Ranking Volatility in Building Energy Consumption using Ensemble Learning and Information Entropy

Kunal Sharma[[1]](#footnote-1)

*AE4699, Summer 2019*

**Energy waste is a common problem that is faced by just about any institution. Certain buildings waste a lot of energy but energy management is often understaffed and not equipped with a system that can analyze every data point regarding energy consumption coming in from the sensors. We will explore a method to first classify a buildings behavior into one of 5 common types. These types are defined by their expected behavior and functionality in the organization or institution. This building behavior is classified for a 24 hour period. The expectation is that building behavior should not change over time. If it does then this behavior is highly suspicious and should be brought to the attention of energy management. Many campuses have multiple buildings so another important feature is to be able to not only measure the volatility of a buildings energy consumption behavior but also be able to rank the buildings so that management can attend to the buildings in the order of priority commensurate to the building’s volatility. Even with limited time and resources, energy management could both detect suspiciously behaving buildings and maximize their impact by inspecting the buildings in the proportionate order. The objective is to reduce the overall energy consumption, detect buildings with energy consumption issues, and maximize the impact of the energy management. Also this helps making the entire campus more sustainable.**

# Nomenclature

*<add definitions if needed>*

# Introduction

Talk about:

What the problem is

Who is facing it

How much does it cost?

Why is it happening

Why isn’t it currently solved

The annual air traffic passenger demand keeps growing in the past few decades. The International Air Transport Association (IATA) predicted that the number could double to 8.2 billion in 2037 [1]. While air travel largely speeds up the modern transportation, it has one intrinsic problem, which is delay. The Federal Aviation Administration (FAA) defined a flight to be delayed when it is 15 minutes later than its scheduled time [2]. In 2018, only 79.6% of flights arrived on-time in US [3].

Delay has many negative impacts, and it costs billions of dollars each year. The economic impacts of airline delays are mainly reflected in three aspects: airline costs, passenger costs, and loss of demand [4]. Since most crew are paid based on the scheduled block time, schedule buffer or flight against schedule brings airlines extra crew costs. From the passenger perspective, delays disrupt passengers’ schedules and may bring extra miscellaneous costs. Because of this, for short and mid-range trips, passengers may choose other transportation methods like road or rail traveling, which have less uncertainty in travel duration.

To track delays, since 1987, the Federal government requires carriers that account for 1 percent of the domestic scheduled passenger revenue to file reports for the on-time performance. The percentage was modified to 0.5 percent in 2018 [5]. Since June 2003, the airlines that report on-time data also report the causes of delays and cancellations to the Bureau of Transportation Statistic (BTS) [6]. The reported causes of delay include: air carrier delay, extreme weather delay, national aviation system (NAS) delay, late-arriving aircraft delay, and security delay. However, carriers always would like to have better statistic. So in response to the regulation and to improve their on-time performance statistic, carriers increase their scheduled gate-to-gate times, and established longer scheduled flight times compared to those required under optimal conditions [7]. This practice is known as “schedule padding”.

Delay is a common problem in air transportation. However, different from what is expected, when looking into the general delay data as shown in Figure 1 [3], the annual average departure delay percentage seems to be quite stable in the past decade, which seems to be quite different from the feelings of general public.

Figure 1: Average Departure Delay

Indeed the percentage data is aggregated without showing the detail information of the on-time performance. If we extract small pieces of information, we can see what the problem is. Table 1 shows the information of two flights took place on the same date of a year with the same origin-destination pair, operating carrier, and departure time. Morning flights are selected to minimize the influence of air traffic congestion. With the improvement of air traffic scheduling system and airplane technology, we expect the flight to be quicker as time passed. However, under the same conditions, the 2018 flight had a longer scheduled flight time.

Table 1: Flight Information Comparison

|  |  |  |
| --- | --- | --- |
| **Flight Date** | 1/1/2008 | 1/1/2018 |
| **Origin – Destination** | ATL-LGA | ATL-LGA |
| **Operating Carrier** | Delta | Delta |
| **Departure Time** | 6:45 AM. | 6:45 AM. |
| **Scheduled Flight Time** | 128 min | 133 min |

To prove that schedule padding exists, we use the 1988 data as reference since that is the first year carriers are required to report their on-time performance. In Figure 2, it can be observed that the peak frequency of delay in 1988 is zero, which means that most of the flights in 1998 did not experience delays. However, the peak shifts to negative in 2018, which means most of the flights in 2018 arrived earlier than the scheduled time. This comparison further proved that schedule padding exists.

Figure 2: January Departure Delay Distributions

Many models are setup to approach the delay problem. Over the years, engineers, statisticians, and even economists have been approaching the problem from different perspectives. However, with the existence of schedule padding, the delay data we currently have does not mean the actual “extra time” needed to perform a flight. The current delay data is blinding people’s eyes. In this case, it is necessary to use modeling schedule padding as an intermediate step to find out the “actual delay” data.

# Problem Formulation

## Literature Review

There are researches that have been done on both schedule padding and delay modeling. Past literatures reveal that the uncertainty involved in flight times is driven by both periodic and stochastic factors, and thus carriers schedule the block times based on their expectations of the performance [7]. Even for the same route, different carriers “pad” their schedules differently [8].

Delay modeling has been approached via many different methods: classical machine learning approaches, artificial neural networks (ANN), Markov Jump Linear System (MJLS), etc. [9]. However, when predicting delays in air traffic networks, the fluctuations in flight scheduled time blocks and air time are usually ignored.

## Problem Scoping

To fulfill the gap between the origin-destination pair study and nationwide overview of schedule padding research, we would like to look into the outbound routes of one specific airport. In this case, the choice of the origin airport is critical. It is expected that the origin airport to be large enough that has a variety of flight routes and relatively big sample size. Hartsfield-Jackson Atlanta International Airport (ATL) is chosen, since it is the world’s busiest airport by passenger traffic since 2000 and the primary hub of Delta Air Lines and some low-cost carriers [10]. In the meantime, it also serves as a connection hub in the US air traffic network. So in this problem, we will focus on the analysis of outbound routes of ATL airport.

## Problem Statement

In this problem, it is expected to model schedule padding with machine learning approaches and to obtain the actual delay data with schedule padding model.

# Approach

The data used in this research is obtained from *BTS* [3]. In this section, definitions of variables are stated. Calculations are performed with the assumptions listed.

## Variable Definitions

CRSDepTime/ CRSArrTime: Computerized Reservation System (CRS) departure/arrival time in local time

DepTime/ ArrTime: Actual departure/arrival time in local time

DepDelay/ArrDelay: Difference in minutes between scheduled and actual departure time; early departures show negative numbers

CRSElapsedTime: CRS elapsed time of flight in minutes

AirTime: Flight time in minutes

## Variable Calculations

*Schedule Padding = CRSElapsedTime – AirTime*

Assumption: AirTime roughly equals to the flight time under the optimal condition without consideration of extra loitering time due to various reasons. This value might be negative.

*Corrected Departure Delay = DepDelay + Schedule Padding*

Assumption: Corrected departure delay is considered as the absolute “extra time” it takes for the carrier to operate a flight. Reference value of 1987 is needed to perform comparison, since it is the first year on-time performance data

## Regression Models

To understand how different factors affect the schedule paddings, we need to setup a high-fidelity regression model. In our case, while the output variable (schedule padding) is numerical/ continuous, various input variables like day of the week or destination airport can be categorical. ANOVA (Analysis of Variance) is chosen to be the technique we would like to use for testing whether different input categories have significantly different values for output variables [11]. When using ANOVA, it is assumed that the outputs in regression are normally distributed and independent. After ANOVA, to further confirm where the differences occur between groups, the post hoc test is needed to be run. One common and popular post-hoc method of ANOVA is Tukey’s test [12]. Tukey’s test can provide us deeper insights into patterns between specific groups.

# Analysis

## Preliminary Analysis

Air traffic system is complicated. Before modeling schedule paddings, preliminary analysis is performed to identify the factors that affect schedule paddings.

The first factor we look into is the destination. As the busiest airport, ATL airport serves around 150 U.S destinations [10]. If the destination airports are ranked by enplanement, the top 5 airports are: MCO (Orlando, FL), FLL (Fort Lauderdale, FL), LGA (New York, NY), LAX (Los Angeles, CA), and TPA (Tampa, FL). The difference in schedule padding between destinations is obvious. In these five destination airports, the maximum schedule padding happens at LAX which has an average of 43.46 minutes, and the minimum is 23.09 minutes at TPA. Both the distance between origin and destination, and the congestion level of the destination airport affect the length of schedule padding.

Besides the destination, the carrier (or airline) also plays an important role. As mentioned above, Delta Air Lines (DL) has its hub at ATL airport, and has the greatest market share. The other carriers, that has more than 5% of carrier share if ranked by number of operations, are Southwestern Airlines (WN), American Airlines (AA), and United Airlines (UA). While different airlines may have different marketing strategies and schedule their flight routes differently, it is still surprised to see that American Airlines has an average of more than 10 minutes more schedule padding compared with Southwestern.

Other factors are also affecting schedule padding. Morning flights usually experience less delay, so the local departure time needs to be taken into consideration. Air traffic congestion is also a seasonal issue. Operation month of the year and day of the week are also two important factors. Table 2 summarizes the factors we are going to consider when modeling schedule padding and whether they are numerical or categorical.

Table 2: Regression Model Factors

|  |  |
| --- | --- |
| **Factor** | **Data Type** |
| Month | Categorical |
| Day of Week | Categorical |
| Operation Carrier | Categorical |
| Destination Airport | Categorical |
| Distance | Numerical |
| Departure Time | Numerical |

## Schedule Padding Model

The regression model used first is the classical linear regression model. The linear regression model summary is attached in Appendix A. Since several of the factors are categorical, ANOVA is conducted to test whether different input categories have a significant impact on the output variable. In ANOVA, interactions are considered. Based on the nature of the factors, month of year, day of week, and departure time are considered as independent variables. However, in ANOVA the interaction between categorical variable and numerical variable is not feasible to be reflected, and the interaction between the operation carrier and the destination airport has little impact on the result. A summary of ANOVA statistical result is given in Table 3. It can be noticed that the Pr(>F) is very small, which indicates that all six factor are statistically significant factors that will impact the mean schedule padding. To further examine the validity of the ANOVA test, the Residuals vs. Fitted plot and Normal Q-Q plot are created to check the homogeneity and normality assumptions [13]. The plots are attached in Appendix B. The outliers are those affecting the homogeneity and normality of variance. It is useful to remove them to increase the fidelity of the model.

Table 3: ANOVA Summary

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variables** | **Df** | **Sum Sq** | **Mean Sq** | **F value** | **Pr(>F)** |
| Month | 10 | 950820 | 95082 | 1868.62 | <2e-16 |
| Day of Week | 6 | 22538 | 3756 | 73.82 | <2e-16 |
| Operation Carrier | 16 | 1230823 | 76926 | 1511.82 | <2e-16 |
| Destination Airport | 163 | 1461090 | 11977 | 28714.43 | <2e-16 |
| Distance | 1 | 1952204 | 1461090 | 235.38 | <2e-16 |
| Departure Time | 1 | 67806 | 67806 | 1332.57 | <2e-16 |

As the ANOVA test is significant, Tukey HSD (Tukey Honest Significant Differences) is computed to perform pairwise-comparison between groups [13]. Appendix C has included the results of Tukey’s test. Figure 3 is a visualization of the result of Tukey’s test. The adjusted p value indicates the significance. Generally speaking, if the adjusted p-value is smaller than 0.05, the category is considered to have significant impact on the output variable. If most of the categories of a factor have large p-values, the factor is considered as less influential. In the figure, that the interval line is greatly off from the center indicates the pair of categories have contrast in mean schedule padding, and either or both of the categories are significant. By observing Tukey HSD, it can be found that month of the year has larger impact on the schedule padding than day of week. Among the 12 months, the impact of February, March, July, and September are more significant. For the operation carriers, major airlines like Delta and United Airlines, have larger impact on schedule padding than smaller airlines.

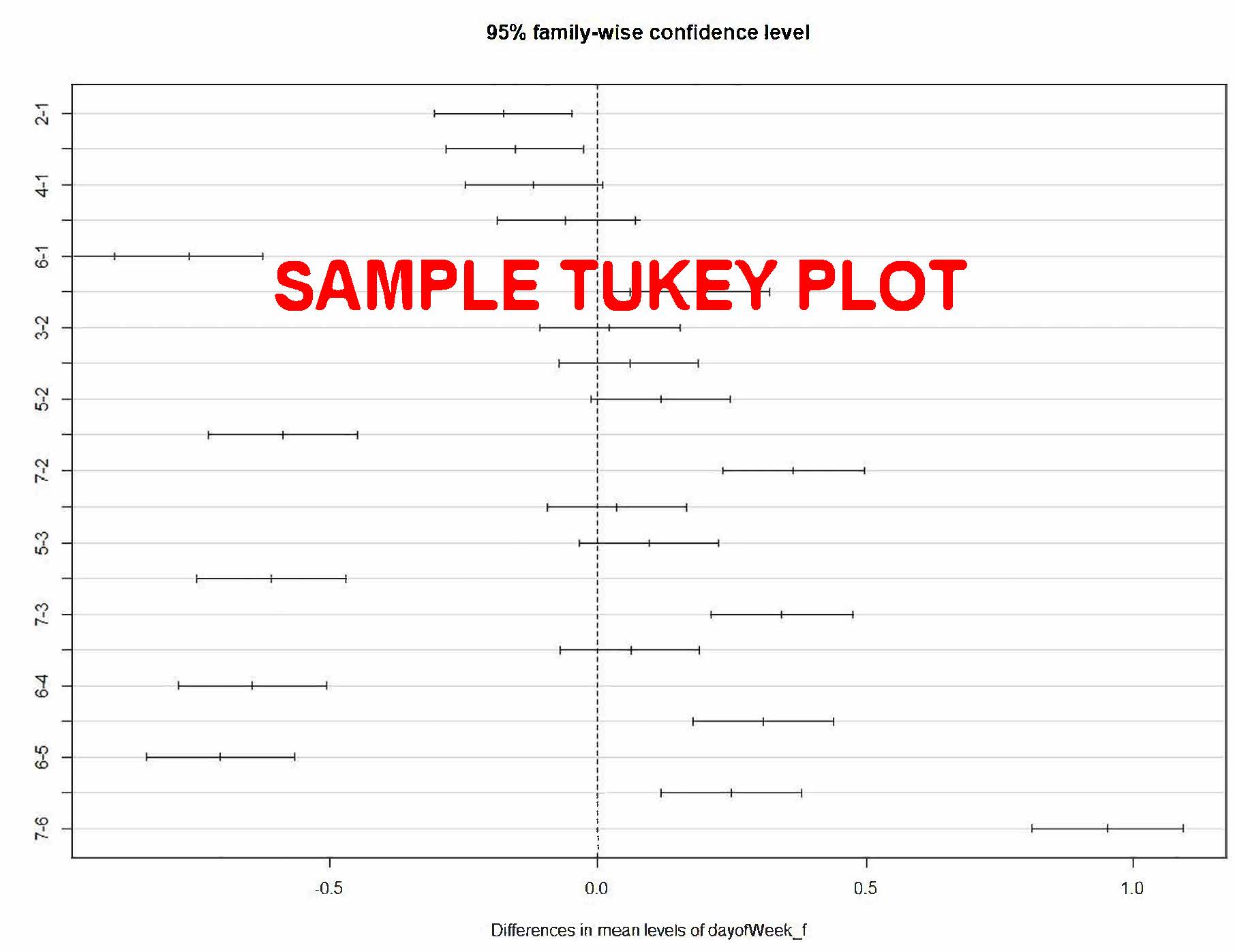


Figure : Sample Tukey's Plot

## Treated Delay Data

From the Tukey’s test, it can be learnt that the schedule padding varies from airlines to airlines. The boxplot shown in Figure 4 also reveals that different airlines have different schedule padding distributions. Indeed, different airlines can have different market focus, flight schedules, and operation aircrafts. Some airlines focus on operating regional flights, while others may have larger aircrafts operating mid-range flights; while major airlines have departures throughout the day, some low-cost airlines tend to have more morning flights. These are all factors may influence schedule padding and delay. However, on-time performance is still one of the most important elements people use to evaluate the airlines. It is no doubt that under the same conditions, people tend to book the flight that has less possibility get delayed.

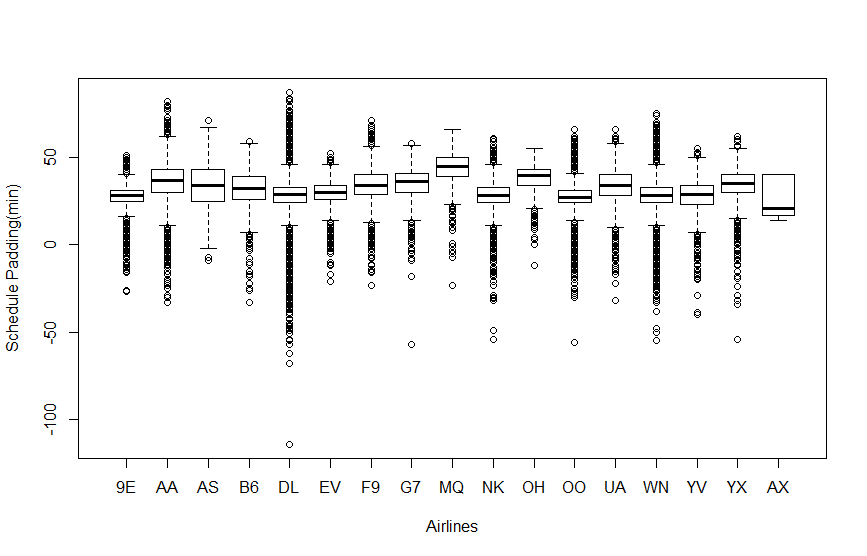


Figure : Schedule Padding Boxplot of Airlines

Table 4 shows how the departure delay percentage and its corresponding rank changes after schedule padding is taken into consideration. For the treated departure percentage, we use 1995 average schedule padding as the reference, since this is the first year actual air time was reported. The treated delay percentage is calculated by adding the additional schedule padding time in addition to the 1995 reference. It can be observed that in major airlines, Delta, United, and Southwest airlines all have an increase in rank, which means they have relatively short schedule paddings. In contrast, American Airlines has an obvious decrease in rank, which means it may take passengers longer time to travel with American Airlines for a similar route.

Table : Airline On-Time Performance

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Airline** | **IATA Airline Code** | **Treated Dep. Delay%** | **Dep. Delay%** | **Treated Rank** | **Original Rank** | **Change in Rank** |
| Endeavor | 9E | 20.75% | 14.47% | 1 | 3 | -2 |
| SkyWest | OO | 24.60% | 17.68% | 2 | 8 | -6 |
| Mesa | YV | 26.99% | 13.95% | 3 | 3 | 0 |
| Spirit | NK | 29.41% | 17.15% | 4 | 7 | -3 |
| Delta | DL | 32.28% | 16.03% | 5 | 6 | -2 |
| ExpressJet | EV | 32.59% | 22.05% | 6 | 11 | -5 |
| Republic | YX | 37.89% | 14.49% | 7 | 4 | 3 |
| United | UA | 41.81% | 21.88% | 8 | 10 | -2 |
| Southwest | WN | 41.83% | 25.08% | 9 | 13 | -4 |
| Alaska | AS | 42.72% | 10.88% | 10 | 1 | 9 |
| American | AA | 48.56% | 14.69% | 11 | 5 | 6 |
| GoJet | G7 | 48.93% | 21.00% | 12 | 9 | 3 |
| Frontier | F9 | 48.95% | 25.92% | 13 | 15 | -2 |
| JetBlue | B6 | 52.18% | 30.84% | 14 | 16 | -2 |
| Trans States | AX | 60.00% | 40.00% | 15 | 17 | -2 |
| PSA | OH | 66.73% | 24.95% | 16 | 12 | 4 |
| American Eagle | MQ | 79.65% | 25.26% | 17 | 14 | 3 |

# Conclusion

In this special problem, we have proved that schedule padding exists and has much influence on delay data. To model delay more accurately, regression model of schedule padding is built as an intermediate step. ANOVA test is conducted to see how the categorical variables affect schedule paddings. Together with Tukey’s test, we figured out that among the categorical variables, month of the year and operation carrier have larger influence toward the schedule padding. After treating delay data with schedule padding, it is noticeable that airlines “hide” real flight time and on-time performance in different degrees.

With the schedule padding model, the actual flight time of future flights can be predicted. The predictions can also be used to validate the model and the method. For future study, we would like to expand the sample size to the entire domestic routes, and improve the schedule padding model by including more factors and refining the regression algorithm.

# References

[1] “IATA Forecast Predicts 8.2 billion Air Travelers in 2037”, *IATA*, 24 Oct. 2018, Press Release No.:62, URL: <https://www.iata.org/pressroom/pr/Pages/2018-10-24-02.aspx> [retrieved 4 Mar. 2019]

[2] “OPSNET: Delays”, *FAA*, last modified on 3 Apr. 2018, URL: <https://aspmhelp.faa.gov/index.php/OPSNET:_Delays> [retrieved 4 Mar. 2019]

[3] “Airline On-Time Performance Data”, *Bureau of Transportation Statistics*, URL: <https://www.transtats.bts.gov/tables.asp?DB_ID=120>

[4] Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., Peterson, E., Sherry, L., Trani, A., Zou, B., “Total Delay Impact Study: A Comprehensive Assessment of the Costs and Impacts of Flight Delay in the United States”, *Nextor*, Oct. 2010

[5] “Code of Federal Regulations, Title 14 – Aeronautics and Space, Part 234 – Airline Service Quality Performance Reports”, 1 Jan. 2018, URL: <https://www.govinfo.gov/content/pkg/CFR-2018-title14-vol4/xml/CFR-2018-title14-vol4-part234.xml>

[6] “Understanding the Reporting of Causes of Flight Delays and Cancellations”, *Bureau of Transportation Statistics*, Updated: 26 Feb. 2018, URL: <https://www.bts.gov/topics/airlines-and-airports/understanding-reporting-causes-flight-delays-and-cancellations>

[7] Skaltsas, G., “Analysis of Airline Schedule Padding on U.S. Domestic Routes”, Jun. 2011 (Master Thesis)

[8] Burgauer, D., Peters, J., “Airline Flight Delays and Flight Schedule Padding: An Investigative Report”, Dec. 2000

[9] Gopalakrishnan, K., Balakrishnan, H., “A Comparative Analysis of Models for Predicting Delays in Air Traffic Networks”, Twelfth USA/Europe Air Traffic Management Research and Development Seminar, 2017

[10] Tharpe, J, "Atlanta Airport Still the "Busiest"". The Atlanta Journal-Constitution. Archived from the original on January 6, 2007

[11] “Chapter 6: Categorical data”, URL: <http://www.mit.edu/~6.s085/notes/lecture6.pdf>

[12] “Post-Hoc Analysis with Tukey’s Test”, URL: <https://rpubs.com/aaronsc32/post-hoc-analysis-tukey>

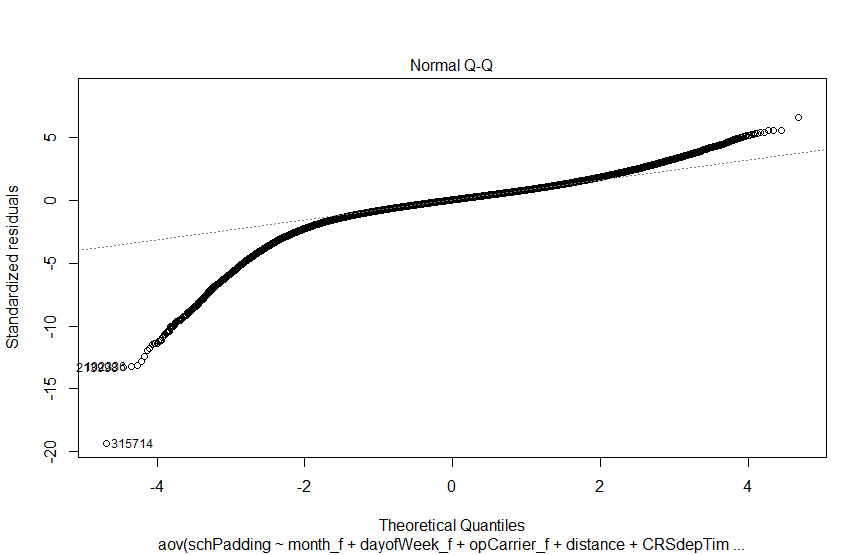
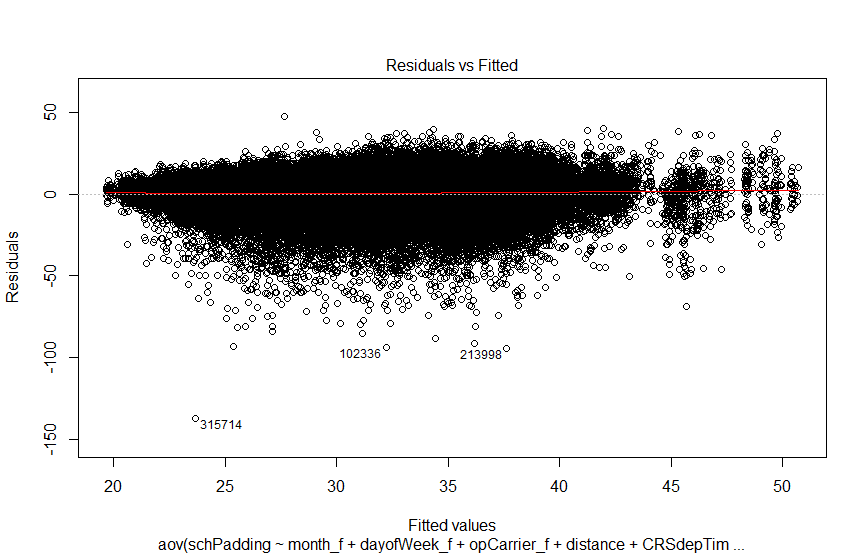
[13] “ANOVA Test in R”, URL: <http://www.sthda.com/english/wiki/two-way-anova-test-in-r>

# Appendix

1. **Linear Regression Summary**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Estimate** | **Std. Error** | **t value** | **Pr(>abs(t))** |
| (Intercept) | 72.9700 | 10.9600 | 6.66 | 2.75E-11 |
| month\_f2 | -0.3889 | 0.0594 | -6.546 | 5.91E-11 |
| month\_f3 | -2.7630 | 0.0569 | -48.544 | <2e-162e-16 |
| month\_f4 | -2.3560 | 0.0572 | -41.158 | <2e-162e-16 |
| month\_f5 | -2.3860 | 0.0568 | -42.021 | <2e-16 |
| month\_f6 | -0.0041 | 0.0569 | -0.072 | 0.94269 |
| month\_f7 | 0.7872 | 0.0565 | 13.936 | <2e-16 |
| month\_f8 | 0.0216 | 0.0564 | 0.384 | 0.70107 |
| month\_f9 | -3.1550 | 0.0578 | -54.554 | <2e-16 |
| month\_f10 | -4.3550 | 0.0571 | -76.338 | <2e-16 |
| month\_f11 | -2.4090 | 0.0579 | -41.588 | <2e-16 |
| dayofWeek\_f2 | -0.1585 | 0.0437 | -3.625 | 0.000289 |
| dayofWeek\_f3 | -0.1336 | 0.0438 | -3.052 | 0.002271 |
| dayofWeek\_f4 | -0.1155 | 0.0436 | -2.651 | 0.008024 |
| dayofWeek\_f5 | -0.0573 | 0.0435 | -1.317 | 0.187752 |
| dayofWeek\_f6 | -0.7776 | 0.0471 | -16.511 | <2e-16 |
| dayofWeek\_f7 | 0.1536 | 0.0445 | 3.453 | 0.000554 |
| opCarrier\_fAA | 2.4010 | 0.1116 | 21.513 | <2e-16 |
| opCarrier\_fAS | -2.7750 | 0.3258 | -8.517 | <2e-16 |
| opCarrier\_fB6 | -1.0350 | 0.1666 | -6.214 | 5.17E-10 |
| opCarrier\_fDL | -1.0040 | 0.0762 | -13.177 | <2e-16 |
| opCarrier\_fEV | 1.3030 | 0.1001 | 13.021 | <2e-16 |
| opCarrier\_fF9 | 1.6570 | 0.1600 | 10.356 | <2e-16 |
| opCarrier\_fG7 | -1.7570 | 0.2546 | -6.903 | 5.12E-12 |
| opCarrier\_fMQ | 7.0170 | 0.3193 | 21.974 | <2e-16 |
| opCarrier\_fNK | -2.7490 | 0.1167 | -23.554 | <2e-16 |
| opCarrier\_fOH | 3.2470 | 0.2516 | 12.905 | <2e-16 |
| opCarrier\_fOO | -1.6690 | 0.0775 | -21.526 | <2e-16 |
| opCarrier\_fUA | -1.1390 | 0.1430 | -7.961 | 1.71E-15 |
| opCarrier\_fWN | -1.8730 | 0.0873 | -21.45 | <2e-16 |
| opCarrier\_fYV | -2.4350 | 0.1962 | -12.408 | <2e-16 |
| opCarrier\_fYX | 2.5420 | 0.1376 | 18.469 | <2e-16 |
| opCarrier\_fAX | -10.0100 | 3.1930 | -3.136 | 0.001712 |
| distance | -0.0605 | 0.0158 | -3.835 | 0.000125 |
| CRSdepTime\_n | 0.0221 | 0.0007 | 31.905 | <2e-16 |
| dest\_fABQ | 37.1100 | 9.0530 | 4.099 | 4.15E-05 |
| dest\_fABY | -36.9100 | 8.6760 | -4.255 | 2.09E-05 |
| dest\_fAEX | -13.9900 | 3.0910 | -4.527 | 5.97E-06 |
| dest\_fAGS | -36.3100 | 8.7050 | -4.171 | 3.04E-05 |
| dest\_fALB | 10.1900 | 2.5120 | 4.057 | 4.97E-05 |
| dest\_fASE | 42.6200 | 9.6310 | 4.425 | 9.66E-06 |
| dest\_fATW | 5.5720 | 1.1640 | 4.786 | 1.70E-06 |
| dest\_fAUS | 7.3520 | 1.8790 | 3.913 | 9.13E-05 |
| dest\_fAVL | -36.2700 | 8.3740 | -4.331 | 1.48E-05 |
| dest\_fAVP | 0.4114 | 0.5526 | 0.744 | 0.456618 |
| dest\_fBDL | 8.8210 | 2.6000 | 3.392 | 0.000693 |
| dest\_fBHM | -39.3600 | 8.8460 | -4.449 | 8.63E-06 |
| dest\_fBMI | -12.2400 | 2.5750 | -4.755 | 1.99E-06 |
| dest\_fBNA | -30.2800 | 7.5860 | -3.992 | 6.56E-05 |
| dest\_fBOS | 19.0600 | 3.9640 | 4.807 | 1.53E-06 |
| dest\_fBQK | -29.6900 | 7.2120 | -4.117 | 3.84E-05 |
| dest\_fBTR | -17.0000 | 3.9030 | -4.356 | 1.33E-05 |
| dest\_fBTV | 17.5200 | 4.2160 | 4.156 | 3.24E-05 |
| dest\_fBUF | 0.4320 | 0.4033 | 1.071 | 0.284075 |
| dest\_fBWI | -6.6530 | 1.8780 | -3.544 | 0.000395 |
| dest\_fBZN | 59.9700 | 14.9200 | 4.02 | 5.81E-05 |
| dest\_fCAE | -33.4700 | 7.9330 | -4.219 | 2.46E-05 |
| dest\_fCAK | -11.3600 | 2.6530 | -4.282 | 1.85E-05 |
| dest\_fCHA | -38.2600 | 9.2880 | -4.12 | 3.79E-05 |
| dest\_fCHO | -14.9900 | 3.7640 | -3.982 | 6.83E-05 |
| dest\_fCHS | -26.6900 | 6.8770 | -3.88 | 0.000104 |
| dest\_fCID | 1.1060 | 0.3418 | 3.236 | 0.001212 |
| dest\_fCLE | -7.8850 | 2.2390 | -3.522 | 0.000428 |
| dest\_fCLT | -22.3200 | 7.3970 | -3.018 | 0.002545 |
| dest\_fCMH | -16.8900 | 3.9180 | -4.311 | 1.63E-05 |
| dest\_fCOS | 32.1100 | 7.7190 | 4.159 | 3.19E-05 |
| dest\_fCRW | -23.0100 | 5.2420 | -4.389 | 1.14E-05 |
| dest\_fCSG | -41.4700 | 9.6520 | -4.296 | 1.74E-05 |
| dest\_fCVG | -21.6200 | 5.0830 | -4.253 | 2.11E-05 |
| dest\_fDAB | -24.4800 | 5.1940 | -4.713 | 2.44E-06 |
| dest\_fDAL | -0.0002 | 0.4918 | 0 | 0.999735 |
| dest\_fDAY | -17.5400 | 4.1570 | -4.219 | 2.45E-05 |
| dest\_fDCA | -11.2700 | 2.3470 | -4.8 | 1.59E-06 |
| dest\_fDEN | 33.3100 | 7.9470 | 4.191 | 2.78E-05 |
| dest\_fDFW | 5.8460 | 0.6230 | 9.384 | <2e-16 |
| dest\_fDHN | -32.2500 | 8.2810 | -3.894 | 9.85E-05 |
| dest\_fDSM | 1.6860 | 0.8263 | 2.041 | 0.04128 |
| dest\_fDTW | -3.3300 | 1.6130 | -2.064 | 0.03898 |
| dest\_fECP | -31.7700 | 7.1770 | -4.426 | 9.60E-06 |
| dest\_fEGE | 35.2000 | 9.7560 | 3.608 | 0.000309 |
| dest\_fELM | -11.7100 | 7.1460 | -1.638 | 0.101344 |
| dest\_fELP | 41.2700 | 9.2580 | 4.458 | 8.27E-06 |
| dest\_fEVV | -23.1500 | 5.4480 | -4.25 | 2.14E-05 |
| dest\_fEWN | -19.6000 | 4.1470 | -4.725 | 2.30E-06 |
| dest\_fEWR | 7.3550 | 0.8447 | 8.707 | <2e-16 |
| dest\_fEYW | -8.3800 | 0.8277 | -10.124 | <2e-16 |
| dest\_fFAR | 25.6400 | 6.9960 | 3.665 | 0.000247 |
| dest\_fFAY | -24.8900 | 5.7460 | -4.332 | 1.48E-05 |
| dest\_fFLL | -8.4160 | 1.8150 | -4.638 | 3.52E-06 |
| dest\_fFNT | -4.5460 | 0.8561 | -5.31 | 1.10E-07 |
| dest\_fFSD | 17.3300 | 4.1110 | 4.216 | 2.49E-05 |
| dest\_fFSM | -10.5200 | 1.8670 | -5.637 | 1.74E-08 |
| dest\_fFWA | -8.8200 | 2.9650 | -2.975 | 0.002931 |
| dest\_fGNV | -25.7700 | 6.2330 | -4.134 | 3.57E-05 |
| dest\_fGPT | -24.8600 | 5.4150 | -4.591 | 4.41E-06 |
| dest\_fGRB | 3.2040 | 1.3080 | 2.45 | 0.014277 |
| dest\_fGRK | 8.1290 | 2.3450 | 3.467 | 0.000526 |
| dest\_fGRR | -4.2670 | 0.9177 | -4.65 | 3.33E-06 |
| dest\_fGSO | -26.4900 | 6.1370 | -4.316 | 1.59E-05 |
| dest\_fGSP | -36.3600 | 8.5470 | -4.254 | 2.10E-05 |
| dest\_fGTR | -30.0300 | 7.1640 | -4.192 | 2.77E-05 |
| dest\_fHDN | 38.1900 | 10.2000 | 3.745 | 0.000181 |
| dest\_fHNL | 247.2000 | 60.0000 | 4.12 | 3.79E-05 |
| dest\_fHOU | -0.5418 | 0.2674 | -2.026 | 0.042757 |
| dest\_fHPN | 2.9180 | 1.3700 | 2.131 | 0.033126 |
| dest\_fHSV | -38.1400 | 8.5790 | -4.445 | 8.77E-06 |
| dest\_fIAD | -10.5200 | 2.5520 | -4.124 | 3.72E-05 |
| dest\_fIAH | 0.4326 | 0.2823 | 1.532 | 0.125487 |
| dest\_fICT | 2.1160 | 1.4110 | 1.5 | 0.133649 |
| dest\_fILM | -21.8000 | 5.0220 | -4.34 | 1.42E-05 |
| dest\_fIND | -17.4700 | 4.1540 | -4.205 | 2.62E-05 |
| dest\_fJAC | 52.2000 | 13.8400 | 3.772 | 0.000162 |
| dest\_fJAN | -25.8500 | 5.5870 | -4.626 | 3.73E-06 |
| dest\_fJAX | -28.9500 | 6.7040 | -4.318 | 1.58E-05 |
| dest\_fJFK | 11.3200 | 1.0600 | 10.677 | <2e-16 |
| dest\_fLAS | 67.2200 | 16.5800 | 4.054 | 5.04E-05 |
| dest\_fLAX | 83.6600 | 19.7300 | 4.239 | 2.24E-05 |
| dest\_fLEX | -26.4900 | 6.1690 | -4.294 | 1.76E-05 |
| dest\_fLFT | -11.0800 | 3.0430 | -3.642 | 0.00027 |
| dest\_fLGA | 9.1530 | 1.0860 | 8.431 | <2e-16 |
| dest\_fLIT | -18.1700 | 3.8250 | -4.749 | 2.04E-06 |
| dest\_fLNK | 7.6220 | 2.3590 | 3.23 | 0.001236 |
| dest\_fMCI | -0.0423 | 0.2705 | -0.156 | 0.875805 |
| dest\_fMCO | -17.3800 | 4.5930 | -3.784 | 0.000154 |
| dest\_fMDT | -5.6860 | 1.2290 | -4.627 | 3.71E-06 |
| dest\_fMDW | -7.2010 | 1.6610 | -4.336 | 1.45E-05 |
| dest\_fMEM | -24.9200 | 5.7280 | -4.35 | 1.36E-05 |
| dest\_fMGM | -37.5400 | 8.6430 | -4.344 | 1.40E-05 |
| dest\_fMHT | 19.9500 | 4.0770 | 4.893 | 9.92E-07 |
| dest\_fMIA | -6.7930 | 1.6130 | -4.212 | 2.53E-05 |
| dest\_fMKE | -1.2200 | 0.4930 | -2.475 | 0.013326 |
| dest\_fMLB | -18.4300 | 3.9990 | -4.608 | 4.07E-06 |
| dest\_fMLI | -5.2990 | 1.0320 | -5.136 | 2.81E-07 |
| dest\_fMLU | -17.8300 | 3.9070 | -4.564 | 5.01E-06 |
| dest\_fMOB | -26.5200 | 6.2010 | -4.277 | 1.89E-05 |
| dest\_fMSN | -0.1702 | 0.3745 | -0.455 | 0.649452 |
| dest\_fMSP | 15.0800 | 3.3510 | 4.5 | 6.81E-06 |
| dest\_fMSY | -18.7700 | 4.2630 | -4.402 | 1.07E-05 |
| dest\_fMTJ | 45.2400 | 10.4300 | 4.336 | 1.45E-05 |
| dest\_fMYR | -26.5900 | 5.9660 | -4.456 | 8.34E-06 |
| dest\_fOAJ | -20.8100 | 4.6770 | -4.449 | 8.62E-06 |
| dest\_fOAK | 93.7500 | 22.6200 | 4.144 | 3.41E-05 |
| dest\_fOKC | 1.6160 | 1.0800 | 1.496 | 0.134602 |
| dest\_fOMA | 6.5380 | 2.0080 | 3.256 | 0.001131 |
| dest\_fORD | 1.5430 | 1.4260 | 1.082 | 0.2792 |
| dest\_fORF | -13.6700 | 2.8350 | -4.822 | 1.42E-06 |
| dest\_fPBI | -11.2100 | 2.3790 | -4.712 | 2.46E-06 |
| dest\_fPDX | 92.2800 | 23.2800 | 3.964 | 7.37E-05 |
| dest\_fPHF | -13.8000 | 2.9650 | -4.656 | 3.23E-06 |
| dest\_fPHL | -1.1110 | 0.5250 | -2.116 | 0.034386 |
| dest\_fPHX | 54.9500 | 14.0600 | 3.908 | 9.30E-05 |
| dest\_fPIA | -6.1990 | 2.0770 | -2.984 | 0.002844 |
| dest\_fPIT | -10.4500 | 2.6770 | -3.903 | 9.50E-05 |
| dest\_fPNS | -28.0700 | 6.6890 | -4.196 | 2.71E-05 |
| dest\_fPVD | 10.9700 | 3.2950 | 3.329 | 0.000871 |
| dest\_fPWM | 20.0000 | 5.2460 | 3.812 | 0.000138 |
| dest\_fRDU | -23.5500 | 5.3490 | -4.403 | 1.07E-05 |
| dest\_fRIC | -14.4700 | 3.3830 | -4.278 | 1.88E-05 |
| dest\_fRNO | 83.2800 | 20.5400 | 4.054 | 5.03E-05 |
| dest\_fROA | -19.6900 | 5.3360 | -3.689 | 0.000225 |
| dest\_fROC | 4.4960 | 0.9141 | 4.918 | 8.74E-07 |
| dest\_fRST | 8.8280 | 2.1820 | 4.046 | 5.21E-05 |
| dest\_fRSW | -13.9100 | 2.8500 | -4.882 | 1.05E-06 |
| dest\_fSAN | 75.6700 | 18.8700 | 4.011 | 6.05E-05 |
| dest\_fSAT | 12.3900 | 2.8340 | 4.37 | 1.24E-05 |
| dest\_fSAV | -30.4600 | 7.5860 | -4.015 | 5.95E-05 |
| dest\_fSBN | -3.6510 | 2.0580 | -1.774 | 0.076122 |
| dest\_fSDF | -23.7700 | 5.9010 | -4.027 | 5.64E-05 |
| dest\_fSEA | 96.3100 | 23.4400 | 4.109 | 3.97E-05 |
| dest\_fSFO | 95.7500 | 22.7600 | 4.207 | 2.59E-05 |
| dest\_fSGF | -7.6360 | 2.1030 | -3.63 | 0.000283 |
| dest\_fSHV | -10.9200 | 2.2880 | -4.771 | 1.84E-06 |
| dest\_fSJC | 89.8300 | 22.4000 | 4.011 | 6.06E-05 |
| dest\_fSJU | 50.4100 | 13.4300 | 3.753 | 0.000175 |
| dest\_fSLC | 56.5700 | 14.1100 | 4.01 | 6.08E-05 |
| dest\_fSMF | 88.4200 | 22.0200 | 4.016 | 5.93E-05 |
| dest\_fSNA | 78.6900 | 19.2900 | 4.079 | 4.53E-05 |
| dest\_fSRQ | -18.4400 | 3.9660 | -4.65 | 3.33E-06 |
| dest\_fSTL | -15.1700 | 3.3360 | -4.548 | 5.43E-06 |
| dest\_fSTT | 52.9300 | 14.2500 | 3.713 | 0.000205 |
| dest\_fSTX | 54.4500 | 14.8900 | 3.657 | 0.000255 |
| dest\_fSYR | 6.1260 | 1.5960 | 3.839 | 0.000123 |
| dest\_fTLH | -32.4100 | 7.4450 | -4.354 | 1.34E-05 |
| dest\_fTPA | -22.0800 | 4.5620 | -4.839 | 1.31E-06 |
| dest\_fTRI | -31.7300 | 7.3820 | -4.298 | 1.72E-05 |
| dest\_fTTN | 6.0350 | 0.5202 | 11.6 | <2e-16 |
| dest\_fTUL | -2.4260 | 0.4474 | -5.422 | 5.89E-08 |
| dest\_fTUS | 53.4200 | 13.3400 | 4.005 | 6.20E-05 |
| dest\_fTYS | -37.5500 | 8.5630 | -4.385 | 1.16E-05 |
| dest\_fVLD | -31.3300 | 7.6840 | -4.077 | 4.56E-05 |
| dest\_fVPS | -26.6600 | 7.0200 | -3.798 | 0.000146 |
| dest\_fXNA | -6.0130 | 1.6960 | -3.546 | 0.000391 |
| dest\_fACY | -4.4870 | 0.6310 | -7.11 | 1.16E-12 |
| dest\_fISP | 1.1460 | 1.8130 | 0.632 | 0.527317 |
| dest\_fANC | 183.7000 | 42.9100 | 4.282 | 1.86E-05 |
| dest\_fFCA | 63.0400 | 18.2000 | 3.465 | 0.000531 |
| dest\_fMSO | 71.8000 | 17.5300 | 4.096 | 4.20E-05 |
| dest\_fRAP | 33.6100 | 8.6090 | 3.904 | 9.47E-05 |

1. **ANOVA Test Validity Plots**



1. **Tukey’s Test Results**

1. Undergraduate, Areospace Design Lab, Georgia Institute of Technology. [↑](#footnote-ref-1)