Visualizing and Quantifying Flight Delay Trends in the US Airline Industry

Andy Eskenazi [*1] **Marek Homola** [*1] **Olivier Ng'weno Kigotho** [*1]

Abstract

The growth of the airline industry over the past decade has led to significant increases in flight delays, resulting in economic, environmental, and safety challenges. Thus, it is imperative to understand the cause of these delays, both from an individual flight and network-wide perspective, including how these delays depend on the day-of-week, airline, and aircraft type. To that end, this work developed an open-source set of tools to help airlines and policy makers visualize and quantify and, eventually, mitigate flight delays. To illustrate further applications of the open-source tools, this work examined delays associated with flight sequences, and investigated a thought experiment where a passenger takes an infinite random walk in the US airline network.

1. Introduction

The exponential growth that the air transportation sector has experienced over the past 10 years (8) has created significant capacity challenges for airports (10). In particular, flight congestion and delays have become prevalent, highlighting the disparity between passenger demand and infrastructure limitations. Thus, there is a pressing need for innovative solutions that optimize airport operations and enhance efficiency while maintaining passenger satisfaction.

The importance of addressing flight delays extends beyond passenger inconvenience, as they have far-reaching economic, environmental, and safety consequences. Economically, delays lead to financial losses, decreased productivity, and supply chain disruptions. For instance, in 2007, flight delays were estimated to be responsible for an economic loss of $32.9 billion in the United States alone (1). Environmentally, delays exacerbate the aviation sector's carbon footprint through increased fuel consumption and emissions (5; 11). From a safety perspective, delays can indirectly compromise operational safety by intensifying congestion and straining air traffic control capacity (12). Moreover, flight delays influence regulatory and policy considerations, as authorities strive to improve the efficiency and resilience of air transportation systems.

Recognizing the multifaceted impact of flight delays, collaboration among stakeholders is crucial to develop innovative solutions for effective delay mitigation. By leveraging advanced technologies, optimizing resources, and promoting data-driven decision-making, the aviation community can work towards minimizing the occurrence and severity of flight delays, benefiting both the consumers, airlines, and the environment alike. Thus, the goal of this study is to work towards this goal by creating a set of open-source, data-driven tools (available in the following online repository) that will enable both passengers, airlines and policy makers to visualize flight delays, specifically taking place within the U.S. Airline Industry. The reason for focusing in the U.S. Airline Industry is two fold: first, US civil aviation accounts for approximately a fifth of the world's total aviation activity, when measured by revenue passenger-kilometers (9); second, as will be further reviewed in the following section, US flight delay data, sourced from the Bureau of Transportation Statistics (BTS) (2; 3) is publicly-accessible, a quality that is necessary to achieve the present study's objective of developing a truly open-source tool.

Overall, through the provision of the open-source delay visualization tools, this work aims to first, enable an easier identification of delay root causes, allowing for the uncovering of unusual trends across different time scales; second, quantify network-wide delays for individual flights and sequences of flights within the U.S. commercial airspace, identifying bottlenecks and critical nodes; and third, examine a thought experiment using Markov chains to model and simulate the dynamics of flight delays, where a passenger takes an infinite random walk in the US airline network.

The remainder of this paper is structured as followed. **Section 2** introduces the data and employed methods to compute the various network delay matrices, as well as to construct the delay visualization tools; **Section 3** presents the results from implementing the methodology in code; and **Section 4** delivers this study's conclusions and future work recommendations.

[*]Equal contribution [1]Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA. Correspondence to: Andy Eskenazi <andyeske@mit.edu>.

2. Methods

The following section introduces the mathematical models supporting the flight delay visualization tools. The first subsection concerns the BTS flight data, while the following three sub-sections describe the equations to construct the various delay matrices, Markov chains, and Bayesian networks (to visualize delays of specific flight sequences or infinite random walks in the network).

2.1. Data Processing

As was alluded to in the introduction, the data utilized in this work was obtained from the Bureau of Transportation Statistics (BTS) (2). For any given month and year, (e.g., January 2023, as in this work), the BTS Marketing Carrier On-time Performance dataset contained detailed information on daily flight movements, operating airlines and tail number (i.e., aircraft registration number), origin and destination airport, passenger volumes, distance traveled, and departing and arrival delays, among other variables. While this dataset possessed sufficient information to calculate route-specific delays, from both the day-of-week and operating airline level of granularity, it did not possess enough data to examine delays from the operating aircraft level of granularity. However, because the dataset contained the aircraft tail number for every flight, it was possible to construct a mapping with another BTS database, the B-43 Inventory (3) (which stored information about 7,923 aircraft in the entire US fleet), which enabled determining the aircraft type operating each flight. This processed flight data (538,837 flights), for January 2023, can be found on this work's online repository). While there exists other works and platforms that have developed flight delay visualization and quantification tools to the day-of-week and airline level of granularity (6; 4), the present study would be the first to also examine the impact of aircraft type.

2.2. Constructing the Delay Matrices

Having completed the data processing, it was possible to start constructing the delay matrices, by analyzing every flight operated in January of 2023. This work utilized the term "delay matrices" to refer cumulatively to three matrices: D, P, and E, for the route-specific average delays, probability of delays, and expected delays, respectively. These three matrices have dim $= 339 \cdot 339$, given that flights operated between 339 airports in the US airline route network. The entries of each of these matrices can be defined via **Eq. (1)** through **Eq. (3)**:

$$\mathrm{D_{i,j}} = \frac{\sum_{q=1}^{k_{i,j}} d_{i,j,q}}{k_{i,j}} \tag{1}$$

where i, the rows of D, represents the route's origin airport, while j, the columns of D, correspond to the route's destination airport. Here, the variable $k_{i,j}$ denotes the total number of flights on the route that had a delay, while $d_{i,j,q}$ is the delay, in minutes, of the specific flight q. For clarification, this work considered a flight to be delayed for as long as $d_{i,j,q} > 0$. Next, for the P matrix, its entries can be defined via:

$$\mathrm{P_{i,j}} = \frac{\sum_{q=1}^{n_{i,j}} \begin{cases} 1, & \text{if } f_{i,j,q} \text{ delayed} \\ 0, & \text{if } f_{i,j,q} \text{ not delayed} \end{cases}}{n_{i,j}} \tag{2}$$

where $n_{i,j}$ represents the total number of flights f that operated between i and j. As can be seen, the numerator of $P_{i,j}$ is precisely $k_{i,j}$, so a simplified expression would be $P_{i,j} = k_{i,j}/n_{i,j}$. Finally, the E matrix, which contains the expected delay (in min) on each route, was simply computed as the piece-wise multiplication of the entries in D and P, as follows:

$$\mathrm{E_{i,j}} = \mathrm{D_{i,j}} \cdot \mathrm{P_{i,j}}. \tag{3}$$

To illustrate these equations, **Figure 1** computes the D, P and E matrices for a toy network consisting of New York (JFK), Miami (MIA) and Boston (BOS). By definition, these matrices will always have 0s along the diagonals, given that there do not exist flights from airport i to i.

Notice that D, P, and E need not be a symmetric matrix, given that the route delay performance from airport i to j could differ significantly than that from j back to i. Additionally, computing these matrices helps collapse the information of 538,837 flights (taking place during January of 2019) into datasets of $339 \cdot 339 = 114,921$ entries, representing a reduction of $\sim$79% in terms of data storage. However, these matrices are mostly sparse, and only have 5,450 non-zero elements, or 4.7% of the total entries. In essence, of the total of $339 \cdot 338/2 = 57,291$ possible routes in the US airline network, only $5,450/2 = 2,725$ are actually operated.

The equations above described how to construct D, P and E using information from all 538,837 flights. However, mentioned previously, it is possible to disaggregate the data into the day-of-week, airline and aircraft type level of granularity. To that end, **Eq. (4)** through **Eq. (6)** below describe the modifications to the above equations to compute these new matrices. As an example, for the airline-level granularity, it is possible to compute $\mathrm{D_a}$, $\mathrm{P_a}$ and $\mathrm{E_a}$ via:

$$\mathrm{D_{a,i,j}} = \frac{\sum_{q=1}^{k_{a,i,j}} d_{a,i,j,q}}{k_{a,i,j}} \tag{4}$$

$$\mathrm{P_{a,i,j}} = \frac{k_{a,i,j}}{n_{a,i,j}} \tag{5}$$

$$\mathrm{E_{i,j,a}} = \mathrm{D_{a,i,j}} \cdot \mathrm{P_{a,i,j}}. \tag{6}$$

where the index a represents the operating airline (15 possible options); $d_{a,i,j,q}$ is the delay of flight q on the route from

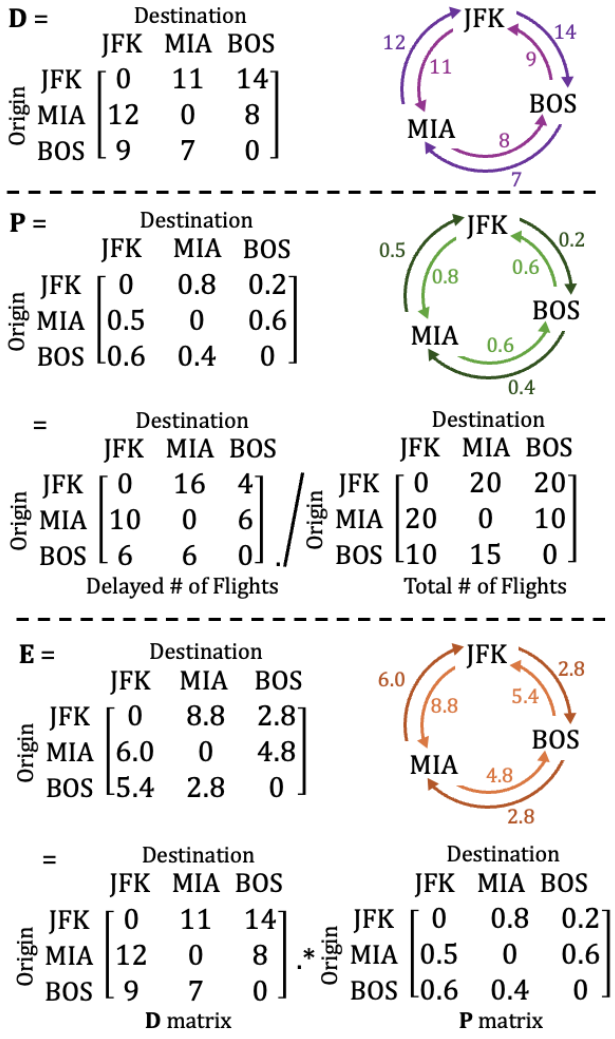


Figure 1. Construction of the D, P and E matrices, using the toy JFK - MIA - BOS network as an example.

i to j (operated by a); $k_{a,i,j}$ is the number of delayed flights on airline a for the route; and $n_{a,i,j}$ is the total number of flights that airline a operated on the route. This procedure can be identically extended to compute $\mathrm{D_w}$, $\mathrm{P_w}$ and $\mathrm{E_w}$, where w represents the day of the week (7 possible options), and $\mathrm{D_p}$, $\mathrm{P_p}$ and $\mathrm{E_p}$, where p denotes the aircraft operating the route (31 possible options). As a result, a total of $1 + 15 + 7 + 31 = 54$ sets of D, P and E matrices can be computed.

On the macroscopic scale, it is also possible to utilize the disaggregated D, P and E matrices to compute the network-wide expected delays and probability of delays for a specific airline, aircraft type, or day-of-week. For instance, the overall probability that an airline a would operate a delayed flight would be given by:

$$\mathrm{p_a} = \frac{\sum_{i=1}^{339}\sum_{j=1}^{339} k_{a,i,j}}{\sum_{i=1}^{339}\sum_{j=1}^{339} n_{a,i,j}} \tag{7}$$

where the numerator represents the addition of all the delayed flights across all routes i to j, while the numerator is simply the total number of flights, both operated by the airline. Similarly, the average delay for an airline a could be described by:

$$\mathrm{d_a} = \frac{\sum_{i=1}^{339}\sum_{j=1}^{339}\sum_{q=1}^{k_{a,i,j}} d_{a,i,j,q}}{\sum_{i=1}^{339}\sum_{j=1}^{339} k_{a,i,j}} \tag{8}$$

where the numerator denotes the addition of all the delays (in min) and the denominator represents the total number of delayed flights, both corresponding to airline a. Finally, the expected delay for airline a would be given by:

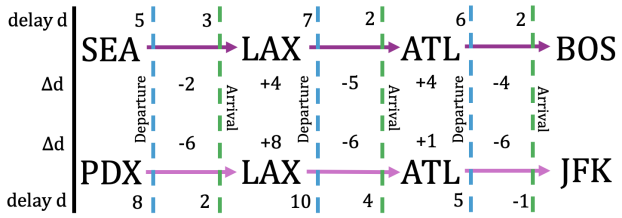
$$\mathrm{e_a} = \mathrm{p_a} \cdot \mathrm{d_a}. \tag{9}$$

Just like before, **Eq. (7)** through **Eq. (9)** can be extended to compute $\mathrm{p_w}$, $\mathrm{d_w}$, and $\mathrm{e_w}$ for every day-of-week w, and $\mathrm{p_p}$, $\mathrm{d_p}$, and $\mathrm{e_p}$ for every aircraft type p.

2.3. Flight Delay Propagation

One of the questions that this work sought to investigate was how delays propagate from one flight to the next. To answer this question, the present study built a Bayesian probabilistic graphical model (PGM) with nodes describing the aircraft in both the ground and flight state.

Using the BTS On-time Performance dataset (2), this work built an algorithm that followed each aircraft from one airport to another, recording the change in the delay while on the ground (Δd_{ground}) and during flight (Δd_{air}). To this end, this study first constructed a $1 \times n$ (with $n = 339$) vector g_{mean} of the mean change in flight delay Δd_{ground}, where each entry corresponded to a particular airport in the network. For example, if 20 aircraft arrived at an airport an average of 8 minutes late, and these same 20 aircraft left the airport an average of 10 minutes late, then the entry for this airport would be $10 - 8 = 2$ minutes. The next step consisted of building an $n \times n$ matrix F_{mean} of the change in delay during flight Δd_{air} with all the flights, where the columns corresponded to the origins and the rows were the destinations airports. Here, if 20 flights left from airport i to airport j an average of 9 minutes late and arrived an average of 7 minutes late, then the i, j^{th} entry on this matrix would be $7 - 9 = -2$ minutes. The last step saw the creation of the vector g_{std} and matrix F_{std}, of the same size and shape as g_{mean} and F_{mean}, respectively, with data recording the standard deviation of each of these datasets. This process is shown in **Figure 2**.



	LAX	LAX→ATL	ATL
Sequence 1	+4	-5	+4
Sequence 2	+8	-6	+1

Figure 2. Data processing used to generate the data for the change in delays Δd (consisting of both arrival and departure delays) for the example case of flights flying from LAX to ATL. As shown in the top half, flights might not be on the same sequence, but merely have overlapping sections. These overlapping sections were used to create the data entries as shown in the table in the bottom of the figure.

In addition, this work assumed these data to be normally distributed, a quality which would enable using a *t*-test to prove whether the deference between each entry and the mean delay was statistically significant. In general, a *t*-test uses the mean value of the sample data and the standard deviation of the sample data to estimate the probability (*p*-value) that the mean population data is the same from a particular value of interest. This work estimated the *p*-value of the mean Δd_{ground} of a particular airport relative to the mean Δd_{ground} across all airports. If the entry did not have a *p*-value of 0.05 or less in the *t*, then a mean value was used across all airports as the entry in g_{mean}. Indeed, this was the case for some regional airports with too few flights to draw statistically significant conclusions. Thus, it was possible to construct a normal distribution for each of the airports and flights using statistically significant evidence stored in g_{mean}, g_{std}, F_{mean}, and F_{std}. Overall, this enabled building a PGM for any given route sequence.

As an example, consider the PGM for a flight scheduled to leave from Boston (BOS), fly to New York (JKF), then fly to Miami (MIA). This flight plan can be modelled as as shown in **Figure 3**, where each of the nodes was modeled as a Gaussian distribution X as follows:

$$X_{k+1} = X_k + N(\mu, \sigma) \tag{10}$$

Here, μ and σ corresponded to the mean and standard deviation of the delays of a particular airport or flight.

2.4. Transition Probability Matrix P

Besides visualizing flight delays in individual routes and examining the complexities of delays on sequences of flights,

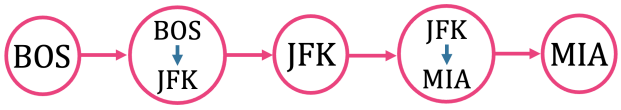


Figure 3. A series of flights for a particular aircraft was modeled with nodes representing both the flights (BOS→JFK, JFK→MIA) and the airports (BOS, JFK and MIA).

this study also sought to explore a thought experiment, in which a passenger embarked itself on an infinite random walk across the entire US airline route network. While this thought experiment is theoretical in nature, in practice it could be applicable to illustrate delay patterns in point-to-point airline route network structures, such as that of Southwest Airlines in the US (7). In essence, an aircraft in Southwest's route network typically starts the day at an airport i and may jump through 3 or 4 stations before ending the day at a different airport j, repeating the same pattern each day. As a result, this thought experiment sought to take this example to the extreme, trying to determine the probability of a flight being delayed between airports i and j, taking an infinite random walk (or number of flights) in between.

To this goal, a new matrix S was introduced, which represented the transition matrix from one state (current airport) to another (next airport). The entries of S are defined as:

$$\mathrm{S_{i,j}} = \frac{k_{i,j}}{\sum_{i=1}^{339} k_{i,j}} \tag{11}$$

Essentially, entry $\mathrm{S_{i,j}}$ divided the number of delayed flights on the route from i to j, over the total number of delayed flights arriving into airport j. Hence, this ratio represents the "delay weight" of each route, or the percentage, out of all the delayed flights arriving into airport j, that came from airport i. These ratios constitute a distribution where the addition of all the $\mathrm{S_{i,j}}$ over all i for a fixed j is equal to 1. By definition, all $\mathrm{S_{i,j}} \geqslant 0$, so it is possible to employ the following observation, which states that there exists a stationary distribution vector $\Pi^* \geqslant 0$ given by $\lim_{n\to\infty} S^n \cdot \Pi_o = \Pi^*$, regardless of the initial distribution $\Pi_o = [\Pi_{o,1} \dots \Pi_{o,339}]$. The distribution vector Π, whose entries add up to 1, can then be understood as well as representing the percentages, out of all the delayed flights arriving into some airport j, that came from airport i. To illustrate the construction of the transition probability matrix, **Figure 4** computes the S, $\mathrm{S^2}$, $\mathrm{S^3}$ and $\mathrm{S^*}$ matrices for the toy network consisting of JFK, MIA and BOS. To understand the i, j entries of the powers of S, it is helpful to reference **Figure A1** and **Table A1** in the Appendix, which show the adjacency matrix of the toy network and also calculates the 1,2 entry (row: JFK, column: MIA) of the S, $\mathrm{S^2}$, $\mathrm{S^3}$ matrices.

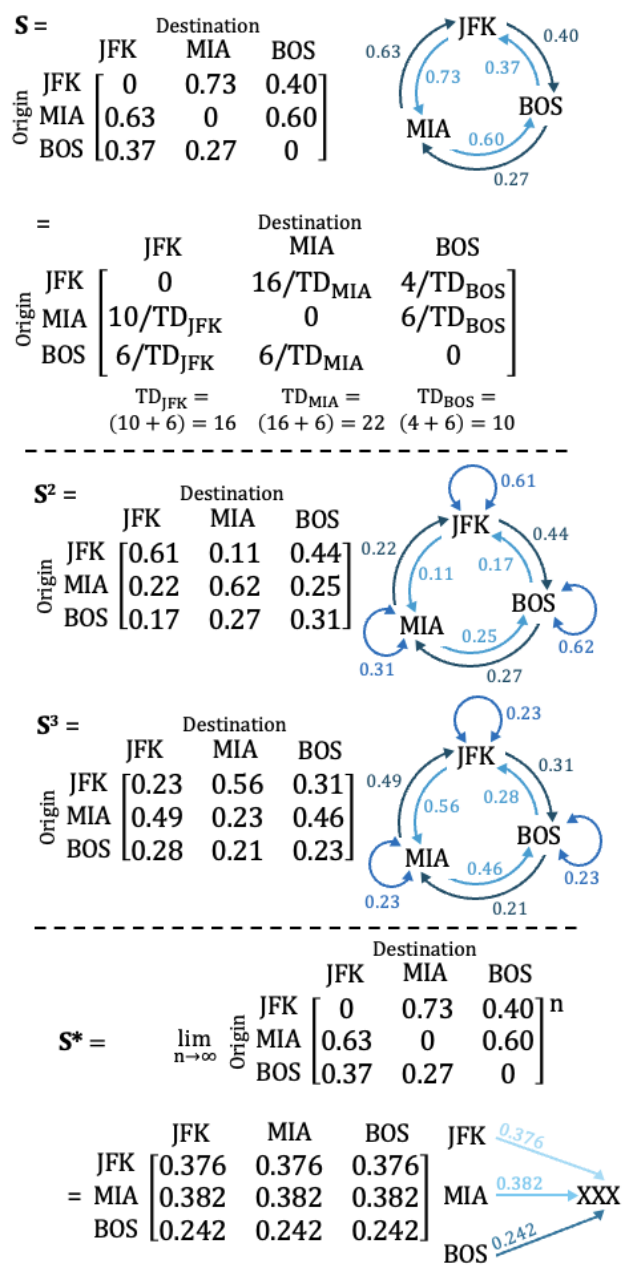


Figure 4. Construction of the S matrix, using the toy JFK - MIA - BOS network as an example.

The toy problem displayed in **Figure 4**, alongside the content in the Appendix, provide valuable insight into the implications of taking random walk in the route network. First, as the graphs demonstrate, both representing S and the network's adjacency matrix A, with higher powers, the diagonal entries of S and A (i.e., where origin i = destination j) are no longer zero, and have an associated probability of delay (i.e., "round-trip flights to nowhere"). Second, and perhaps most importantly, from examination, all columns C^* of S^* are identical, and thus $S^* \cdot \Pi_o = S^* \cdot [\Pi_{o,1} \dots \Pi_{o,339}] = \mathrm{C}^*$ (since $\sum_{g=1}^{339} \Pi_{o,g} = 0$). Equivalently, all rows in S^* have share the same entry, so if the rows represented the origin airports i and the columns the destination airport j, this result means that, in the infinite random walk, the outcome will be fully determined by the origin, not the destination. In the toy problem, all destinations "XXX" shared the same flight delays distribution, with 37.6% of all delayed flights coming from JFK, 38.2% from MIA, and 24.2% from BOS. Hence, the interpretation of this result is that, if a passenger arrived delayed to some random destination "XXX", chances are highest that it started its journey in JFK, followed by MIA, and then BOS.

3. Results and Discussion

The following section presents the results from implementing the mathematical models introduced in the previous section. The utilized code, functions and scripts are available on the open-source online repository.

3.1. Visualizing Individual Flight Delays

The first application of the mathematical framework (specifically, **Eq. (1)** through **Eq. (6)**) corresponded to the visualization of individual flight delay statistics, disaggregated to the day-of-week, airline, and aircraft level of granularity. Querying the various D, P and E matrices enabled the creation of **Figure 5**, which shows the expected arrival delays (in the blue bar columns) and probability of delays (in the orange lines) for the example San Francisco (SFO) to Honolulu (HNL) route.

Here, the top plot displays the delays (in min) on the route for a user-selected day of the week (e.g., Friday) and the average weekday, and compares it against the average of all flights arriving into the destination airport. The middle plot displays the delays (in min) for all airlines operating the route. Finally, the bottom plot conveys the same information, but segmenting the results by the aircraft that operated the routes. To compare and contextualize the data, each bar possesses a label that denotes the number of flights that were operated.

As the figure elucidates, there is large variability in all examined dimensions, i.e., day-of-week, airline, and aircraft type. For instance, flights taking place on a Friday (21 in the month, or 12.5% of the total) performed worse than the average day-of-week (168 in total), although the probability of delay increased; in its turn, the average arrival delays into Honolulu airport was the lowest, highlighting that this route (SFO - HNL) in particular was a bad performant out of all routes into HNL. On the airline front, Alaska had a significant better on-time performance than Hawaiian and

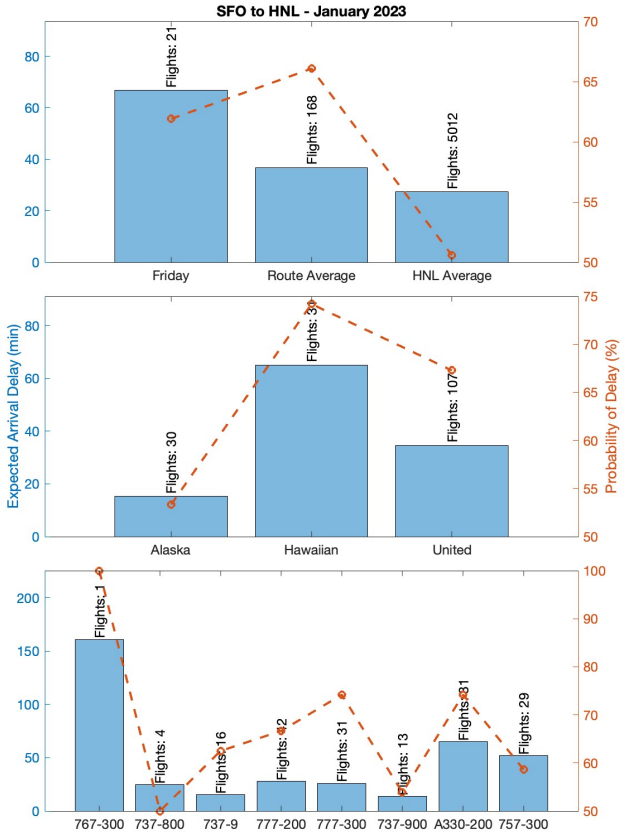


Figure 5. Route specific expected arrival delays (left y-axis) and probability of delays (right y-axis), using San Francisco (SFO) to Honolulu (HNL) as an example.

.

United, both of which had higher expected arrival delays and probabilities of delay; however, Alaska operated the least number of flights (30 in the month), being slightly outnumbered by Hawaiian (31) and dominated by United (107). Finally, the best performant aircraft was the 737-800, while the worst was the 767-300 (although notice that the number of flights in each case was not significant). As can be seen, these results offer a wealth of information, although it is critical to examine them in the context of the number of flights, given that naturally, with higher flight volumes, the probability of a delay can increase.

The second application of the mathematical framework (**Eq. (7)** through **Eq. (9)**) corresponded to the visualization of macroscopic variables, that is, the network-wide (i.e., all routes) expected delays and probability of delays for a specific airline and aircraft type. This resulted in **Figure 6** and **Figure 7**, respectively.

On the airline perspective, Frontier had the largest expected arrival delay (e_a), while Hawaiian instead had the highest

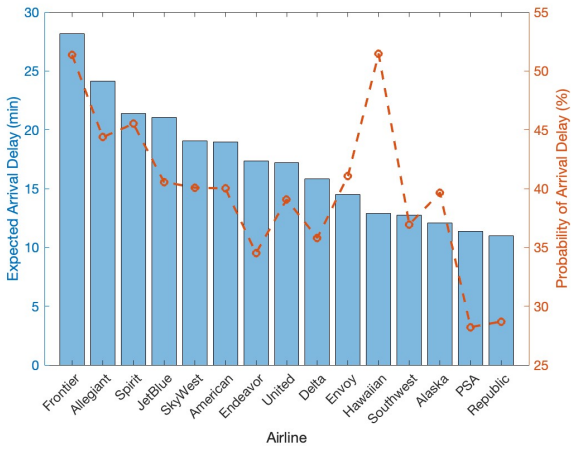


Figure 6. U.S. Airline Network expected arrival delays (left y-axis) and probability of delay (right y-axis), segmented by 15 airlines.

probability of arrival delay (p_a). In contrast, regional operators Republic and PSA each saw the lowest expected arrival delay and probabilities of delay, respectively. On the aircraft perspective, the worst performance was seen by the A350-900, which had both the largest e_p and p_p. On the opposite end of the spectrum, the 717-200 (operated only by Delta and Hawaiian) was the most reliable airliner.

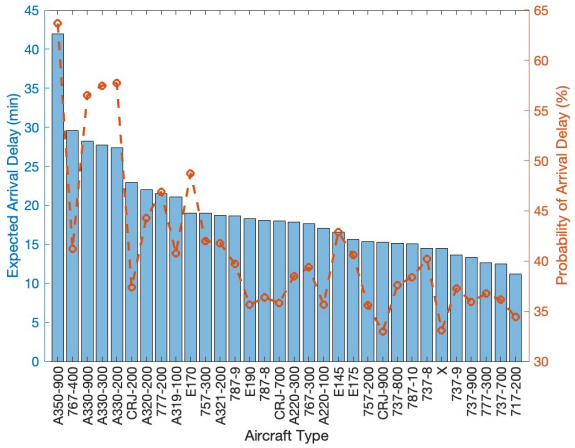


Figure 7. U.S. Airline Network expected arrival delays (left y-axis) and probability of delay (right y-axis), segmented by 31 aircraft types. Here, the aircraft type "X" refers to an uncategorized type.

3.2. Visualizing Delay Propagations

The third application of the mathematical framework was the creation of a Monte Carlo simulation to quantify how delays propagated through the network. Here, two cases were

considered, one where the initial delay at the beginning of the network was fixed, and the other where the initial delay was unknown and modeled using a normal distribution representing the mean and standard deviation of the departure delay from that given airport. As an example, the PGM of an aircraft flying from BOS to JFK then to MIA was considered, as shown in **Figure 8**. In the figure, it is possible to see that knowing the initial condition significantly reduced the uncertainty in predicting the delay arriving into MIA.

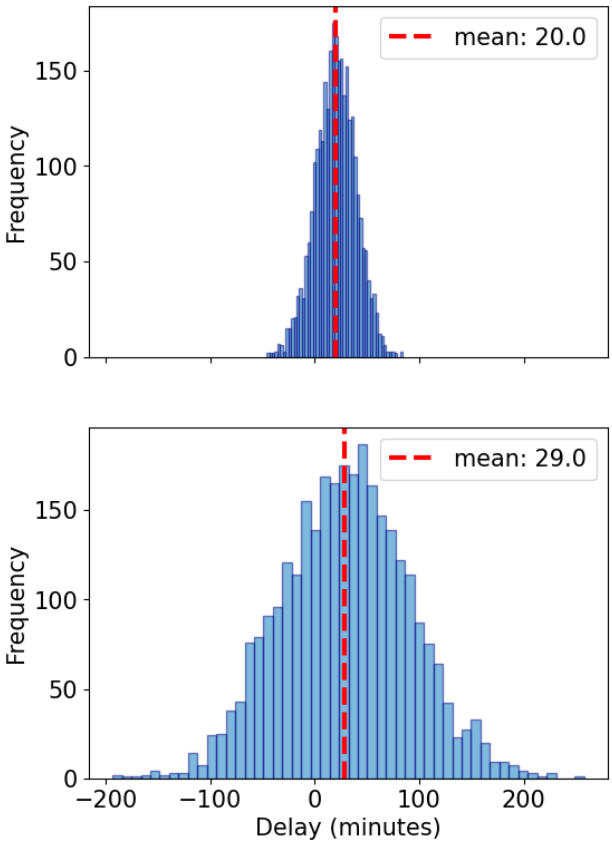


Figure 8. Flights flying from BOS to JFK then to MIA. The top plot shows the distribution of arrival times in MIA if the flight leaves on-time from BOS. The bottom distribution has an unknown departure time from BOS.

In addition, this study found that aircraft were able to gain or lose a significant amount of time while stopping at certain airports, most commonly at regional airports. The airports that induced the largest gains and losses, provided that the value was statistically significant from the mean (p-value less than 0.05), are plotted in **Figure 9**. Here, it is possible to observe that the most extreme values are at regional airports, or airports in the state of New York, such as LGA and JFK. This result suggests that airline operators scheduling flights at regional airports were more likely to overestimate or underestimate the necessary time spent on the ground than those estimating times for larger airports.

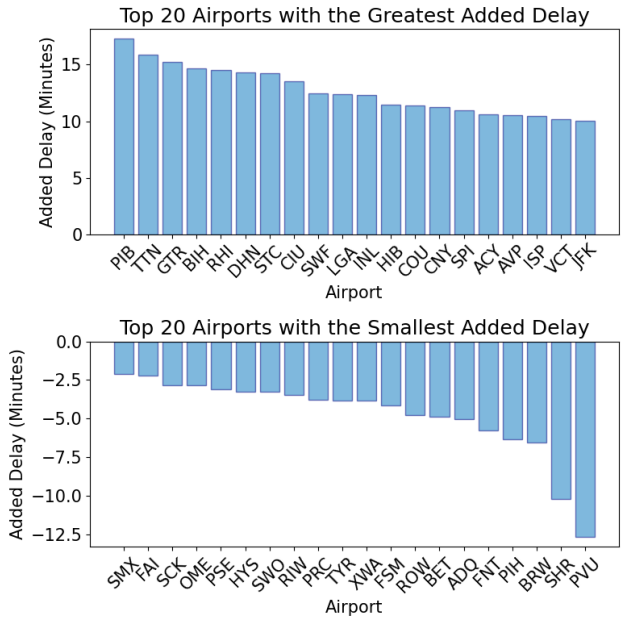


Figure 9. Mean change in the delay time at certain airports. The time corresponds to the mean difference between the arrival delay and the departure delay of an aircraft passing through an airport for airports with a statistically significant number of samples (p-value less than 0.05).

Meanwhile, it is possible to conduct a similar analysis for flights in the air. In **Figure 10**, this study shows that the routes with the largest delay had an average lost time of over 50 minutes, while those with the largest gains were over 75 minutes.

3.3. Random Walks in the Airline Network

The fourth application of the mathematical framework corresponded to the random walks thought experiment. After computing the S matrix for the entire network, it was possible to start evaluating $S^n \cdot \Pi_o$ for various values of n and random vectors Π_o, satisfying the condition that the entries of add up to 1. Starting at $n = 20$, it was possible to observe that $S^n \cdot \Pi_o = \Pi_{20}$ started to converge towards a stationary S^* distribution. Thus, for the results presented in **Figure 11**, a value of $n = 30$ was selected (which ultimately meant that n need not be large for the results to start behaving as infinite walks, making the thought experiment a more realistic application to the Southwest Airlines example). As the figure suggests, the highest probability of arrival delay was seen by Denver (DEN), followed by Atlanta (ATL), Chicago (ORD), Dallas (DFW) and Las Vegas (LAS). Just like with the toy problem above, what these results sug-

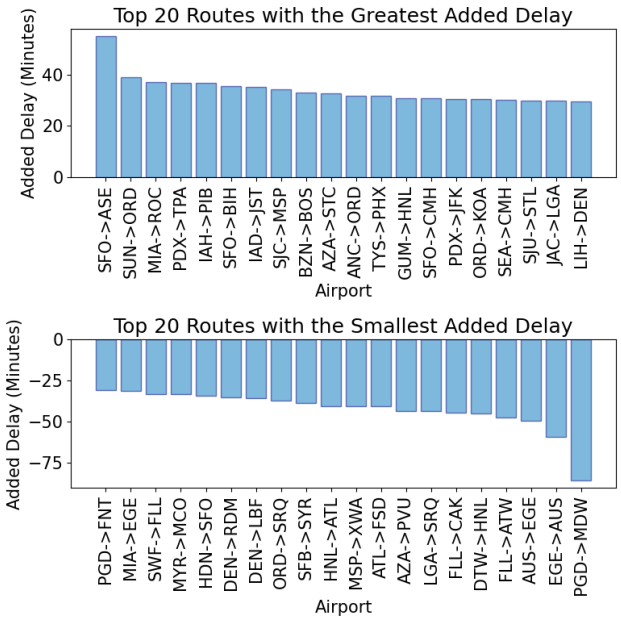


Figure 10. Mean change in the delay time on certain routes. The time corresponds to the mean difference between the departure delay and the arrival delay for routes with a statistically significant number of samples (p-value less than 0.05).

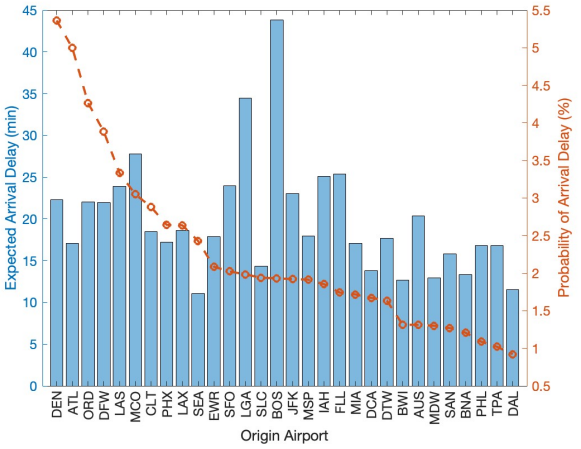


Figure 11. Expected arrival delays (left y-axis) and probability of delay (right y-axis) distribution for infinite random walks in the U.S. Airline Network, for the 30 biggest airports.

gest is that all random destinations "XXX" shared the same flight delays distribution, with 5.37% of all delayed flights coming from DEN, 5.00% from ATL, 4.26% from BOS, 3.88% from DFW, and 3.33% from LAS. Hence, in this thought experiment, if a passenger arrived delayed to some random destination "XXX" (after taking an "infinite random walk"), chances are highest that it started its journey in DEN, followed by the other four airports. In fact, the 30 airports displayed in the figure account for over 60% of these probabilities, meaning more generally that there is +60% chance that a passenger arriving delayed into some destination "XXX" came from these airports.

4. Conclusion

As the airline industry continues to grow, it is critical to pay close attention to the evolution of flight delays, due to their safety, economic, and environmental effects. By providing open-source tools, this study contributes to the efforts in improving delay visualization and quantification, both for passengers, policy makers, and airlines in the United States. While this work provides valuable insights surrounding delay trends in the US airline industry, there are several limitations and opportunities for future expansion. First, all the analyses were performed offline, using the BTS January 2019 data. Developing an interactive online graphical user interface (GUI) with real-time, up-to-date delay predictions would enhance the practical applicability of this work's findings, allowing stakeholders to identify causes of delays and evaluate mitigation strategies more rapidly. Secondly, integrating weather data into the analysis could significantly improve the accuracy of delay predictions, as weather conditions play a crucial role in flight operations. Finally, data availability-permitting, expanding the analysis to encompass the global aviation network, including international flights, would offer a more comprehensive understanding of delay patterns and their impact on a worldwide scale.

References

[1] Michael Ball, Cynthia Barnhart, Martin Dresner, Mark Hansen, Kevin Neels, AR Odoni, Everett Peterson, Lance Sherry, Antonio Trani, and Bo Zou. Total delay impact study: a comprehensive assessment of the costs and impacts of flight delay in the united states. 2010.

[2] Bureau of Transportation Statistics. Marketing carrier on-time performance, 2024. Online: https://www.transtats.bts.gov/Fields.asp?gnoyr_VQ=FGK.

[3] Bureau of Transportation Statistics. Schedule b-43 inventory, 2024. Online: https://www.transtats.bts.gov/DL_SelectFields.aspx?gnoyr_VQ=GEH&QO_fu146_anzr=Nv4%20Pn44vr4%20Sv0n0pvny.

[4] Chen Chen, Chenhui Li, Junjie Chen, and Changbo Wang. Vfdp: Visual analysis of flight delay and propagation on a geographical map. *IEEE Transactions on*

Intelligent Transportation Systems, 23(4):3510–3521, 2020.

[5] DMMS Dissanayaka, V Adikariwattage, and HR Pasindu. Evaluation of co emission from flight delays at taxiing phase in bandaranaike international airport (bia). *Transportation Research Procedia*, 48:2108–2126, 2020.

[6] Flight Aware. Total delays within, into, or out of the united states today, 2024. Online: hhttps://www.flightaware.com/live/airport/delays.

[7] Xiaowen Fu, Huan Jin, Shaoxuan Liu, Tae H Oum, and Jia Yan. Exploring network effects of point-to-point networks: An investigation of the spatial patterns of southwest airlines' network. *Transport Policy*, 76:36–45, 2019.

[8] International Civil Aviation Organization. Future of aviation, 2022. Online: https://www.icao.int/Meetings/FutureOfAviation/Pages/default.aspx.

[9] International Civil Aviation Organization. El mundo del transporte aéreo en 2019, 2024. Online: https://www.icao.int/annual-report-2019/Pages/the-world-of-air-transport-in-2019.aspx.

[10] Megan S Ryerson and Amber Woodburn. Build airport capacity or manage flight demand? how regional planners can lead american aviation into a new frontier of demand management. *Journal of the American Planning Association*, 80(2):138–152, 2014.

[11] Ingrid Sekelová, Peter Korba, Simona Pjurová, Siva Marimuthu, and Utku Kale. Reducing the environmental impact of aviation by minimizing flight delays. In *International Symposium on Energy Management and Sustainability*, pages 711–717. Springer, 2022.

[12] Huawei Wang, Jun Gao, et al. Bayesian network assessment method for civil aviation safety based on flight delays. *Mathematical Problems in Engineering*, 2013, 2013.

A. More on the Transition Probability Matrix

Link to Github repo containing this work's code: https://github.com/andyeske/Flight-delays.

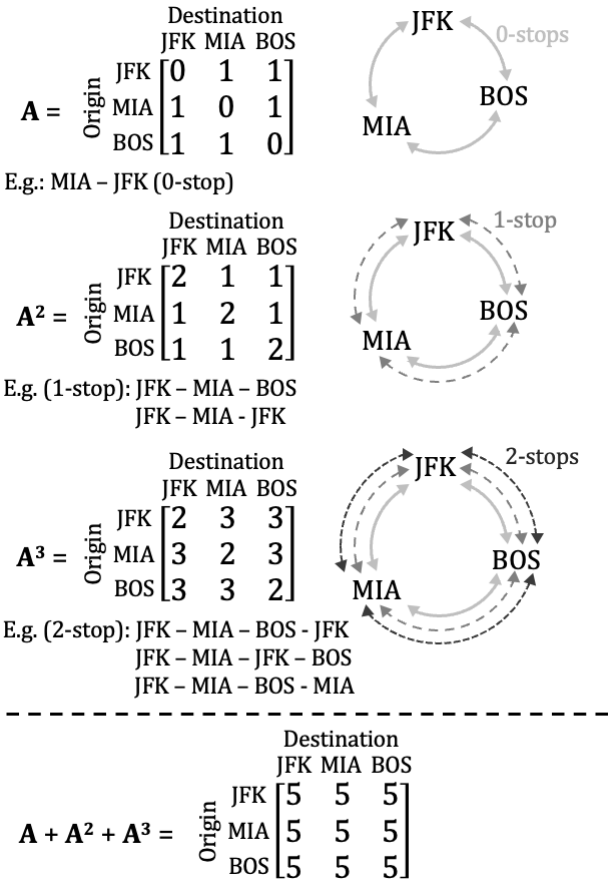


Figure A1: Adjacency matrix representation of the toy JFK - MIA - BOS network.

Stops	Path	Probability
0	JFK – MIA	$P_{JFK \to MIA} = 0.73$
1	JFK – BOS – MIA	$P_{JFK \to BOS} \cdot P_{BOS \to MIA} = 0.40 \cdot 0.27 = 0.108$
2	JFK – MIA – BOS – MIA JFK – BOS – JFK – MIA JFK – MIA – JFK – MIA	$P_{JFK \to MIA} \cdot P_{MIA \to BOS} \cdot P_{BOS \to MIA} = 0.73 \cdot 0.60 \cdot 0.27 = 0.118$ $P_{JFK \to BOS} \cdot P_{BOS \to JFK} \cdot P_{JFK \to MIA} = 0.40 \cdot 0.37 \cdot 0.73 = 0.108$ $P_{JFK \to MIA} \cdot P_{MIA \to JFK} \cdot P_{JFK \to MIA} = 0.73 \cdot 0.63 \cdot 0.73 = 0.336$ $P_{total} = 0.118 + 0.108 + 0.336 = 0.562$

Table A1: Route-specific probability computation, examining the propagation of delays with increasing flight sequence length.