Collegiate Aviation Review International

Volume 42 | Issue 1 Peer Reviewed Article #4

03-29-2024

Comparison of Taxi-in and Taxi-out Times by Airport Hub Classification and Number of Hot Spots

Jiansen Wang
Purdue University

Shantanu Gupta
Purdue University

Mary E. Johnson, Ph.D.
Purdue University

Taxi time has been identified as a significant factor that may affect airport capacity, congestion, fuel burn, and emissions. Reducing taxi time at airports may contribute to increasing airport efficiency and capacity, and reducing fuel consumption and emissions. In this paper, average quarter-hour taxi-time data from a sample of 33 U.S. airports was analyzed to explore the difference between taxi-out time and taxi-in time. Using parametric and non-parametric statistical tests, this research found that the mean and median taxi-out time was significantly different from the mean and median taxi-in time for each of the three airport hub classifications (small, medium, large), each of the six numbers of airport hot spots (0, 1, 2, 3, 4, 5), and for each combination of hub classification and numbers of hot spots. The results of this research may provide a better understanding of taxi time at small, medium, and large hubs airports with hot spots. The results may be useful to airport managers and decision makers to improve airport efficiency when faced with competing airport improvement initiatives or projects.

Recommended Citation:
Wang, J., Gupta, S., Johnson, M. E. (2024). Comparison of taxi-in and taxi-out times by airport hub classification and number of hot spots. *Collegiate Aviation Review International*, *42*(1), 75–96. Retrieved from https://ojs.library.okstate.edu/osu/index.php/CARI/article/view/9678/8600

Introduction

The U.S. National Airspace System (NAS) forecasts show an annual enplanement growth rate of 1.7% to 3.5% in the future (FAA, 2023). Due to airport configurations and limitations of land use, some airports cannot expand their aircraft operating area as a way to meet the increasing demand. This limited operating area may lead to longer taxi time, congestion, flight delays, cancellations, and increased aircraft emissions at airports. With increased operations and constrained airport landside space, aircraft taxiway routing and queuing time may increase. Reducing taxi time at airports contributes to increasing airport efficiency and capacity, and reducing fuel consumption and emissions. Previously, multiple factors have been studied for their potential effects on taxi time, such as weather conditions, number of flights, airport runway configurations, taxi distance, and type of operations. The Federal Aviation Administration (FAA, 2022 March) identifies areas on airport diagrams that need attention by pilots, ground vehicle drivers, or air traffic controllers due to the potential for collisions or incursions because of previous incidents or potential risks of incidents. However, there is limited research on the relationships between airport taxi time, airport hub classification, and the number of airport hot spots. This paper contains a comparison of taxi-out to taxi-in time for each of the three airport hub classifications (small, medium, and large) and for each of the six numbers of hot spots (0, 1, 2, 3, 4, 5) on airports.

This research focuses on exploring whether taxi-out time is equal to taxi-in time at airports for each of the three hub classifications (S/M/L) and for each of the six numbers of hot spots (0, 1, 2, 3, 4, 5). In this study, the 11 busiest airports (by number of total operations) from each of the three National Plan of Integrated Airport System (NPIAS) hub classification categories (Small, Medium, Large) (FAA, 2022 October) were selected from the Aviation System Performance Metrics (ASPM) dataset (FAA, n.d.-a). The 20 busiest days (by number of daily operations) were selected for each of the 33 airports and the average quarter-hour taxi-in and
taxi-out time were sampled from the ASPM dataset (FAA, n.d.-a) between 06:00 AM to 10:00 PM from May 01, 2022, to September 30, 2022, for each of the 33 airports. Statistical (parametric and non-parametric) and graphical analyses were used to answer research questions related to comparing taxi-out and taxi-in time. Minitab® and IBM® SPSS® are commercially available statistical packages that were used for statistical and graphical tests.

By exploring the difference between airport taxi-out and taxi-in time for each of the three hub classifications and for each of the six hot spot numbers, the research aims to identify factors affecting taxi operations. Reducing aircraft taxi time may contribute to reducing aircraft emissions, fuel burn, and improving airport capacity. The results of this research may be useful to airport managers and decision makers to improve airport efficiency when faced with competing airport improvement initiatives or projects.

Background and Literature Review

The FAA ASPM dataset describes taxi-in time as the time difference between Wheels On time, Gate In time, and taxi-out time as the time difference between Gate Out time and Wheels Off time (FAA, n.d.-b). Aircraft taxi operations and taxi time length have several impacts on air travel, such as flight delays, congestion, excessive fuel burn, and aircraft emissions. The estimated total cost of flight delays in U.S. airports was 33 billion dollars in 2019, with a rise of 9% compared to 2018 (FAA, n.d.-c). A study found that a majority (60%) of the total flight delays at U.S. major airports were related to taxi-out delays (Laskey et al., 2006).

Longer taxi time contributes to increased engine exhaust emissions, brake dust and tire deposits, and congestion at airports. Due to the limitations of land use and airport configurations, some airports cannot expand their operating area to meet the increasing demand; therefore, expanding the capacity of airports may be restricted. For example, New York LaGuardia (LGA) airport is constrained by the limited space for new runways and ramps (Burgain, 2010). With the growing air traffic flow in the U.S., airports may expect more congestion and longer taxi time, along with associated flight delays, congestion, excessive fuel burn, and aircraft emissions.

When aircraft engines are operating, pollutants are emitted at the airport and surrounding areas. Aircraft operating at airports are significant sources of pollutant emissions that can impact humans (Dissanayaka et al., 2023). The carbon dioxide and nitrogen oxides generated by aviation operations around airports are expected to increase significantly by 2025 and beyond (Burgain, 2010). Aircraft that have to stop and go while taxiing contribute to approximately 18% of the fuel consumption, representing an increase of around 35% compared to the fuel consumption during uninterrupted aircraft operations (Nikoleris et al., 2011). The cost of fuel consumed during taxi operations has a major impact on airline operating costs (Murner, 2012). A better understanding of taxi time and a reduction in taxi time may help save fuel, improve the financial health of airlines, and reduce the impacts on humans and the environment.

Investigating the factors affecting taxi time and constructing taxi time models may reveal patterns of taxi operations that may be used to reduce taxi time and, therefore, mitigate congestion and reduce fuel burn. By reducing total taxi time, airport operation efficiency may be improved, and aircraft emissions may be reduced (Atkin et al., 2010; Atkin et al., 2011). When taxi time is shorter at airports, there is the potential for expanding the capacity of airports while not expanding the aircraft movement area. Less congestion and shorter taxi time contribute to less engine exhaust substances emitted at airports and surrounding areas. Identifying patterns in taxi-in and taxi-out time at airports may assist in reducing the operational costs for airports and airlines.

For statistical modeling and predicting airport taxi time, researchers have studied eleven factors as independent variables in the regression models that may affect taxi time at airports. These factors include weather conditions (Mirmohammadsadeghi et al., 2019; Park & Kim, 2023; Srivastava, 2011), number of flights (Balakrishna et al., 2010; Clewlow et al., 2010; Mirmohammadsadeghi et al., 2019; Park & Kim, 2023; Ravizza et al., 2013; Srivastava, 2011), airport runway configurations (Clewlow et al., 2010; Lordan et al., 2016; Mirmohammadsadeghi et al., 2019; Srivastava, 2011), taxi distance (Lee et al., 2015; Lee et al., 2016; Lordan et al., 2016; Ravizza et al., 2013; Srivastava, 2011), type of operations (Ravizza et al., 2013), taxi speed (Ravizza et al., 2013; Wang et al., 2021), number of turns during taxi operation (Wang et al., 2021), aircraft weight (Lee et al., 2016; Wang et al., 2021), time of the day (Balakrishna et al., 2010; Lee et al., 2016), terminal concourse and gates (Lee et al., 2015; Lee et al., 2016), and aircraft type (Wang et al., 2021). Studies have shown that weather conditions, especially wind, air pressure, and precipitation (Park & Kim, 2023), can significantly affect taxi-in time and taxi-out time at airports (Srivastava, 2011). Another study showed that including weather in the runway capacity simulation model brought an enhancement to the accuracy of the simulation results (Mirmohammadsadeghi et al., 2019). Studies have shown that an increasing the number of flights can result in longer taxi time (Balakrishna et al., 2010; Mirmohammadsadeghi et al., 2019; Park & Kim, 2023; Ravizza et al., 2013; Srivastava, 2011) and that the number of arrival flights is significantly correlated to taxi-out time (Clewlow et al., 2010). When the airport runway configuration has more interactions between arrivals and departure aircraft (intersections or conflicts), the number of arrival flights has a more significant impact on taxi-out time (Clewlow et al., 2010). Researchers have concluded that taxi distance (Lee et al., 2015; Lee et al., 2016) and taxi speed (Wang et al., 2021) significantly contribute to taxi time. In this paper, two potential factors that may affect taxi time at airports are explored – airport hub classification and number of hot spots.

Airport hub classification may be studied as a broad indicator of the number of annual enplanements, number of flights, and size of land area or movement area. Previous studies (Balakrishna et al., 2010; Clewlow et al., 2010; Mirmohammadsadeghi et al., 2019; Park & Kim, 2023; Ravizza et al., 2013; Srivastava, 2011) have studied the number of flights as a factor when analyzing taxi time or developing taxi time prediction models. The number of flights may be affected by the airport's physical size, the number of enplanements, and the aircraft types operating at that airport. In terms of hub classifications, the FAA categorizes publicly owned airports with a minimum of 2,500 annual enplanements as commercial service airports (FAA, n.d.-d). These airports are further divided into Large Hub, Medium Hub, Small Hub, and Non hub airports based on the number of enplanements (FAA, n.d.-d). The FAA defines hub airports as:

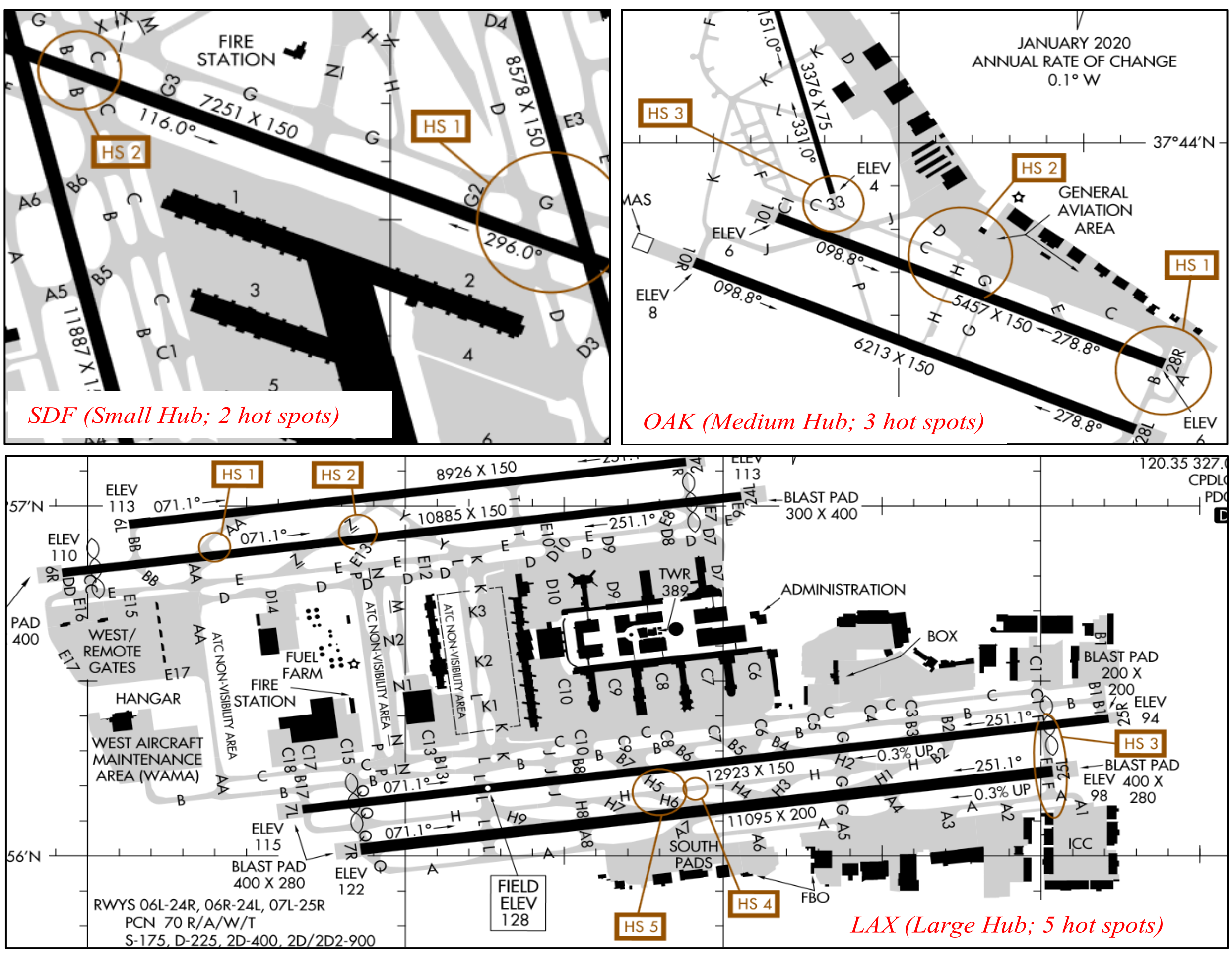
- Large Hub airports as those that receive 1% or more of the U.S. annual commercial enplanements,
- Medium Hub airports as those that receive 0.25% to 1% of the annual enplanements,
- Small Hub airports as those that receive 0.05% to 0.25% of the annual enplanements, and
- Non-hub airports are those that receive less than 0.05% but more than 10,000 of the annual enplanements (FAA, n.d.-d).

Therefore, in this research, airport hub classifications are considered as a broad indicator of the size of the airport. Annual enplanements are associated with the number of flights. This paper focuses on small, medium, and large hub airports as classified by the FAA in the National Plan of Integrated Airport System (NPIAS) report (FAA, 2022 October).

Per the FAA, a hot spot is a location in an airport movement area with a history of collision or runway incursion, therefore needing attention by the pilots or air traffic controllers (FAA, 2022 March). The presence of hot spots may influence pilots to slow down or even stop for safety reasons, affecting the aircraft taxi time. According to the FAA, "hot spots are complex or confusing taxiway/taxiway or taxi/runway intersections" (FAA, 2022 March, para. 2); the criteria for determining hot spots in airports include, but are not limited to, "airport layout, traffic flow, airport marking, signage and lighting, situational awareness, and training" (FAA, 2022 March, para. 2). Researchers have studied airport complexity factors such as the number of intersections (Clewlow et al., 2010), airport layout (Lordan et al, 2016), and traffic flow (Balakrishna et al., 2010; Park & Kim, 2023; Srivastava, 2011) that may contribute to taxi time at airports. Given that these factors are considered when determining airport hot spots, the number of hot spots in airports may be studied as a simplified and broad indicator of airport configuration complexity and the number of complex or confusing intersections in analyzing taxi time. Therefore, this research includes the number of hot spots on airport diagrams as an independent factor to study its effects on taxi time at small, medium, and large hub airports. Airport hot spots are marked on the FAA airport diagrams, and the hot spot data is also published in the Digital Chart Supplement (d-CS) by the FAA (FAA, 2022 March). Figure 1 shows examples of hot spots on SDF (small hub), OAK (medium hub), and LAX (large hub) airport diagrams (FAA, 2022 May).

Figure 1

Examples of Hot Spots on SDF, OAK, and LAX Airport Diagrams



Note. The images are adapted from the FAA Airport Diagrams (26 JAN 2023 to 23 FEB 2023) of SDF, OAK, and LAX airports (FAA, 2022 May).

Gupta et al. (2023) studied the effect of airport hub classification and the number of hot spots on taxi time by conducting one-way ANOVA parametric and Kruskal-Wallis non-parametric tests. Separate tests were conducted for taxi-out time and taxi-in time as response variables, and airport hub classifications (3 levels – S/M/L) and the number of hot spots (6 levels – 0, 1, 2, 3, 4, 5) as factors. The study found that taxi-out and taxi-in time each varies by airport hub classification, i.e., taxi-out time was different for small, medium, and large hubs, and taxi-in time was different for small, medium, and large hubs. It also found that taxi-out time at airports with (0, 3, 4, 5) hot spots was different from the taxi-out time at airports with (1, 2) hot spots. Taxi-in time at airports with (1, 2, 3, 4) hot spots was different from the taxi-in time at airports with (0, 5) hot spots (Gupta et al., 2023). However, Gupta et al. (2023) did not test if the mean of taxi-out time equals the mean of taxi-in time for airports with the same hub classification and for airports with the same number of hot spots.

This paper builds upon the results of Gupta et al. (2023) to further explore taxi time and compare taxi-out time with taxi-in time. In this study, the difference between taxi-out time and taxi-in time is analyzed in 1) a combined sample of the same 33 airports, 2) three separate tests for each airport hub classification (S/M/L), and 3) six separate tests for each number of hot spots at airports (0, 1, 2, 3, 4, 5). Specifically, this research aims to answer three research questions:

RQ1: Is taxi-out time equal to taxi-in time for airports (without considering hubs or hot spots)?
RQ2: Is taxi-out time equal to the taxi-in time for each of the three airport hub classifications?
RQ3: Is taxi-out time equal to the taxi-in time for each of the six numbers of hot spots?
RQ4: Is taxi-out time equal to the taxi-in time for each combination of the three airport hub classifications and the six numbers of spots?

Methodology

This section presents the data sources, collection, consolidation, and data analysis used to answer the research questions.

Data Sources: For this paper, the researchers collected airport-related data from ASPM for airport operations data (FAA, n.d.-e) and for taxi time (FAA, n.d.-f), NPIAS 2023-2027 for hub classification (FAA, 2022 October), and FAA airport diagrams for hot spots (FAA, 2022 May).

Data Collection: The researchers used the ASPM (FAA, n.d.-a) dataset for the list of airports, dates, and the number of daily Departures and Arrivals for Metric Computation for the selected timeframe between 05/01/2022 and 09/31/2022 to capture a busy travel season. The hub classification of airports was found from the NPIAS 2023-2027 (FAA, 2022 October) dataset (*Appendix A: List of NPIAS Airports* (FAA, 2022 October)). The ASPM dataset and the NPIAS dataset were crossmatched to find the hub classification (Small/Medium/Large) for each of the 77 ASPM airports.

Based on the total number of Departures and Arrivals from 05/01/2022 to 09/30/2022, the researchers formed a sample of 33 airports by selecting the 11 busiest airports from each of the three hub classifications (S/M/L). Researchers selected the 20 busiest days for each of the 33 airports based on the total daily number of departures and arrivals at each airport. Note: The dates of the busiest days may differ among the airports. The researchers collected quarter-hour taxi-time data between 6:00 AM and 10:00 PM (local time) from the ASPM dataset (FAA, n.d.-a) for each of the 20 busiest days for each of the 33 airports. The FAA Airport Diagrams

(01/26/2023 to 02/23/2023) (FAA, 2022 May) were used to obtain the number of hot spots at each of the 33 sampled airports.

Data Consolidation: Data from 33 airports, obtained from the ASPM dataset, NPIAS dataset, and FAA airport diagrams, were combined into a single spreadsheet. Tabulated data for each airport included quarter-hour taxi-in and taxi-out time between 6:00 AM and 10:00 PM (local time) for the 20 busiest days, departures and arrivals in each quarter, NPIAS hub classification, and the number of hot spots at airports. This consolidated data repeats that of Gupta et al. (2023), who conducted ANOVA on taxi-out and taxi-in time (as separate response variables), and hub classification and the number of hot spots as factors. This paper focuses on comparing taxi-out time with taxi-in time for each of the three hub classifications and six numbers of hot spots. In total, 39,268 observations were collected across ten fields, as shown in a snippet of the spreadsheet in Figure 2.

Figure 2

Sample of Consolidated Data Table Using Three Sources of Data

ASPM Dataset (FAA, n.d.-f)								**NPIAS 2023-2027** (FAA, 2022, Oct)	**FAA Airport Diagrams** (FAA, 2022, May)
Facility	Date	Quarter	Hour	Departures For Metric Computation	Average taxi out time	Arrivals For Metric Computation	Average taxi in time	NPIAS Hub Classification	Number of Hot Spots
PVD	09/01/2022	3	10	3	8.67	0	0	S	5
PVD	09/01/2022	3	12	3	28	1	10	S	5
PVD	09/01/2022	3	13	2	11.5	1	5	S	5
PVD	09/01/2022	3	14	2	9.5	2	5.5	S	5
ANC	8/24/2022	4	8	0	0	3	5	M	2
ANC	8/24/2022	4	9	3	11.33	1	6	M	2
ANC	8/24/2022	4	11	4	15.75	3	9.67	M	2
ANC	8/24/2022	4	12	4	12.75	3	6.33	M	2
CLT	6/6/2022	1	9	35	22.66	4	6.75	L	3
CLT	6/6/2022	1	10	3	14	30	12.6	L	3
CLT	6/6/2022	1	11	38	22.18	2	5.5	L	3
CLT	6/6/2022	1	12	5	15	20	11.85	L	3

Note. The data in this figure has the same headings as Gupta et al. (2023)

Data Analysis: To answer the three research questions, a parametric test (Paired Samples *t* Test) and a non-parametric test (Wilcoxon Signed Rank Test) were selected to compare means and medians for quarter-hour taxi-out and taxi-in time. The researchers identified outliers and unequal variance in the samples during data processing and, therefore, approached the research questions using both parametric and non-parametric methods to analyze the mean and median quarter-hour taxi time. The researchers used commercially available statistical packages (Minitab® and IBM® SPSS®) to analyze and visualize the data. There were zeros in the taxi-time data for the sampled airports in each of the three hub classifications. Zeros in taxi-in data represent no arrival operations during that quarter hour; likewise, zeros in taxi-out data represent no departures during that quarter hour. The number of zeros in taxi-in and taxi-out data was highest for small hub airports (taxi-out: 3,195 zeros; taxi-in: 3,318 zeros), followed by medium hub (taxi-out: 994 zeros; taxi-in: 999 zeros) and large hub airports (taxi-out: 45 zeros; taxi-in: 41

zeros). In this paper, zeros have been removed from the data, i.e., any quarter-hour taxi time paired with either taxi-out or taxi-in time reported as zero was removed from the data. Therefore, the taxi time used in the analysis is based on the condition that there was at least one departure and one arrival at the airport in that specific quarter-hour.

For RQ1, taxi-out and taxi-in time were treated as a combined single-paired sample without considering the hub classifications or hot spots. For RQ2, the data was segregated according to hub classifications (S/M/L) to test if taxi-out and taxi-in time were equal for each of the three NPIAS hub classifications. Individual tests were conducted to compare mean and median taxi time for small, medium, and large hub airports. For RQ3, the data was segregated according to the number of hot spots (0, 1, 2, 3, 4, 5) to test if taxi-out and taxi-in time were equal for airports with different numbers of hot spots. Individual tests were conducted to compare mean and median taxi time for 0, 1, 2, 3, 4, and 5 hot spots in airports. For RQ4, the data was segregated according to each of the airport hub classifications and hot spot combinations (for example, small hub airports with four hot spots, medium hub airports with one hot spot, large hub airports with three hot spots, etc.). A total of 16 combinations were segregated from the sample data. For each combination, paired *t*-tests and Wilcoxon Signed Rank tests were used to compare taxi-out time with taxi-in time. Table 1 lists the statistical tests, the null and alternate hypotheses, and the data samples.

Table 1
Statistical Tests, Hypotheses, and Data Samples to Answer the Research Questions

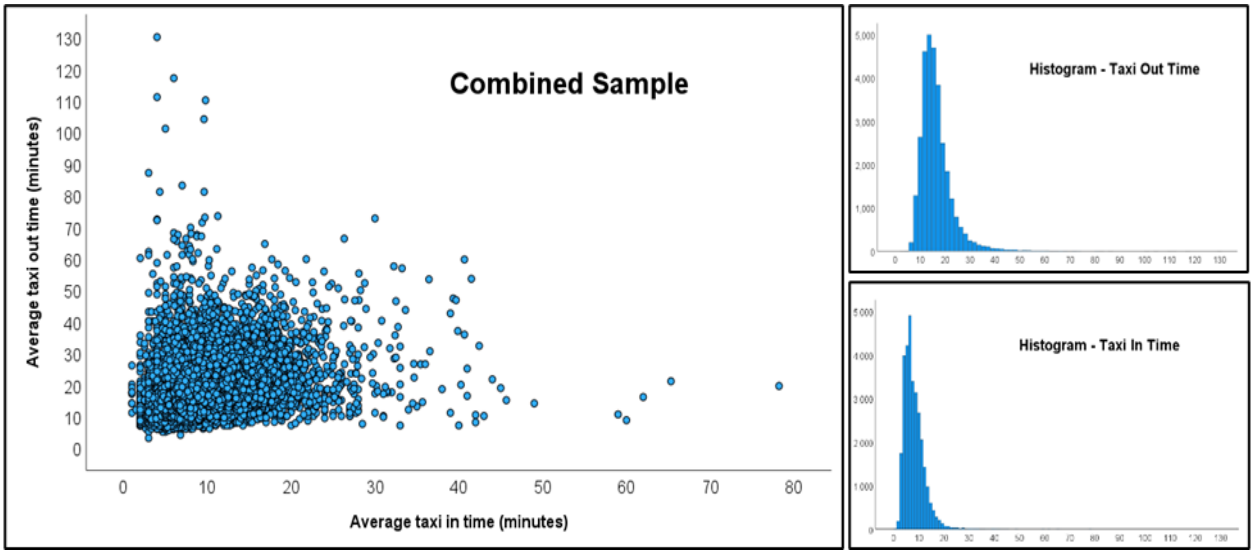
Test	Hypotheses	Data Sample
Paired Samples *t*-test	H_o: $\mu_{taxi\text{-}out} - \mu_{taxi\text{-}in} = 0$ H_a: $\mu_{taxi\text{-}out} - \mu_{taxi\text{-}in} \neq 0$	RQ1: combined sample; RQ2: data segregated by hub classification; RQ3: data segregated by number of hot spots; RQ4: data segregated by each combination of hub classification and number of hot spots
Wilcoxon Signed Rank Test	H_o: $\eta_{taxi\text{-}out} - \eta_{taxi\text{-}in} = 0$ H_a: $\eta_{taxi\text{-}out} - \eta_{taxi\text{-}in} \neq 0$	

Results

This section presents the results of the statistical tests used to answer research questions. The researchers compared the taxi-in and taxi-out time as a combined sample at the three NPIAS hub classifications (S/M/L) and across the different numbers of hot spots at airports. Results of statistical tests and charts are shown within each research question.

RQ1. Is taxi-out time equal to taxi-in time for airports (without considering hubs or hot spots)? To answer this question, taxi-in time and taxi-out time were treated as a combined single paired sample without considering the hub classifications or number of hot spots. Figure 3 shows the scatter plot and histograms for taxi-out and taxi-in time. Table 2 presents the consolidated results of the statistical tests.

Figure 3
Scatter Plot and Histograms for Taxi-Out and Taxi-In Time from All 33 Airports



Note. In the scatterplots, the scales for the x-axis (zero to 80 minutes) and for the y-axis (zero to 120 minutes) are different. Zeros have been removed from the taxi-time data.

The sample mean and median taxi-out time (M = 16.32 minutes, Mdn = 15.09 minutes) was greater than the sample mean and median taxi-in time (M = 8.00 minutes, Mdn = 7.22 minutes) for the sampled airports. The sample variation in taxi-out time (SD = 6.35 minutes) was greater than the sample variation of taxi-in time (SD = 4.00 minutes).

<u>Paired *t*-test:</u> Using the data collected, an alpha of 0.05, and the paired samples *t*-test to compare means, the researchers *rejected the null hypothesis* (p<0.001) that the population mean taxi-out time was equal to the population mean taxi-in time at airports. The 95% confidence interval for the difference $\mu_{taxi\text{-}out} - \mu_{taxi\text{-}in}$ was (8.25, 8.39) minutes.

<u>Wilcoxon Signed Rank Test:</u> Using the data collected, an alpha of 0.05, and the Wilcoxon Signed Rank Test to compare medians, the researchers *rejected the null hypotheses* (p<0.001) that the population median taxi-out time was equal to the population median taxi-in time. The 95% confidence interval for the median difference $\eta_{taxi\text{-}out} - \eta_{taxi\text{-}in}$ was (7.78, 7.89) minutes.

Table 2

Consolidated Results of RQ1 (taxi-out vs. taxi-in time for the combined sample)

Average Taxi Time	**Descriptive Statistics** (all times are in minutes)				**Difference Tests for Mean and Median** (alpha = 0.05) (all times are in minutes)						
					Paired *t*-test Ho: $\mu_{taxi\text{-}out} - \mu_{taxi\text{-}in} = 0$ Ha: $\mu_{taxi\text{-}out} - \mu_{taxi\text{-}in} \neq 0$				**Wilcoxon Signed Rank Test** Ho: $\eta_{taxi\text{-}out} - \eta_{taxi\text{-}in} = 0$ Ha: $\eta_{taxi\text{-}out} - \eta_{taxi\text{-}in} \neq 0$		
	n	Mean time (*M*)	Median time (*Mdn*)	*SD*	Mean time (*M*)	*SD*	95% *CI* for mean diff	*p*	Median	95% confidence achieved	*p*
Taxi-out	30,676	16.32	15.09	6.35	8.32	6.11	(8.25, 8.39)	<0.001	7.39	(7.78, 7.89)	<0.001
Taxi-in	30,676	8.00	7.22	4.00							

RQ2. Is taxi-out time equal to taxi-in time for each of the three airport hub classifications? This question was answered individually for small, medium, and large hub airport data. Paired *t*-test and Wilcoxon Signed Rank test were conducted for each of the three hub classifications. Figure 4 shows scatter plots and histograms for taxi-out and taxi-in time at each of the three airport hub classifications. Table 3 presents the consolidated results of the statistical tests.

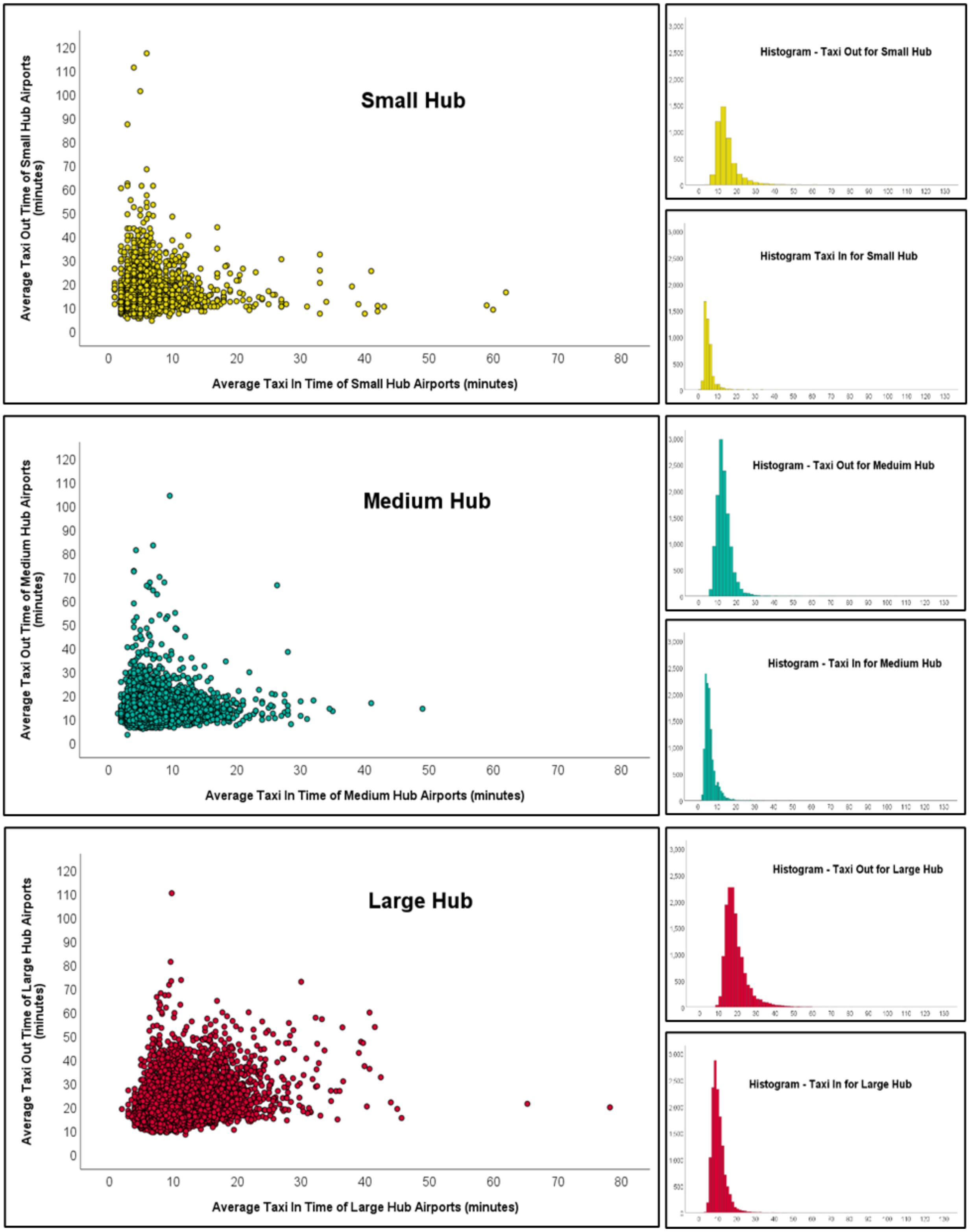
The sample mean and median taxi-out time (minutes) for large hub airports (*M* = 19.30 minutes, *Mdn* = 17.78 minutes) were longer than that of medium hubs airports (*M* = 13.46 minutes, *Mdn* = 12.75 minutes), which were longer than small hub airports (*M* = 14.75 minutes, *Mdn* = 13.00 minutes). The sample mean and median taxi-in time for large hub airports (*M* = 10.32 minutes, *Mdn* = 9.63 minutes) were longer than that of medium hub airports (*M* = 6.29 minutes, *Mdn* = 5.50 minutes), which were longer than small hub airports (*M* = 5.49 minutes, *Mdn* = 5.00 minutes). Small hub airports were found to have the highest sample variation in taxi-out (*SD* = 6.62 minutes) time. Large hub airports were found to have the highest sample variation in taxi-in (*SD* = 3.49 minutes) time.

Paired *t*-test: Using the data collected, an alpha of 0.05, and the paired-t test to compare means, the researchers *rejected the null hypotheses* ($p<0.001$) that the population mean taxi-out time was equal to the population mean taxi-in time at each of the airport hub classifications (S/M/L). The sample mean difference in the taxi-out and taxi-in time was found for small hubs (9.26 minutes), medium hubs (7.17 minutes), and large hub airports (8.89 minutes). The 95% confidence interval for the mean difference $\mu_{taxi\text{-}out} - \mu_{taxi\text{-}in}$ was (9.05, 9.47) minutes at small hub airports, (7.08, 7.26) minutes at medium hub airports, and (8.88, 9.08) minutes at large hub airports.

Wilcoxon Signed Rank Test: Using the data collected, an alpha of 0.05, and the Wilcoxon Signed Rank Test to compare medians, the researchers *rejected the null hypotheses* ($p<0.001$) that the population median taxi-out time was equal to the population median taxi-in time at each of the airport hub classifications (S/M/L). The sample median difference in the taxi-out and taxi-in time was found for small hubs (8.67 minutes), medium hubs (7.00 minutes), and large hub airports (8.36 minutes). The 95% confidence interval for the median difference $\eta_{taxi\text{-}out} - \eta_{taxi\text{-}in}$ was (8.50, 8.80) minutes at small hub airports, (6.92, 7.06) minutes at medium hub airports, and (8.28, 8.45) minutes at large hub airports.

Figure 4

Scatter Plot and Histograms for Taxi Time at Small, Medium, and Large Hub Airports



Note. In the scatterplots, the scales for the x-axis (zero to 80 minutes) and for the y-axis (zero to 120 minutes) are different. Zeros have been removed from the taxi-time data.

Table 3

Statistical results of RQ2 (taxi-out vs. taxi-in time by airport hub classifications)

Airport Hub Classification		Descriptive Statistics[a] (all times are in minutes)				Difference Tests for Mean and Median (alpha = 0.05) (all times are in minutes)						
						Paired *t*-test $H_o: \mu_{taxi\text{-}out} - \mu_{taxi\text{-}in} = 0$ $H_a: \mu_{taxi\text{-}out} - \mu_{taxi\text{-}in} \neq 0$				**Wilcoxon Signed Rank Test** $H_o: \eta_{taxi\text{-}out} - \eta_{taxi\text{-}in} = 0$ $H_a: \eta_{taxi\text{-}out} - \eta_{taxi\text{-}in} \neq 0$		
		n	Mean time (*M*)	Median time (*Mdn*)	*SD*	Mean time (*M*)	*SD*	95% *CI* for mean diff	*p*	Median	95% confidence achieved	*p*
S	Taxi out	4,707	14.75	13.00	6.62	9.26	7.35	(9.05, 9.47)	<0.001	8.67	(8.50, 8.80)	<0.001
	Taxi in	4,707	5.49	5.00	3.48							
M	Taxi out	11,977	13.46	12.75	4.50	7.17	5.23	(7.08, 7.26)	<0.001	7.00	(6.92, 7.06)	<0.001
	Taxi in	11,977	6.29	5.50	2.97							
L	Taxi out	13,992	19.30	17.78	6.30	8.89	6.21	(8.88, 9.08)	<0.001	8.36	(8.28, 8.45)	<0.001
	Taxi in	13,992	10.32	9.63	3.49							

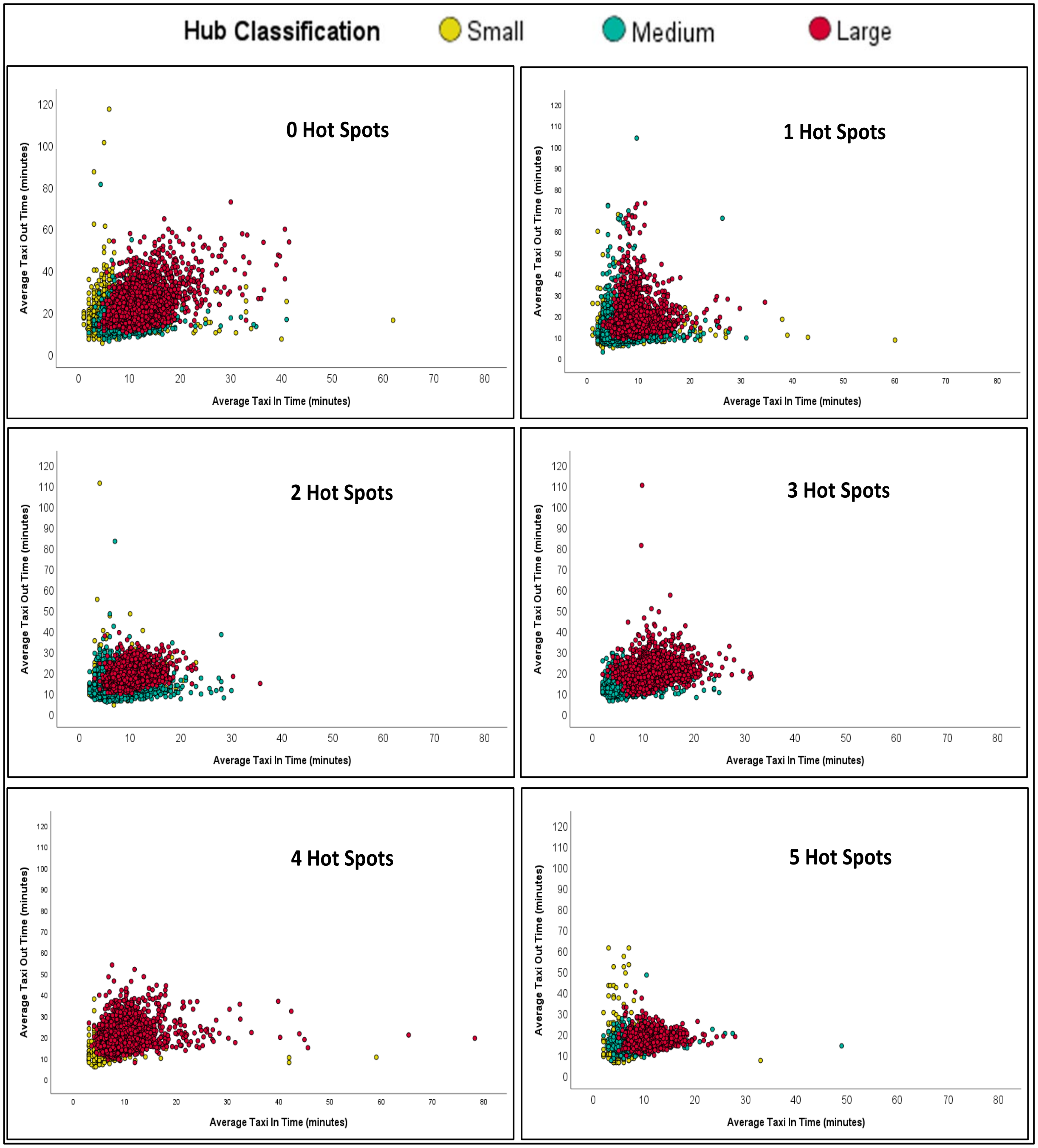
Note. [a]Descriptive statistics are from Gupta et al. (2023). The sample size *n* represents the number of all quarter-hours with at least one departure and one arrival, i.e., zeros have been removed from the taxi-time data.

RQ3. Is taxi-out time equal to taxi-in time for each of the six numbers of hot spots? This question was answered individually for each number of hot spots (0, 1, 2, 3, 4, 5). Paired *t*-test and Wilcoxon Signed Rank test were conducted for each of the hot spots. Figure 5 shows scatter plots and histograms for sample taxi-out and taxi-in time at each of the six hot spot numbers. Table 4 presents the consolidated results of the statistical tests.

Table 4 shows that the sample mean and median taxi-out time (minutes) were longer than the mean and median taxi-in time for airports with (0, 1, 2, 3, 4, and 5) hot spots. The sample mean and median taxi-out time were highest for airports with four hot spots (*M* = 19.23 minutes, *Mdn* = 18.24 minutes). The sample mean taxi-out time was lowest for airports with two hot spots (*M* = 14.87 minutes), and median taxi-out time was lowest for airports with one hot spot (*Mdn* = 13.22 minutes). The sample mean and median taxi-in time were highest for airports with 0 hot spot (*M* = 9.36 minutes, *Mdn* = 8.88 minutes) and lowest for airports with one hot spot (*M* = 6.62 minutes, *Mdn* = 6.00 minutes). Airports with 0 hot spots were found to have the highest sample variation in taxi-out (*SD* = 8.21 minutes) and taxi-in (*SD* = 4.93 minutes) time.

Figure 5

Scatter Plot of Taxi-Out vs Taxi-In Time at Airports with Different Number of Hot Spots



Note. In the sampled data, there were no small hub airports with three hot spots and no medium hub airports with four hot spots. In the scatterplots, the scales for the x-axis (zero to 80 minutes) and for the y-axis (zero to 120 minutes) are different. Zeros have been removed from the taxi-time data.

<u>Paired *t*-test:</u> Using the data collected, an alpha of 0.05, and the paired-t test to compare means, the researchers *rejected the null hypotheses* ($p<0.001$) that the population mean taxi-out time was equal to the population mean taxi-in time at airports with 0, 1, 2, 3, 4, and 5 hot spots. The minimum sample mean difference in the taxi-out and taxi-in time was 4.87 minutes at airports

with two hot spots, and the maximum sample mean difference was 10.07 minutes at airports with four hot spots. The 95% confidence interval for the mean difference $\mu_{taxi\text{-}out} - \mu_{taxi\text{-}in}$ was (9.42, 9.84) minutes at airports with 0 hot spots, (8.10, 8.41) minutes at airports with one hot spot, (7.42, 7.64) minutes at airports with two hot spots, (7.47, 7.75) minutes at airports with three hot spots, (9.84, 10.30) minutes at airports with four hot spots, and (7.49, 7.87) minutes at airports with five hot spots.

Wilcoxon Signed Rank Test: Using the data collected, an alpha of 0.05, and the Wilcoxon Signed Rank Test to compare medians, the researchers *rejected the null hypotheses* (p<0.001) that the population median taxi-out time was equal to the population median taxi-in time at airports with 0, 1, 2, 3, 4, and 5 hot spots. The minimum sample median difference in the taxi-out and taxi-in time was 7.34 minutes and occurred at airports with three hot spots, and the maximum sample median difference was 9.86 minutes and occurred at airports with four hot spots. The 95% confidence interval for the median difference $\eta_{taxi\text{-}out} - \eta_{taxi\text{-}in}$ was (8.75, 9.10) minutes at airports with 0 hot spots, (7.41, 7.62) minutes at airports with one hot spot, (7.27, 7.46) minutes at airports with two hot spots, (7.22, 7.46) minutes at airports with three hot spots, (9.68, 10.04) minutes at airports with four hot spots, and (7.25, 7.51) minutes at airports with five hot spots.

Table 4

Statistical Results of RQ3 (taxi-out vs taxi-in time by different number of hot spots)

Number of Hot Spots		**Descriptive Statistics**[a] (all times are in minutes)				**Difference Tests for Mean and Median** (alpha = 0.05) (all times are in minutes)						
						Paired *t*-test H₀: $\mu_{taxi\text{-}out} - \mu_{taxi\text{-}in} = 0$ Hₐ: $\mu_{taxi\text{-}out} - \mu_{taxi\text{-}in} \neq 0$				**Wilcoxon Signed Rank Test** H₀: $\eta_{taxi\text{-}out} - \eta_{taxi\text{-}in} = 0$ Hₐ: $\eta_{taxi\text{-}out} - \eta_{taxi\text{-}in} \neq 0$		
		n	Mean time (*M*)	Median time (*Mdn*)	*SD*	Mean time (*M*)	*SD*	95% *CI* for mean diff	*p*	Median	95% confidence achieved	*p*
0	Taxi out	5,183	18.98	17.00	8.21	9.63	7.64	(9.42, 9.84)	<0.001	8.94	(8.75, 9.10)	<0.001
	Taxi in	5,183	9.36	8.88	4.93							
1	Taxi out	7,022	14.88	13.22	6.72	8.26	6.76	(8.10, 8.41)	<0.001	7.50	(7.41, 7.62)	<0.001
	Taxi in	7,022	6.62	6.00	3.36							
2	Taxi out	7,775	14.87	14.07	4.65	7.53	4.87	(7.42, 7.64)	<0.001	7.37	(7.27, 7.46)	<0.001
	Taxi in	7,775	7.34	6.69	3.09							
3	Taxi out	4,781	16.48	15.80	5.52	7.61	4.95	(7.47, 7.75)	<0.001	7.34	(7.22, 7.46)	<0.001
	Taxi in	4,781	8.87	8.50	4.27							
4	Taxi out	2,882	19.23	18.24	5.77	10.07	6.18	(9.84, 10.30)	<0.001	9.86	(9.68, 10.04)	<0.001
	Taxi in	2,882	9.16	8.30	4.38							
5	Taxi out	3,033	15.80	15.40	4.77	7.68	5.32	(7.49, 7.87)	<0.001	7.38	(7.25, 7.51)	<0.001
	Taxi in	3,033	8.12	7.63	3.46							

Note. [a]Descriptive statistics are from Gupta et al. (2023). The sample size *n* represents the number of all quarter-hours with at least one departure and one arrival, i.e., zeros have been removed from the taxi-time data.

RQ4. Is taxi-out time equal to taxi-in time for each of the three airport hub classifications and six of the numbers of hot spots? This question was answered individually for each combination of the airport hub classification and number of hot spots. The data was segregated into 16 groups. Paired *t*- tests and Wilcoxon Signed Rank tests were conducted for each of the groups. Table 5 presents the consolidated results of the statistical tests.

As shown in Table 5, the sample mean and median taxi-out time (minutes) were longer than the sample mean and median taxi-in time for each of the 16 combinations of hub classification and number of hot spots. The sample mean and median taxi-out time was highest for large hub airports with 0 hot spots (M = 22.29 minutes, Mdn = 19.50 minutes) and lowest for medium hub airports with one hot spot (M = 12.67 minutes, Mdn = 11.60 minutes). The sample mean taxi-in time was highest for large hub airports with 0 hot spots (M = 11.89 minutes), while the sample median taxi-in time was highest for large hub airports with three hot spots (Mdn = 11.4 minutes). The sample mean taxi-in time was lowest for small hub airports with five hotspots (M = 4.85 minutes), while the sample median taxi-in time was lowest for small hub airports with one hot spot (Mdn = 4.00 minutes). Small hub airports with five hot spots were found to have the highest sample variation in taxi-out time (SD = 8.62 minutes). Small hub airports with four hot spots were found to have the highest sample variation in taxi-in time (SD = 4.58 minutes), primarily due to a relatively lower sample size. The sample mean taxi-out time at small hub airports with 0, 1, or 2 hot spots were longer than the sample mean taxi-out time at median hub airports with 0, 1, or 2 hot spots. In addition, the researchers found some interesting and counter-intuitive results in the collected data sample. For example, as shown in Table 5, small and large hub airports with 0 hot spots were found to have the highest sample mean difference between taxi-out and taxi-in time as compared to their other combinations. While at medium hub airports, the highest sample mean difference between taxi-out and taxi-in time was at airports with five hotspots.

Paired *t*-test: Using the data collected, an alpha of 0.05, and the paired-t test to compare means, the researchers *rejected the null hypotheses* ($p<0.001$) that the population mean taxi-out time was equal to the population mean taxi-in time at each combination of small, medium, and large airports and 0, 1, 2, 3, 4, and 5 hot spots. The minimum sample mean difference in the taxi-out and taxi-in time was 5.71 minutes at medium hub airports with 0 hot spots, and the maximum sample mean difference was 10.97 minutes at small hub airports with 0 hot spots.

Wilcoxon Signed Rank Test: Using the data collected, an alpha of 0.05, and the Wilcoxon Signed Rank Test to compare medians, the researchers *rejected the null hypotheses* ($p<0.001$) that the population median taxi-out time was equal to the population median taxi-in time at small, medium, and large airports with 0, 1, 2, 3, 4, and 5 hot spots. The minimum sample median difference in the taxi-out and taxi-in time was 5.42 minutes at medium hub airports with 0 hot spots, and the maximum sample median difference was 10.33 minutes and occurred at small hub airports with 0 hot spots.

Table 5 *Statistical Results of RQ4 (taxi-out time vs. taxi-in time for each combination of hub classification and number of hot spots)*

Hub Classification	Number of Hot spots		Descriptive Statistics (all times are in minutes)				Difference Tests for Mean and Median (alpha = 0.05) (all times are in minutes)						
							Paired *t*-test H$_o$: $\mu_{taxi\text{-}out} - \mu_{taxi\text{-}in} = 0$ H$_a$: $\mu_{taxi\text{-}out} - \mu_{taxi\text{-}in} \neq 0$				Wilcoxon Signed Rank Test H$_o$: $\eta_{taxi\text{-}out} - \eta_{taxi\text{-}in} = 0$ H$_a$: $\eta_{taxi\text{-}out} - \eta_{taxi\text{-}in} \neq 0$		
			n	Mean time (*M*)	Median time (*Mdn*)	*SD*	Mean time (*M*)	*SD*	95% *CI* for mean diff	*p*	Median	95% confidence achieved	*p*
S	0	Taxi out	1595	16.50	15.00	7.52	10.97	8.16	(10.57, 11.37)	<0.001	10.33	(1.00, 10.60)	<0.001
		Taxi in	1595	5.53	5.00	3.81							
	1	Taxi out	1014	13.18	12.00	5.00	8.13	6.43	(7.73, 8.53)	<0.001	8.00	(7.67, 8.17)	<0.001
		Taxi in	1014	5.05	4.00	3.99							
	2	Taxi out	1208	14.23	13.33	5.31	8.22	5.73	(7.90, 8.54)	<0.001	7.88	(7.65, 8.09)	<0.001
		Taxi in	1208	6.01	5.50	2.55							
	4	Taxi out	323	13.42	12.75	4.48	7.57	6.54	(6.85, 8.28)	<0.001	7.50	(7.00, 8.00)	<0.001
		Taxi in	323	5.86	5.00	4.58							
	5	Taxi out	567	14.53	13.00	8.62	9.68	8.83	(8.95, 10.41)	<0.001	8.50	(8.00, 8.75)	<0.001
		Taxi in	567	4.85	4.50	1.98							
M	0	Taxi out	1044	14.73	14.00	4.99	5.71	6.29	(5.33, 6.09)	<0.001	5.42	(5.11, 5.75)	<0.001
		Taxi in	1044	9.02	8.00	3.86							
	1	Taxi out	3462	12.67	11.60	5.50	7.25	6.06	(7.04, 7.45)	<0.001	6.79	(6.67, 6.92)	<0.001
		Taxi in	3462	5.43	4.67	2.63							
	2	Taxi out	4017	13.61	13.00	4.23	7.09	5.09	(6.93, 7.25)	<0.001	6.99	(6.84, 7.12)	<0.001
		Taxi in	4017	6.52	6.00	2.86							
	3	Taxi out	2268	12.93	12.50	3.09	7.32	3.90	(7.16, 7.48)	<0.001	7.25	(7.09, 7.38)	<0.001
		Taxi in	2268	5.61	5.00	2.39							
	5	Taxi out	1186	15.11	14.96	2.99	8.21	3.72	(8.00, 8.43)	<0.001	8.24	(8.05, 8.42)	<0.001
		Taxi in	1186	6.90	6.33	2.62							
L	0	Taxi out	2544	22.29	19.50	8.34	10.40	7.28	(10.11, 10.68)	<0.001	9.56	(9.29, 9.83)	<0.001
		Taxi in	2544	11.89	11.00	4.32							
	1	Taxi out	2546	18.56	16.44	7.23	9.68	7.49	(9.39, 9.97)	<0.001	8.55	(8.34, 8.8)	<0.001
		Taxi in	2546	8.88	8.30	2.70							
	2	Taxi out	2550	17.17	16.36	4.03	7.90	3.88	(7.75, 8.05)	<0.001	7.68	(7.54, 7.82)	<0.001
		Taxi in	2550	9.27	8.75	2.73							
	3	Taxi out	2513	19.68	19.04	5.27	7.88	3.88	(7.65, 8.10)	<0.001	7.47	(7.28, 7.66)	<0.001
		Taxi in	2513	11.81	11.40	3.36							
	4	Taxi out	2559	19.96	18.78	5.49	10.39	6.06	(10.15, 10.62)	<0.001	10.15	(9.96, 10.34)	<0.001
		Taxi in	2559	9.57	8.67	4.17							
	5	Taxi out	1280	17.00	16.63	3.20	6.29	3.96	(6.07, 6.51)	<0.001	6.17	(5.98, 6.36)	<0.001
		Taxi in	1280	10.71	10.25	2.74							

Note. The sample size *n* represents the number of all quarter-hours with at least one departure and one arrival, i.e., zeros have been removed from the taxi-time data.

Discussion

Taxi time has been identified as a significant factor that may affect airport capacity, congestion, fuel burn, and emissions. In this paper, average quarter-hour taxi-time data (30,676 data points) were collected from 33 U.S. airports (11 each from small, medium, and large hubs) and analyzed to compare taxi-out time with taxi-in time for each of the three hub classifications (S/M/L) and each of the six numbers of hot spots (0, 1, 2, 3, 4, 5). For the research questions in this paper, the researchers used paired *t*-tests to compare the mean taxi-out time with mean taxi-in time. Due to the large number of outliers in the data, the researchers also used the Wilcoxon Signed Rank Test to compare the median taxi-out time with median taxi-in time. The results were consistent between the two tests.

This research uses the following assumptions:

1. The average quarter-hour taxi time data collected from the FAA ASPM dataset represent the real taxi time at the U.S. hub airports.
2. The sample of 33 chosen airports can be representative of U.S. hub airports.
3. The NPIAS hub classification can reflect the number of flights.
4. The number of hot spots can reflect the airport physical configuration complexity.
5. Pairs of average quarter-hour taxi time data with either taxi-out or taxi-in time reported as zero were removed from the data. The taxi time used in the analysis is based on the condition that there was at least one departure and one arrival in that specific quarter-hour.

This research is subjected to the following limitations and delimitations:

1. The taxi time samples in this research are collected from 11 of the busiest U.S. hub airports for each of the three hub classifications (small, medium, and large hubs).
2. The population of this research is the hub airports in the U.S. in the ASPM 77.
3. The number of hot spots in U.S. airports ranges from 0 to 9. The number of hotspots for the 33 airports selected from the list of ASPM 77 airport range from 0 to 5. There were no small hub airports with three hot spots and no medium hub airports with 4 hot spots in the sampled 33 airports, which could bias the analysis results or reduce the statistical power of this research.
4. The taxi time data collected from the ASPM dataset are average quarter-hour taxi time, not the specific taxi time of each flight. Therefore, the standard deviation of taxi time is expected to be lower if individual taxi time data were available.
5. The data collected resulted in violations of normality and constant variance assumptions for the *t*-tests. The large sample size (n = 30,676) in this study may improve the validity of the *t*-tests in terms of normality assumption with respect to the Central Limit Theorem.
6. This research does not consider the variations in taxi time caused by different airport operational capacities or demands, airport management, equipment, weather, aircraft type, operation type (maintenance or refueling), or airport physical size and configurations.

Researchers must pay attention to the definitions for taxi time, the procedures to measure taxi time across airports, and the exact events used to start and stop the time measurement. According to the FAA ASPM dataset, taxi-in time is the time difference between Wheels On time and Gate In time; and taxi-out time is the time difference between Gate Out time and Wheels Off time (FAA, n.d.-b). International Civil Aviation Organization (ICAO) defines taxi/ground idle time as the time differences between “initial starting of the propulsion engine(s) and the initiation of the take-off roll; and between the time of runway turn-off and final shutdown of all propulsion engine(s)" (ICAO, 2017, p. I-1-2).

Some of the results in this paper do not align with the researchers' intuitions. For example, the mean and median taxi-in and taxi-out time at airports with 0 hot spots are longer than those at airports with 1, 2, 3, and 4 hot spots; the mean and median taxi-in and taxi-out time at airports with four hot spots are longer than those at airports with five hot spots. This could be a result of the methods used to determine the hot spots by the FAA. For example, as one of the biggest and busiest airports in the world, the Dallas Fort Worth International Airport (DFW) has 0 hot spots on the airport diagram. This could be a reason that the taxi time means and medians at an airport with 0 hot spots are longer than airports with 1, 2, 3, and 4 hotspots. During this research, no literature or documentation was found regarding the methodology to determine the hot spots in airports other than the occurrence of incursions or collisions. Therefore, the researchers are not able to assess the feasibility of using the number of hot spots as an indicator of airport configuration complexity and number of intersections.

It is commonly known that taxi-out time is generally longer than taxi-in time due to queueing for take-off and to traffic managers' priority used for takeoffs versus landings. The objective of this research is to explore if taxi out is longer than taxi in when hub classification and the number of hot spots are taken into consideration. The results of this paper show that taxi-out time is longer than taxi-in time regardless of airport hub classification and number of hot spots. The results suggest that despite the number of enplanements and the number of complex intersections, taxi-out time is still longer than taxi-in time at hub airports.

Conclusion

In this research, the taxi-out times were compared to the taxi-in times for each of the three airport hub classifications (Small/Medium/Large), each of the six numbers of hot spots (0, 1, 2, 3, 4, 5), and for each combination of hub classification and number of hot spots. The researchers collected taxi time data from the FAA ASPM dataset, the hub classification information from the FAA NPIAS report, and the hot spot data from the FAA airport diagrams. The researchers selected 11 airports with the highest numbers of total operations in each hub classification. The taxi time for each of the 20 days with the highest number of daily operations in the timeframe were selected for each of the 33 sampled airports as the sample data. The data collected are in pairs, and the zeros were removed from the data, meaning that for each quarter hour, both a taxi-out and a taxi-in time had to be reported.

Consistent with the experience of those familiar with airport operations, the mean and median airport taxi-out time was found to be significantly longer than the mean and median taxi-in time in the combined sample (without considering hub classification or hot spots). When analyzed separately, the mean and median taxi-out time was found to be significantly longer than the mean and median taxi-in time for each of the three airport hub classifications (S/M/L). For each of the six numbers of hot spots (0, 1, 2, 3, 4, 5) at airports, the mean and median taxi-out time was found to be significantly longer than the mean and median taxi-in time. When the samples were divided by both hub classifications and number of hot spots, the results showed that mean and median taxi out time are significantly longer than taxi in time for each combination of hub classification and number of hot spots. By observing individual airport data for the number of hotspots at different hub sizes, the researchers noticed that for some airports in the same category, the number of hotspots might vary widely, e.g., large hub airports such as DFW (0 hotspots) and LAX (5 hotspots).

Future work may reduce the target sample to focus on specific airports with currently constrained capacity or that are predicted to be congested in the future. In this paper, average quarter-hour taxi time were collected from the ASPM dataset. Other taxi time data sources (such as Airport Surface Detection Equipment - Model X) that provide more specific taxi information, including routes, distances, and taxi time of each operation in minutes, may be used to improve the statistical power of the tests. Future research could also consider other factors that might affect taxi time, such as weather conditions or airport configuration, among others, or other methodologies, such as simulation analysis or Bayesian analyses.

References

Atkin, J., Burke, E. K., & Greenwood, J. S. (2010). TSAT allocation at London Heathrow: the relationship between slot compliance, throughput and equity. *Public Transport*. https://doi.org/10.1007/s12469-010-0029-2

Atkin, J., Burke, E. K., & Greenwood, J. S. (2011). A comparison of two methods for reducing take-off delay at London Heathrow airport. *Journal of Scheduling*, *14*(5), 409–421. https://doi.org/10.1007/s10951-011-0228-y

Balakrishna, P., Ganesan, R., & Sherry, L. (2010). Accuracy of reinforcement learning algorithms for predicting aircraft taxi-out times: A case-study of Tampa Bay departures. *Transportation Research Part C: Emerging Technologies*, *18*(6), 950–962. https://doi.org/10.1016/j.trc.2010.03.003

Burgain, P. (2010). *On the control of airport departure operations* [Doctoral Dissertation, Georgia Institute of Technology]. http://hdl.handle.net/1853/37261

Clewlow, R. R. L., Simaiakis, I., & Balakrishnan, H. (2010). Impact of arrivals on departure taxi operations at airports. *AIAA Guidance, Navigation, and Control Conference*. https://doi.org/10.2514/6.2010-7698

Dissanayaka, M., Ryley, T., Spasojević, B., & Caldera, S. (2023). Evaluating methods that calculate aircraft emission impacts on air quality: A Systematic Literature review. *Sustainability*, *15*(12), 9741. https://doi.org/10.3390/su15129741

Federal Aviation Administration. (2022, March 16). *Hot Spot Standardized Symbology.* U.S. Department of Transportation, Washington, D.C. Retrieved September 24, 2023, from https://www.faa.gov/newsroom/hot-spot-standardized-symbology

Federal Aviation Administration. (2022, May 17). *FAA Airport Diagrams.* U.S. Department of Transportation, Washington, D.C. Retrieved September 24, 2023, from https://www.faa.gov/airports/runway_safety/diagrams/

Federal Aviation Administration. (2022, October 06). *National Plan of Integrated Airport Systems (NPIAS) 2023–2027: Appendix A - List of NPIAS Airports.* U.S. Department of Transportation, Washington, D.C. Retrieved September 24, 2023, from https://www.faa.gov/airports/planning_capacity/npias/current/2023_NPIAS_Appendix_A

Federal Aviation Administration. (2023, May 08). *FAA Aerospace Forecast FY 2023 - 2024.*

U.S. Department of Transportation, Washington, D.C. Retrieved October 03, 2023, from https://www.faa.gov/data_research/aviation

Federal Aviation Administration. (n.d.-a). *Aviation System Performance Metrics (ASPM).* U.S. Department of Transportation, Washington, D.C. Retrieved September 24, 2023, from https://aspm.faa.gov/aspmhelp/index/Aviation_System_Performance_Metrics_(ASPM).html#Definitions_of_Variables

Federal Aviation Administration. (n.d.-b). *ASPM Taxi Times: Definitions of Variables.* U.S. Department of Transportation, Washington, D.C. Retrieved September 24, 2023, from https://aspm.faa.gov/aspmhelp/index/ASPM_Taxi_Times__Definitions_of_Variables.html

Federal Aviation Administration. (n.d.-c). *Cost of Delay Estimate.* U.S. Department of Transportation, Washington, D.C. Retrieved September 24, 2023, from https://www.faa.gov/sites/faa.gov/files/data_research/aviation_data_statistics/cost_delay_estimates.pdf

Federal Aviation Administration. (n.d.-d). *Airport Categories.* U.S. Department of Transportation, Washington, D.C. Retrieved September 24, 2023, from https://www.faa.gov/airports/planning_capacity/categories

Federal Aviation Administration. (n.d.-e). *Aviation System Performance Metrics: Airport Analysis.* U.S. Department of Transportation, Washington, D.C. Retrieved September 24, 2023, from https://aspm.faa.gov/apm/sys/AnalysisAP.asp

Federal Aviation Administration. (n.d.-f). *Aviation System Performance Metrics: Airport Taxi Times.* U.S. Department of Transportation, Washington, D.C. Retrieved September 24, 2023, from https://aspm.faa.gov/apm/sys/TaxiTimes.asp

Gupta, S., Johnson, M. E., & Wang, J. (2023, June). An Investigation of the Effect of Number of Hot Spots on Taxi Time at U.S. Hub Airports. *2023 ASEE Annual Conference & Exposition*. https://peer.asee.org/42635

International Civil Aviation Organization. (2017, July). *International Standards and Recommended Practices, Annex 16, Environmental Protection, Volume II — Aircraft Engine Emissions Fourth Edition*. Retrieved December 08, 2023, from https://ffac.ch/wp-content/uploads/2020/10/ICAO-Annex-16-Environmental-protection-Vol-II-Aircraft-Engine-Emissions.pdf

Laskey, K. B., Xu, N., & Chen, C. (2006). Propagation of delays in the national airspace system. *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence*. https://doi.org/10.48550/arXiv.1206.6859

Lee, H. (2015). Taxi time prediction at Charlotte Airport using Fast-Time simulation and machine learning techniques. *15th AIAA Aviation Technology, Integration, and Operations Conference*. https://doi.org/10.2514/6.2015-2272

Lee, H., Malik, W., & Jung, Y. (2016). Taxi-Out time prediction for departures at Charlotte Airport using machine learning techniques. *16th AIAA Aviation Technology, Integration, and Operations Conference*. https://doi.org/10.2514/6.2016-3910

Lordan, O., Sallán, J. M., & Valenzuela-Arroyo, M. (2016). Forecasting of taxi times: The case of Barcelona-El Prat airport. *Journal of Air Transport Management*, *56*, 118–122. https://doi.org/10.1016/j.jairtraman.2016.04.015

Mirmohammadsadeghi, N., Hu, J., & Trani, A. (2019). Enhancements to the runway capacity simulation model using the ASDE-X data for estimating airports throughput under various wake separation systems. *AIAA Aviation 2019 Forum*. https://doi.org/10.2514/6.2019-3044

Murner, R. (2012, February). Future fuel-efficient gas turbine jet engines run hotter, make component testing more critical. *Mechanical Engineering-CIME*, *134*(2), 50. https://link.gale.com/apps/doc/A279378130/AONE?u=anon~ca5c5a4d&sid=googleScholar&xid=d0eef22b

Nikoleris, T., Gupta, G., & Kistler, M. (2011). Detailed estimation of fuel consumption and emissions during aircraft taxi operations at Dallas/Fort Worth International Airport. *Transportation Research Part D: Transport and Environment*, *16*(4), 302–308. https://doi.org/10.1016/j.trd.2011.01.007

Park, D. K., & Kim, J. K. (2023). Influential factors to aircraft taxi time in airport. *Journal of Air Transport Management*, *106*, 102321. https://doi.org/10.1016/j.jairtraman.2022.102321

Ravizza, S., Atkin, J., Maathuis, M. H., & Burke, E. K. (2013). A combined statistical approach and ground movement model for improving taxi time estimations at airports. *Journal of the Operational Research Society*, *64*(9), 1347–1360. https://doi.org/10.1057/jors.2012.123

Srivastava, A. (2011). Improving departure taxi time predictions using ASDE-X surveillance data. *2011 IEEE/AIAA 30th Digital Avionics Systems Conference*. https://doi.org/10.1109/dasc.2011.6095989

Wang, X., Brownlee, A. E. I., Woodward, J. R., Weiszer, M., Mahfouf, M., & Chen, J. (2021). Aircraft taxi time prediction: Feature importance and their implications. *Transportation Research Part C: Emerging Technologies*, *124*, 102892. https://doi.org/10.1016/j.trc.2020.102892