSA Act "virtually all Covered Points continued to receive service in excess of minimum levels required by the [CRRSAA]" (21).

Time period 1 in our study timeline was used as a reference for airport characteristics for pre-COVID-19 conditions. Time period 2 was meant to represent the reaction period to the shock event that the aviation system went through. During this time period, travel restrictions were still in place and operations were influenced by the CARES Act requirements. Time period 3 was meant to represent the system's reaction while travel restrictions were loosened and the CRRSA Act was in effect.

4.3.1 Datasets

The first set of variables provide the *"Pre-COVID-19 Airport Characteristics"* that could impact their response and outcomes to the shock event. This includes information about carrier service (i.e. major, low-cost carrier, cargo), multi-airport city status, airport hub size (i.e. small, medium, large (24)), and the ratio of arriving passengers that use the airport for a connection flight to those who use it as a final destination. In our data, flights by Delta, United, and American airlines accounted for major carriers. Low Cost Carrier (LCC) flights were flights by Southwest, Spirit, JetBlue, and Frontier. Cargo flights were flights by UPS and FedEx.

The second set indicates *"Airport Metrics Changes"*. The operational metrics include, peak capacity, peak throughput, and total operations count. Throughput is the number of actual operations that took place at the airport. Capacity is represented by the number of arrivals and departures that the airport can accept during a time period, this is also known as the called rate. Called rates are usually estimated values reported by the air traffic control (ATC) personnel managing air traffic at the airport based on weather conditions, separation requirements, and active runway configuration (23).

The peak throughput and capacity metrics were estimated consistent with the definitions set by the ICAO GANP KPIs (22). Both peak throughput and capacity were measured as the 95^{th} percentile of data for the study time period using the ASPM data reported at 15-minute intervals. The hourly peak throughput was calculated on a 15-minute rolling hour basis, which means that the one-hour time frame moved at 15-minute intervals instead of every one hour on the clock. Rolling is recommended since demand does not necessarily align with the top of the hour, for example, a busy "hour" at an airport could refer to the time frame of 7:30 - 8:30 AM. On the other hand, hourly peak capacity was calculated without rolling since capacity has less variability compared to throughput.

Under the CARES Act, airlines were required to continue a minimum level of service to "points" or cities in their network. If an airline served two airports in a multi-airport city prior to COVID-19, it only had to continue serving one of these airports during the regulation period. For the purposes of this study, an airport was given a multi-airport city indicator of 1 if there was another large airport in its city, not counting itself. While part of the CARES Act service requirements were based on multi-airport definitions, the data shows that there are other airport characteristics that had significant relationships to the changes in airport metrics (capacity, throughput, and total operations). Figure 4.2 shows the correlation between pre-COVID-19

airport characteristics and change in airport metrics during the shock event while Figure 4.3 shows the correlation with the change in airport metrics during the recovery.

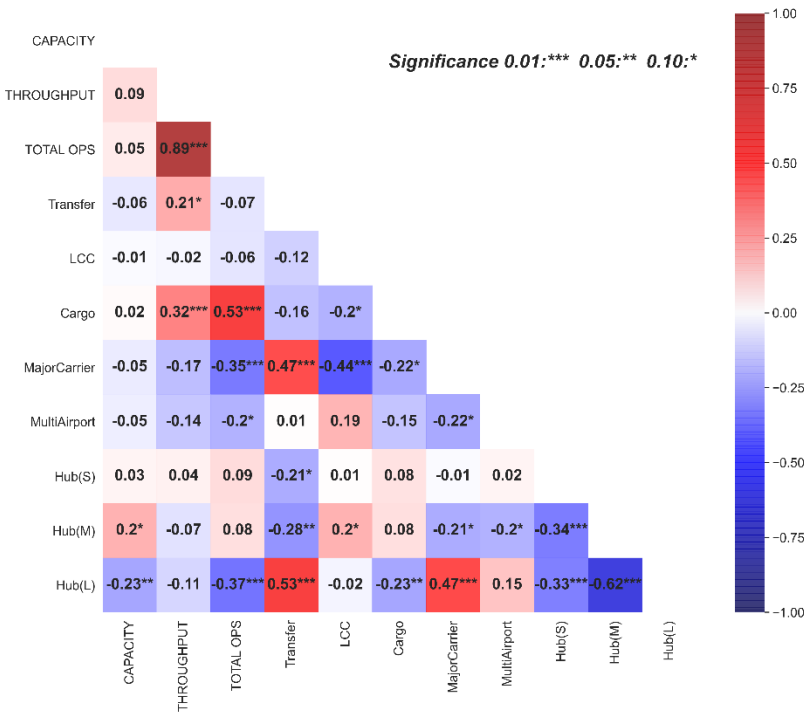

Figure 4.2 Correlation matrix of airport characteristics and metrics changes (Time period 2)

The data showed that the percentage of cargo operations at an airport was positively correlated to the change in total operations. As shown in Figure 4.2, airports with higher cargo percentage were able to retain more operations during the shock event. Transfer rate was positively correlated to peak throughput, meaning that a higher percentage of transferring passengers could be contributing to airlines peaking their schedules. It is more efficient for airlines to coordinate the arrival of connecting passengers coming from different origins within a short period of time in order to board the same outbound flight. The multi-airport indicator was negatively correlated to the change in total operations as expected due to the minimum service requirement set by the CARES Act.

The data showed that the percentage of major carriers operations was negatively correlated to the change in total operations. Similarly, the large hub airport indicator was negatively correlated to the total operations, this was expected since there is a large percentage of major carriers operations in large hubs. The percentage of Low Cost Carrier (LCC) operations did not have a significant correlation with changes in throughput or total operations. This is most likely due to being correlated with both the major carrier and cargo variables.

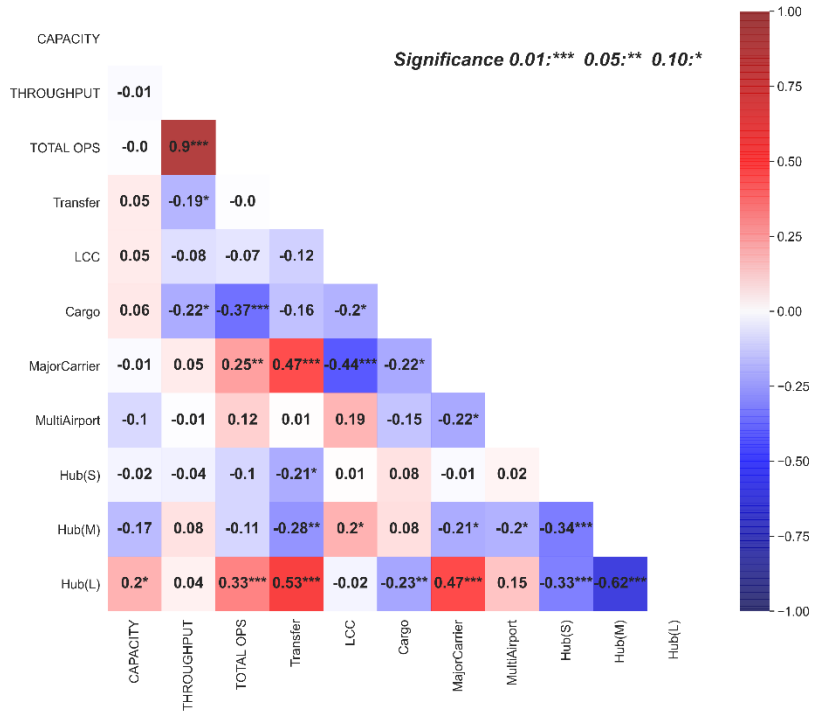

Figure 4.3 Correlation matrix of airport characteristics and metrics changes (Time period 3)

As shown in Figure 4.3, cargo operations were negatively correlated with total operations and peak throughput during recovery. This could be because cargo was the least impacted segment during the shock event, therefore, it was not expected to have a positive impact on recovering operations that were mostly passenger traffic. Major carrier operations were positively correlated to total operations recovery. Similarly, the large hub indicator was positively correlated to both total operations and throughput peak. Lastly, the multi-airport indicator was not significant during the recovery although it was significant under the CARES Act. This trend indicates that airlines scheduling was the major factor in recovery trends since airports covered under the CRRSA Act were able to meet and exceed the minimum service requirements anyway due to demand improvement in 2021, as previously stated.

The third set of metrics measure the *"Operational Performance"* at airports, particularly related to airport delays. We calculated arrivals delay as the difference between the scheduled gate-in time and the actual gate-in time reported in ASPM data. Additionally, we calculated punctuality as the percent of flights that arrived "on-time". A flight is considered "on-time" as long as the delay is less than 15 minutes, consistent with the FAA definition (2). In this analysis, we used punctuality, delay per delayed flight, and delay per flight as metrics of operational performance. The goal of these metrics is to provide insight into how much airlines drive airport operational performance metrics verses the airports themselves. In other words, how much of the delay is due to airport issues versus airline schedule peaking.

4.4 Methodology

A K-means clustering methodology is followed to group airports on the second set of metrics *"Airport Metrics Changes"* to see which airports responded similarly to the shock event and

recovery. These K-means clusters are then joined with the first set of metrics to see what pre-COVID-19 characteristics of airports influence airport response and increase/decrease operational resilience to shock events. The clusters are also compared to the third set, the delay and punctuality data, to provide insight on the ability to increase traffic while managing the increase in delays.

The clustering analysis was carried out twice for time periods shown in Figure 4.1, the shock event comparing time period 2 to pre-COVID, the recovery event comparing time period 3 to time period 2. The elbow method was used to determine the optimal number of clusters to partition the data in the K-means algorithm. As shown in Figure 4.4, the sum of squared errors (SSE) indicated that a seven-cluster solution is appropriate since the reduction in SSE starts to diminish after seven clusters.

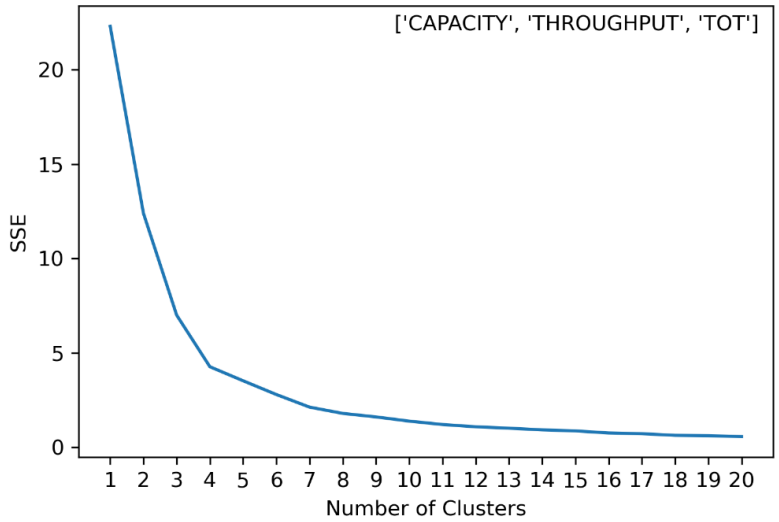

Figure 4.4 Number of Clusters vs SSE

4.5 Airport Metrics Data Summary

While this study includes the 77 ASPM airports, this section will focus on presenting the Main 34 airports of 2019, which are the airports with the highest number of operations (arrivals plus departures). This section presents a summary of the three variables that were used to group airports in our clustering analysis. Subsection 4.5.1 presents how the total operations evolved, which is a metric of demand. Subsection 4.5.2 presents peak throughput which is a measure of how peaky the schedule at the airports was during the busiest hours. Finally, Subsection 4.5.3 shows peak capacity as a measure of the highest declared capacity at the airport.

4.5.1 Total Operations

Figure 4.5 shows the change in number of arrivals plus departures (i.e. total operations) for three time periods covered in this study. While all airports lost operations in 2020, the impact varied greatly across the airports, ranging from a reduction of 14.8% at MEM to 81.5% at LGA. Some airports were more resilient than others in retaining flight demand, in particular, Anchorage International Airport (ANC) and Memphis International Airport (MEM), two major cargo airports in the US, each saw less than a 16% reduction in flights. The airports that lost the

most percent operations were large hub and international airports such as Hartsfield-Jackson Atlanta (ATL), Boston Logan (BOS), Ronald Reagan Washington National (DCA), Newark Liberty (EWR), New York John F. Kennedy (JFK), and New York LaGuardia (LGA).

While air travel demand was recovering in time period 3, some airports recovered more than others. LGA and JFK had the highest recovery rates of 160% and 162%, respectively, after being among the airports that were impacted the most during the shock event. Interestingly, there was a group of airports that reached a total number of operations in the recovery period higher than pre-COVID-19 numbers. These airports include ANC, Phoenix Sky Harbor International airport (PHX), and Salt Lake City International airport (SLC).

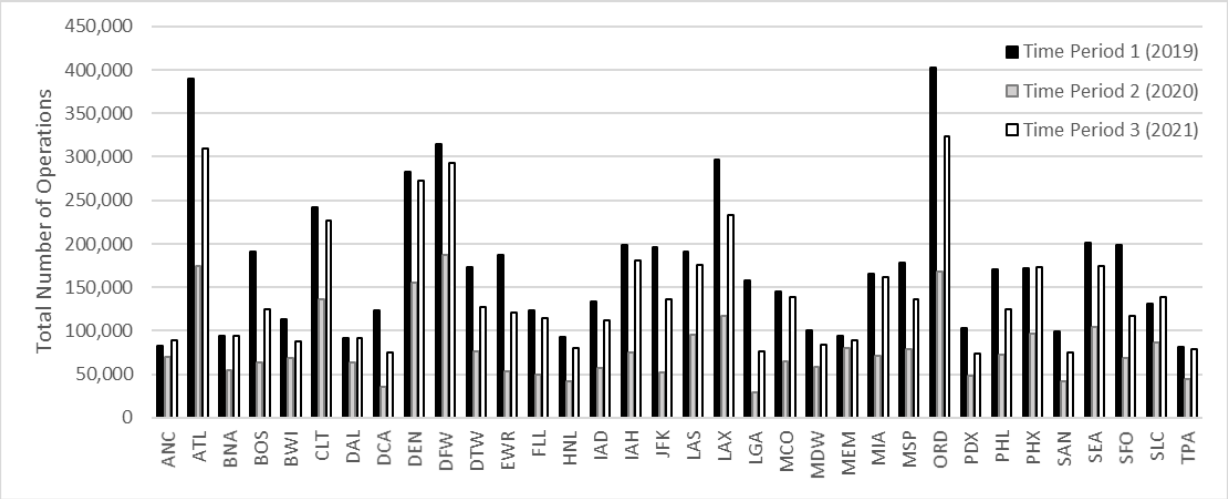

Figure 4.5 Change in total operations (Main 34 airports)

4.5.2 Peak Throughput

Peak throughput for arrivals and departures followed similar patterns. Therefore, total peak throughput (departures + arrivals) was used in this study. Figure 4.6 shows total peak throughput at the Main 34 airports for the study time periods. During the shock event period, airports with minimal changes to total operations, ANC and MEM, did not change their schedule peaks as much as other airports. Similarly, some of the airports with large reduction in operations had a proportional reduction in their peak throughput, examples include LGA, JFK, and ATL. On the other hand, some airports with significant reductions in their total operations did not see a similar drop in peak throughput. For example, the peak throughput operations at Charlotte Douglas International (CLT) was reduced by 10% during the shock event although the total number of operations dropped by 43% in the same time period. The different trends in reaction to the shock event are mostly driven by airlines operating at the airport and how they schedule flights to support their hub-and-spoke network and allow for passenger connections.

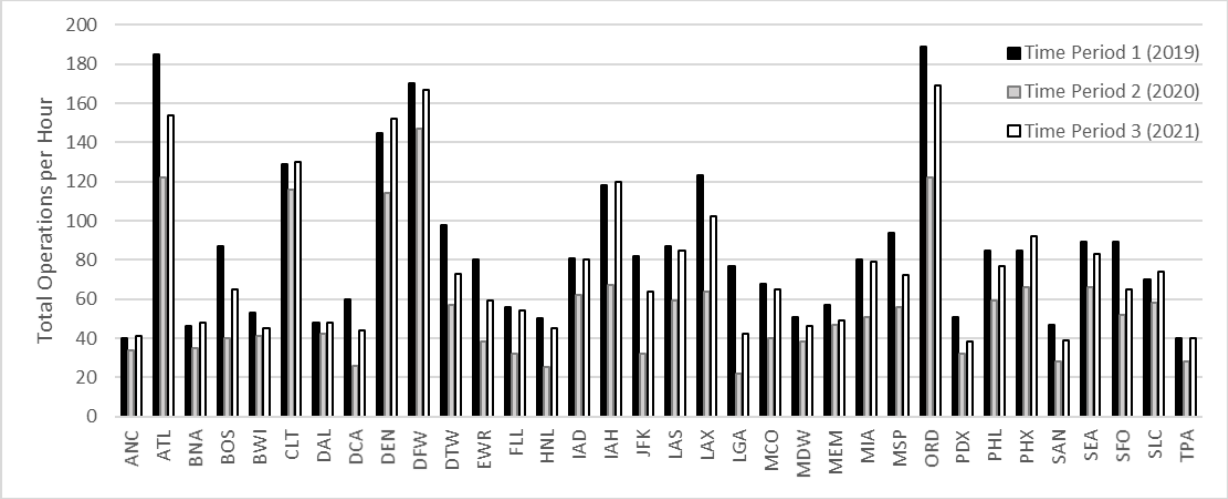

Figure 4.6 Change in total peak throughput (Main 34 airports)

The data showed that airports which recovered to higher then pre-COVID total operations also increased their peak throughput. These airports include ANC, SLC, and PHX as shown in Figure 4.6. However, there were airports like CLT, Denver International (DEN), and George Bush Houston Intercontinental (IAH) that peaked their schedules higher during the recovery period although their total operations did not recover proportionately. This can be linked to the demand characteristics at the airports which are impacted by factors such as the type of airlines (major or low cost carrier) and the passenger transfer rate. Section 4.6 discusses the common factors between airports that showed similar trends in change in their KPIs.

4.5.3 Peak Capacity

The changes in peak total capacity (arrivals + departures) at the Main 34 airports are shown in Figure 4.7. Peak capacity tends to have less variability than throughput since capacity is not driven by demand at the airport. Declared capacity at an airport is impacted by various factors including weather, air traffic control staffing, gate capacity, and runway capacity. Weather conditions impact the airport's capacity when aircraft separation rules change during bad weather for safety. The effect of weather conditions in our analysis wan not major because our data was for the same time periods across the three years which eliminated seasonality differences.

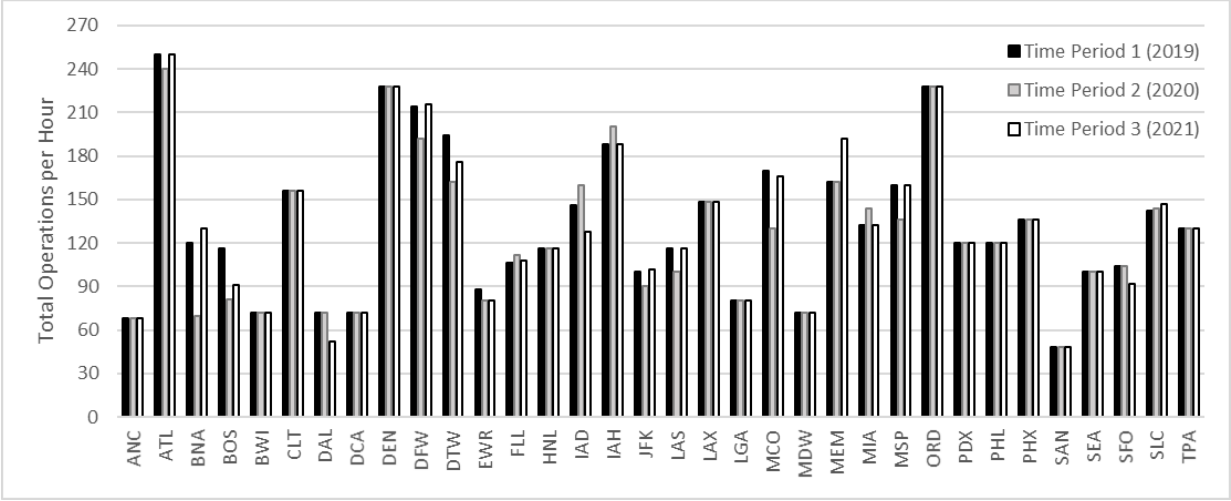

Figure 4.7 Change in peak total capacity (Main 34 airports)

Construction projects often have significant impacts on capacity, it can be a negative impact due to temporary closures during construction and positive when the construction is complete and new infrastructure becomes available. Some airports had ongoing construction projects during our study time periods that could explain significant changes in their peak capacity shown in Figure 4.7. For example, a new terminal is under construction at Orlando International Airport (MCO) which is expected to add 15 gates to the airport's capacity (25). Nashville International Airport (BNA) is having a major terminal expansion project that will add gate capacity to the airport, however, the project has been temporarily impacting the flow of passengers and capacity while construction is taking place (26). Additionally, various construction projects were taking place during our study time periods at MEM including apron area improvements and a consolidated deicing facility (27).

4.6 Results and Discussion

The K-means clustering algorithm was applied to the *Airport Metrics Changes* variables, namely the percent change in each of the throughput peak, capacity peak, and total operations for the ASPM 77 airports. Clustering was done twice to analyze the reaction to each of the shock event and the recovery time periods as described in Figure 4.1. The first clustering run was for the changes from time period 1 to time period 2. The second run was for changes from time period 2 to time period 3. Characteristics of the resulting groups are discussed in detail in Sections 4.6.1 and 4.6.2.

4.6.1 Shock Event Results (Time period 2 compared to time period 1)

The seven clusters were created based on the *Airport Metrics Changes* variables as percent change of time period 2 compared to time period 1. This section of the analysis looks into patterns and factors that influenced reactions by airports to the COVID-19 shock event. The resulting groups and their average airport characteristics are shown in Table 4.1, the highest number in each column is highlighted. While the overall decline in operations was 52% during the COVID-19 shock time period, it can be seen that airports had different responses when comparing across the groups. For example, Clusters 1 and 4 had similar reductions in the number of operations (-56% and -57%), however Cluster 4 airports reported an average 18%

reduction in peak capacity. Cluster 1 was the largest cluster with 25 airports, they had the highest percentage of major carriers and were mostly medium and large hubs.

Cluster 3 had 24 airports, mainly single-airport cities with the highest average transfer rate. These airports tended to keep their schedules peaked which was indicated by the relatively smaller reduction in peak throughput compared to the reduction in total operations. This trend might be due to airlines having to coordinate flight schedules to accommodate transferring passengers which create busy time periods, i.e. peaks, to make the transfer process more efficient. Cluster 0 had the most resilient flight demand of all clusters, it was made up of airports that on average had a higher cargo or general aviation flight share. The cargo segment was the least impacted by the COVID-19 pandemic since demand on shipping increased due to most people working from home and businesses shifting to online shopping for safety purposes.

Table 4.1 Airport Characteristics of Clusters (Shock Event)

	Cluster Size	Average Airport Metrics Changes			Average Airport Characteristics						
		Capacity	Throughput	Total Ops	Transfer Rate	LCC	Cargo	Major Carriers	Multi-Airport	Hub (M)	Hub (L)
0	11	0%	-13%	-19%	1.06	17%	19%	18%	36%	55%	0%
1	25	3%	-39%	-56%	1.21	25%	2%	48%	48%	36%	44%
2	1	93%	-31%	-50%	1.05	50%	1%	29%	0%	100%	0%
3	24	1%	-26%	-41%	1.32	30%	3%	43%	33%	33%	38%
4	8	-18%	-43%	-57%	1.22	35%	2%	46%	25%	38%	63%
5	2	-36%	-30%	-44%	1.11	39%	3%	40%	0%	50%	50%
6	6	-3%	-61%	-72%	1.10	12%	1%	37%	67%	33%	50%

There were two clusters that had a small sample size in each. Cluster 2 included only Sacramento International Airport (SMF) which had a large increase in capacity. The capacity at SMF was originally reduced during time period 1 due a renovation project of runway17R/35L from April to October 2019. This means that the 93% increase shown in Table 4.1 is capacity gained back after the runway was reopened. Another small cluster was Cluster 5 which contained only two airports, Nashville International Airport (BNA) and San Antonio International Airport (SAT). These two airports had a capacity peak reduction of 36% on average, the largest drop in capacity among all groups. This was caused by major terminal construction works at BNA (26) and various taxiway and apron projects at SAT (28).

Table 4.2 shows the operational performance data along with the same clusters from Table 4.1. Cluster 0 did not see the operational improvements that other clusters experienced, given the smaller decrease in total and peak operations. This was expected since this cluster was the highest in cargo operations and cargo demand was increasing during the COVID-19 pandemic. Cluster 0 was the only group of airports that had a decrease in punctuality on average, however it saw minor improvement in flight delays.

Table 4.2 Operational Performance of Clusters (Shock Event)

	Cluster Size	Average Airport Metrics Changes			Average Airport Characteristics		
		Capacity	Throughput	Total Ops	Punctuality	Delay/delayed flight	Delay/flight
0	11	0%	-13%	-19%	-1%	-9%	-7%
1	25	3%	-39%	-56%	11%	-22%	-47%
2	1	93%	-31%	-50%	4%	16%	-18%
3	24	1%	-26%	-41%	8%	-25%	-41%
4	8	-18%	-43%	-57%	13%	-33%	-58%
5	2	-36%	-30%	-44%	10%	-32%	-48%
6	6	-3%	-61%	-72%	24%	-31%	-63%

The largest improvement in punctuality and delay per delayed flight were found in Cluster 6, 24% and 63%, respectively. This group of airports had the most reductions in their total operations as well as peak throughput with minor changes in peak capacity. Cluster 6 included airports from multi-airport cities such as JFK, EWR, and LGA which could be related to the reduced stress levels at these airports compared to cargo or single airport cities with high transfer rates. Cluster 3 improved their punctuality by 8% on average, which is small compared to the 41% average reduction in total operations. This could be related to the high transfer rates in this group that can cause schedules to peak during busy hours. The data showed that the average improvement in punctuality was 9.5% for all airports which is considered a small improvement when compared to the percent decrease in total operations shown in Table 4.2.

4.6.2 Recovery Results (Time Period 3 compared to time period 2)

The seven clusters were created based on the Airport Metrics Changes variables as percent change of time period 3 compared to time period 2 to analyze patterns of recovery among airports. The resulting groups and their average airport characteristics are shown in Table 4.3, the highest number in each column is highlighted.

Table 4.3 Airport Characteristics of Clusters (Recovery)

	Cluster Size	Average Airport Metrics Changes			Average Airport Characteristics						
		Capacity	Throughput	Total Ops	Transfer Rate	LCC	Cargo	Major Carriers	Multi-Airport	Hub (M)	Hub (L)
0	18	-1%	21%	44%	1.34	32%	6%	34%	44%	56%	22%
1	5	1%	89%	145%	1.20	22%	1%	44%	80%	0%	80%
2	11	0%	64%	106%	1.08	21%	2%	43%	55%	27%	55%
3	1	0%	267%	257%	1.00	7%	1%	29%	0%	100%	0%
4	36	-1%	39%	72%	1.22	28%	3%	47%	28%	42%	39%
5	5	4%	-4%	7%	1.01	8%	28%	12%	40%	20%	0%
6	1	86%	37%	71%	1.18	44%	1%	39%	0%	0%	100%

During the recovery time period there was more variation in the changes in airport metrics compared to the shock event time period. Air travel demand recovered during this time period after some travel related COVID-19 restrictions were loosened. The data shows that the total operations at the ASPM 77 airports increased by 75% during the recovery time period. Table 4.3 shows that total operations recovery ranged from 7% to 257% among clusters.

The largest group of airports with similar recovery metrics formed Cluster 4 which contained 36 out of the 77 airports. Cluster 4 was made up of medium and large hubs with the highest percentage of operations by major carriers. This group had a 72% increase in their total operations while managing to increase their peak throughput by only 39% on average. The second largest group was airports with high transfer rates and LCC operations, they were in Cluster 0 which increased total operations by 44% while keeping throughput peak at 21% increase on average. In contrast, Cluster 5 contained cargo airports that had only minor changes in airport metrics. That was expected since cargo flights were the least impacted segment of commercial air transportation during the shock event. The airport with the least changes in both directions was MEM, the largest cargo airport in our study, operations were reduced by 14.8% in the shock and increased by 12% in recovery.

There were two airports with unique changes during the recovery time period that put each of them in their own single-airport cluster. Kahului airport in Hawaii (OGG) was in Cluster 3, there was a 257% increase in demand during the recovery time period. OGG is located on the island of Maui, a tourist destination, which put it in a unique position to receive high demand after the loosening of COVID-19 travel restrictions. Cluster 6 contained BNA which was unique in its 86% increase in peak capacity. This increase in capacity can be attributed to construction projects at the airport. With the exception of capacity, BNA was similar in other metrics to Cluster 4, the largest group of the recovery time period.

Table 4.4 shows the operational performance data along with the same clusters from Table 4.3. Delay increased across all groups ranging from 12% to 195%, punctuality decreased in all groups except Cluster 5 which was the cargo airports. One observation from Table 4.4 is that the increase in delays per flight was disproportionate to the increase in total operations. This trend is a major difference between the shock and the recovery time periods. It can be an indicator to the resiliency of the aviation system and its ability to manage delays in a unique environment of increased air traffic after disruption in demand. The changes in punctuality were smaller compared to changes in delays per flight, this gives insight that the majority of delays causing the high percentages in delays per flight were due to delays under 15 minutes keeping those flights “on time” according to the FAA definition.

Table 4.4 Operational Performance of Clusters (Recovery)

	Cluster Size	Average Airport Metrics Changes			Average Airport Characteristics		
		Capacity	Throughput	Total Ops	Punctuality	Delay/delayed flight	Delay/flight
0	18	-1%	21%	44%	-12%	27%	96%
1	5	1%	89%	145%	-18%	38%	157%
2	11	0%	64%	106%	-10%	29%	119%
3	1	0%	267%	257%	-11%	26%	195%
4	36	-1%	39%	72%	-11%	29%	101%
5	5	4%	-4%	7%	3%	5%	12%
6	1	86%	37%	71%	-17%	45%	169%

The highest increase in delay per flight was in Cluster 3, the single-airport cluster that contained OGG. The second highest increase in delay per flight was in Cluster 5, the other single-airport cluster that contained BNA. Cluster 1 had the most negative impacts on operational performance. This cluster consisted of large hub airports located in multi-airport cities, they had 18% decrease in punctuality and 38% increase in delays per delayed flight. Punctuality on average decreased by 10.7% for all airports which brought it back to levels similar to pre-COVID even though air traffic was not fully recovered.

4.6.3 Clustering Results Summary

This section presents a side-by-side summary of the clustering results from the previous two sections. Figure 4.8 shows a visualization of the flow of airports between clusters from the shock event to the recovery. Each cluster is labeled by its number and size, the thickness of each line is proportional to the number of airports in that flow between clusters. It can be seen that Cluster 4 of the recovery time period, the largest cluster, consolidated multiple airports from clusters 1,3, and 4 from the shock time period.

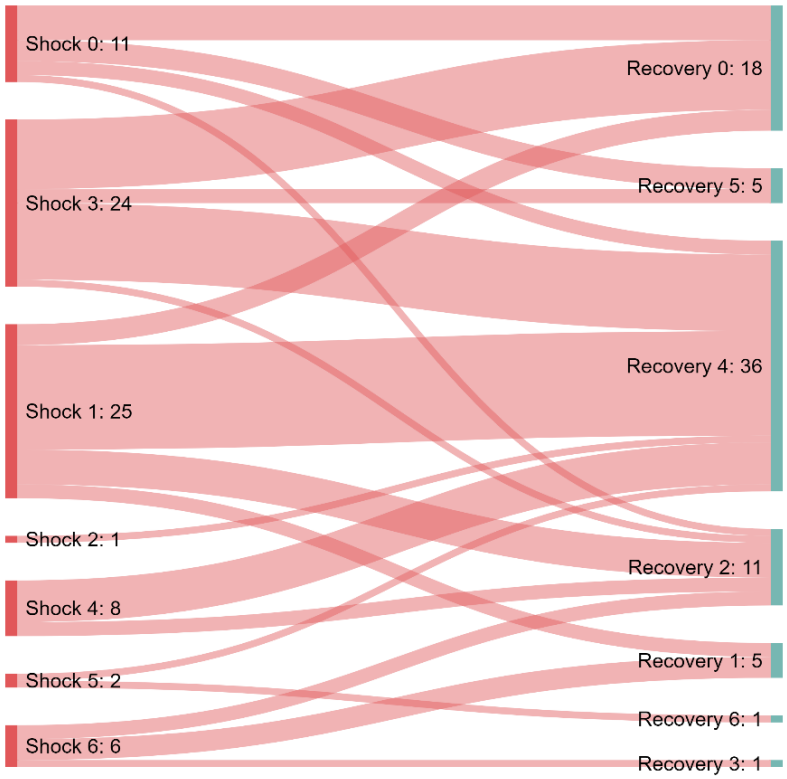

Figure 4.8 Flow of airports between clusters from shock to recovery

Figure 4.8 shows how Cluster 2 in the recovery was made up of airports from five clusters in the shock event, these airports increased their peak throughput causing them to cluster together away from their previous groups. For example, Los Angeles International (LAX) split from Cluster 1 to join Cluster 2 of the recovery. Although LAX clustered previously with other major airports in California such as San Francisco (SFO) and San Diego (SAN), it moved to Cluster 2 of the recovery which increased its throughput more aggressively compared to Cluster 4 where SFO and SAN clustered.

Table 4.5 shows all ASPM 77 airports and their clusters from the shock event time period and the recovery time period. The table is ordered by the recovery cluster numbers to give insight on how the groupings changed. Airports are listed as their three-character FAA location identifier, a list of the ASPM 77 airports and their full airport names are found on the ASPM website (29).

Table 4.5 Summary of clustering results

Airport	Time Period		Airport	Time Period		Airport	Time Period	
	Shock	Recovery		Shock	Recovery		Shock	Recovery
ANC	0	0	HPN	1	2	SMF	2	4
DAL	0	0	LAX	1	2	BHM	3	4
GYY	0	0	MIA	1	2	DEN	3	4
ONT	0	0	PSP	3	2	IND	3	4
SDF	0	0	BOS	4	2	ISP	3	4
OAK	1	0	MCO	4	2	OMA	3	4
PDX	1	0	DCA	6	2	PHX	3	4
PIT	1	0	EWR	6	2	SEA	3	4
ABQ	3	0	OGG	6	3	SJU	3	4
BWI	3	0	PBI	0	4	SLC	3	4
CLT	3	0	VNY	0	4	SNA	3	4
CVG	3	0	ATL	1	4	TPA	3	4
DFW	3	0	BUF	1	4	BDL	4	4
HOU	3	0	BUR	1	4	DTW	4	4
MDW	3	0	CLE	1	4	LAS	4	4
MKE	3	0	DAY	1	4	MSP	4	4
STL	3	0	IAD	1	4	MSY	4	4
TUS	3	0	JAX	1	4	SJC	4	4
FLL	1	1	LGB	1	4	SAT	5	4
IAH	1	1	MCI	1	4	MEM	0	5
JFK	6	1	ORD	1	4	OXR	0	5
LGA	6	1	PHL	1	4	RFD	0	5
TEB	6	1	PVD	1	4	MHT	3	5
RSW	0	2	RDU	1	4	SWF	3	5
AUS	1	2	SAN	1	4	BNA	5	6
HNL	1	2	SFO	1	4			

Table 4.5 provides details on how some airports in multi-airport areas reacted similarly to the shock event but recovered differently. For example, JFK, LGA, and EWR clustered together during the shock event in Cluster 6, however, JFK and LGA increased their peak throughput and operations more than the rest of the group during the recovery leading them to separate to a different cluster. On the other hand, there was a group of 15 airports that remained in the same group, i.e., reacted and recovered similarly. These airports were in Cluster 1 (shock) and stayed together in Cluster 4 (recovery), they include the busiest airports in the US such as ATL, Chicago O'Hare International (ORD), Philadelphia International (PHL), and San Francisco International (SFO).

Cargo airports exhibited interesting grouping, Cluster 0 contained four of the top cargo airports during the shock event, namely, ANC, MEM, Louisville Muhammad Ali International (SDF), and Chicago/Rockford International (RFD). During recovery, ANC and SDF went to Cluster 0 instead of Cluster 5, the cargo cluster that contained MEM and RDF. Both ANC and SDF increased their total operations during recovery to levels exceeding 2019 levels of the same time period. This significant increase made their recovery patterns more aligned with passenger traffic airports such as SFO and Chicago Midway (MDW) from a percent change perspective.

4.7 Conclusions

Our analysis of the ASPM 77 airports data showed that the COVID-19 pandemic resulted in a 67% decline in operations at the lowest point during the pandemic. The impact during the shock time period varied greatly across the airports, ranging from a reduction of 14.8% at MEM to 81.5% at LGA. During the recovery period, the increase in total number of operations ranged from 12% at MEM to 257% at OGG. The latter was an example of increasing demand as a result of loosening travel restrictions related to COVID-19.

We performed a clustering analysis to study airports' response to the COVID-19 pandemic, clustering was done for both the response to the shock event (May-September 2020) and the recovery afterwards (May-September 2021). The airport metrics variables that were used to cluster the airports were the change in peak capacity, peak throughput, and total number of operations. There was a number of airport characteristics that were correlated to the changes in airport metrics including cargo operations, operations by major carriers, multi-airport status, hub size, and the rate of transferring passengers. We found that the cargo airports were the most resilient in retaining air traffic demand which was a result of the shift to working from home and businesses moving towards online shopping during the pandemic. Additionally, the percentage of major carriers operations (Delta, United, and American) was significantly correlated to the change in total operations. The data showed that the rate of transferring passengers was a significant variable as it correlated to the change in peak throughput.

The data showed that being located in a multi-airport city was significantly correlated to the decrease in operations during the shock, however, it was not significant in the recovery trends. Although air carriers that received financial assistance were subject to MSOs, the main difference is that they were able to meet and exceed the MSOs during the recovery period due to normal demand improvement. This can be interpreted that being located in a multi-airport city was only significant when air transportation demand levels were lower than the required minimums, i.e. during the shock time period. Airports in multi-airport areas reacted similarly to the shock event but recovered differently. For example, New York airports JFK, LGA, and EWR clustered together during the shock event, however, LGA and JFK increased their peak throughput by 160% and 162% compared to 125% at EWR.

Our analysis showed that delays in the system did not change proportionately to the change in operations. Similarly, there were only minor improvements in punctuality. On-time flights at the ASPM 77 airports increased by 9.5% while operations declined by 52% during the shock event time period compared to pre-COVID. Part of this phenomenon was a result of schedule peaking which caused delays due to creating busy hours at the airports. Additionally, a

portion of the delays could be attributed to COVID-19 reasons such as increased safety measures at airports, increased passenger boarding times due to social distancing, and staffing shortages.

4.8 Future Work

Future work could include evaluating a shorter time span, possibly a month, to see if it returns different results. Current results may be capturing a trend for a small amount of time. Since the peak takes a 95th percentile, it would only take an airport to operate in a unique way for a few days during the study period to skew the results. Including other upper percentile alternatives could be beneficial.

Second, including additional airport characteristics that may be impacting the schedule changes and operational performance. One aspect of this analysis could be related to staffing at airports. Staffing was affected due to social distancing and safety, causing many airports to have capacity changes in the initial months of COVID-19.

References

1 https://www.airlines.org/dataset/impact-of-covid19-data-updates/
2 FAA/EUROCONTROL. Comparison of Air Traffic Management-Related Operational Performance: U.S./Europe. 2019.
3 https://www.airlines.org/dataset/world-airlines-traffic-and-capacity-2/
4 https://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/systemops/perf_analysis/slot_administration/
5 Di Mascio P, Cervelli D, Correra AC, Frasacco L, Luciano E, Moretti L, Nichele S. A critical comparison of airport capacity studies. Journal of Airport Management. 2020
6 Bubalo B, Daduna JR. Airport capacity and demand calculations by simulation—the case of Berlin-Brandenburg International Airport. NETNOMICS: Economic Research and Electronic Networking. 2011
7 Gilbo EP. Airport capacity: Representation, estimation, optimization. IEEE Transactions on control systems technology. 1993
8 Zografos KG, Madas MA, Androutsopoulos KN. Increasing airport capacity utilisation through optimum slot scheduling: review of current developments and identification of future needs. Journal of Scheduling. 2017
9 Jacquillat A, Odoni AR. A roadmap toward airport demand and capacity management. Transportation Research Part A: Policy and Practice. 2018 Aug 1;114:168-85.
10 Odoni A, Morisset T, Drotleff W, Zock A. Benchmarking airport airside performance: FRA vs. EWR. In9th USA/Europe Air Traffic Management R&D Seminar 2011.
11 Ignaccolo M. A simulation model for airport capacity and delay analysis. Transportation Planning and Technology. 2003 Apr 1;26(2):135-70.
12 Hotle S, Mumbower S. The impact of COVID-19 on domestic US air travel operations and commercial airport service. Transportation Research Interdisciplinary Perspectives. 2021 Mar 1;9:100277.
13 Monmousseau P, Marzuoli A, Feron E, Delahaye D. Impact of Covid-19 on passengers and airlines from passenger measurements: Managing customer satisfaction while putting the US Air Transportation System to sleep. Transportation Research Interdisciplinary Perspectives. 2020 Sep 1;7:100179.

14 Guo J, Li Y, Yang Z, Zhu X. Quantitative method for resilience assessment framework of airport network during COVID-19. Plos one. 2021 Dec 3;16(12):e0260940.
15 Aviation Systems Performance Metrics database (ASPM): https://aspm.faa.gov/aspmhelp/index/Aviation_Performance_Metrics_%28APM%29.html
16 FAA Airports 2020 CARES Act Grants: https://www.faa.gov/airports/cares_act/
17 FAA Airports 2021 CRRSA Act Grants: https://www.faa.gov/airports/crrsaa/
18 U.S. Department of Transportation, Continuation of Certain Air Service, Docket DOT-OST-2020-0037. https://www.regulations.gov/document/DOT-OST-2020-0037-0047
19 U.S. Department of Transportation, Continuation of Certain Air Service, Docket DOT-OST-2020-0037. https://www.regulations.gov/document/DOT-OST-2020-0037-0390
20 U.S. Department of Transportation, Continuation of Certain Air Service, Docket DOT-OST-2020-0037. https://www.regulations.gov/document/DOT-OST-2020-0037-0456
21 U.S. Department of Transportation, Continuation of Certain Air Service, Docket DOT-OST-2020-0037. https://www.regulations.gov/document/DOT-OST-2020-0037-0432
22 International Civil Aviation Organization (ICAO) Key Performance Indicator (KPI) Overview: https://www4.icao.int/ganpportal/ASBU/KPI
23 Gentry J, Duffy K, Swedish WJ. Airport capacity profiles. Washington DC. 2014 May.
24 Passenger Boarding (Enplanement) and All-Cargo Data for U.S. Airports https://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/passenger/
25 https://www.constructiondive.com/news/orlando-international-airport-high-tech-low-touch-terminal/620033/
26 Construction projects overview at Nashville International Airport: https://bnavisionnashville.com/construction/
27 Construction projects overview at Memphis International Airport: https://flymemphis.com/construction/
28 Construction projects overview at San Antonio International Airport : https://flysanantonio.com/business/about-saas/construction-development/
29 List of ASPM 77 airports: https://aspm.faa.gov/aspmhelp/index/ASPM_77.html

CHAPTER 5

Conclusions

This dissertation proposal included three research studies. These studies included analyses of airport performance metrics related to terminal airspace, deicing, and throughput and capacity. The results of our research answer questions that help in understanding and improving different aspects of airport operational efficiency. The scope of our research was related to Key Performance Indicators (KPIs) that are defined by the Federal Aviation Administration (FAA) and the International Civil Aviation Organization (ICAO) to assess the performance of various elements of the aviation system. The first study focused on the key performance indicator KPI08 which measures the efficiency of arrival flights within the terminal airspace, typically the last 100 nautical miles (nmi) of flight. The second study focused on the efficiency of deicing operations. Deicing delays can impact the departure phase of flights in terms of punctuality and additional time in taxi-out, which are defined as KPI01 and KPI02, respectively. Finally, the third study utilizes the airport peak capacity (KPI09) and peak throughput (KPI10) to assess the response of airports to demand and capacity changes during the COVID-19 pandemic and the resilience of airports in the recovery stage afterwards.

In our first study, we examined the methodology behind KPI08. This indicator measures the additional time arrivals spend in the terminal airspace, which is also referred to as A100. Our analysis showed that while the four characteristics currently used by the FAA to benchmark flight A100 times are significant, many other variables outside the control of ATC are also significant. This means that the A100 delay metrics combine the results of external factors and ATC performance. Any change over time in an airport's A100 performance using the FAA methodology requires additional analyses as to the cause. Our results showed that our models are able to predict flight times within the terminal airspace with median RMSE improvements ranging from 39 to 50 seconds in the airport-level models. Our models can be a useful tool to airport management since they use independent variables that are known when the flight is at the 100 nmi crossing.

In the second study, we developed an algorithm that uses ASDE-X data to detect usage of deicing facilities at airports. Our algorithm analyzes speed profiles of aircraft inside a deicing facility to extract deicing, waiting, and taxiing times. Our analysis of 1,504 departures at Chicago O'Hare (ORD) showed that 15.5% of all departures used the centralized deicing facility. There were 19 unique aircraft types in our deicing data, we presented statistics of deicing times per aircraft type including mean deicing times and deicing rates in aircraft/hour. Our results showed that smaller aircraft took longer deicing times on average compared to larger aircraft. Our work fills a gap in the currently available data by providing deicing time estimates that can be useful for other researchers in the area of airport planning and simulation.

The deicing study also presented a discrete-event simulation model that uses the extracted deicing times to simulate the usage of a centralized deicing facility with multiple deicing pads. We ran the simulation model to explore the impacts of changing the capacity utilization strategy at ORD. Our results show that the amount of time spent inside the deicing system can be reduced the most if the airport implements a policy based on our First Come First Served with individual

queues (FCFSi) strategy. The improvement was a result of minimizing the time periods that deicing pads were underutilized by assigning aircraft to available capacity regardless of the operating airline.

In the baseline scenario, aircraft spent 33.9 minutes on average. The simulation results showed that implementing our FCFSi strategy saved 8.5 minutes per aircraft on average compared to the baseline, this is equivalent to 25.1% of aircraft minutes and 26.8% of seat minutes. Our First Come First Served with a combined queue (FCFSc) strategy saved 6.0 minutes per aircraft on average, which is 17.8% of aircraft minutes and 20.3% of seat minutes. The FCFSc improvement to the baseline was less compared to the FCFSi, the latter did not leave a lot of inefficiency in queueing and capacity utilization due to having independent queues per deicing pad. Finally, assigning aircraft to deicing pads based on their deicing speed (slow, medium, fast) showed similar results to the baseline. This was due to certain airlines dominating specific aircraft types in the data, both scenarios showed similar queueing behavior with a small decrease in efficiency in our deicing speed scenario. The deicing speed scenario increased aircraft minutes in the system by 1.3% and aircraft seat minutes by 1.5% on average compared to the baseline.

In the third study, we analyzed airport data for the ASPM 77 airports to study their response to the COVID-19 shock event and recovery period. Our analysis showed that the COVID-19 pandemic resulted in a 67% decline in operations at the lowest point during the pandemic. The impact during the shock time period varied greatly across the airports, ranging from a reduction of 14.8% at MEM to 81.5% at LGA. During the recovery period, the increase in total number of operations ranged from 12% at MEM to 257% at OGG. The latter was an example of increasing demand as a result of loosening travel restrictions related to COVID-19.

We performed a clustering analysis to study airports' response to the COVID-19 pandemic, clustering was done for both the response to the shock event (May-September 2020) and the recovery afterwards (May-September 2021). The airport metrics variables that were used to cluster the airports were the change in peak capacity, peak throughput, and total number of operations. We then compared the airport metrics of the resulting clusters to pre-COVID airport characteristics. There was a number of airport characteristics that were correlated to the changes in airport metrics including cargo operations, operations by major carriers, multi-airport status, hub size, and the rate of transferring passengers. We found that the cargo airports were the most resilient in retaining air traffic demand which was a result of the shift to working from home and businesses moving towards online shopping during the pandemic. Additionally, the percentage of major carriers operations (Delta, United, and American) was significantly correlated to the change in total operations. The data showed that the rate of transferring passengers was a significant variable as it correlated to the change in peak throughput.

The data showed that being located in a multi-airport city was significantly correlated to the decrease in operations during the shock, however, it was not significant in the recovery trends. This finding showed that although air carriers received financial assistance under both the CARES Act and the CRRSA Act, being located in a multi-airport city was only significant when air transportation demand levels were lower than the required minimums, i.e. during the shock time period. Airports in multi-airport areas reacted similarly to the shock event but recovered

differently. For example, New York airports JFK, LGA, and EWR clustered together during the shock event, however, LGA and JFK increased their peak throughput by 160% and 162% compared to 125% at EWR.

Our analysis showed that delays in the system did not change proportionately to the change in operations. Similarly, there were only minor improvements in punctuality. On-time flights at the ASPM 77 airports increased by 9.5% while operations declined by 52% during the shock event time period compared to pre-COVID. Part of this phenomenon was a result of schedule peaking which caused delays due to creating busy hours at the airports. Additionally, a portion of the delays could be attributed to COVID-19 reasons such as increased safety measures at airports, increased passenger boarding times due to social distancing, and staffing shortages.

In summary, our research helped providing insight on potential improvements to the aviation system performance using data analysis and simulation techniques. We studied the performance of the aviation system on multiple levels, namely the arrivals phase, the departures phase, and airports efficiency as a whole. We provided an enhanced methodology to produce more representative benchmarks and a more robust A100 metric. We developed an algorithm that analyzed ASDE-X data to provide realistic deicing time distributions for individual aircraft types. We simulated multiple scenarios of managing aircraft queues at a centralized deicing facility, and showed potential improvements in the time spent in the system. Additionally, we presented a clustering analysis of airports' response to the COVID-19 pandemic from an operational performance metrics perspective. Our results provided insight on significant factors behind the response and recovery patterns in addition to insight on delay and punctuality trends.