Basel Mostafa

bmostafa@caltech.edu

Queue Scheduling Model of SFO's September 2019 Flight Delays Due to Construction

Problem

"In September 2019, San Francisco International Airport (SFO) closed one of its main runways, 28L, for several weeks leading to many flight cancellations and multi hour delays. In this problem, work out the delays suffered by flights on an hour-by-hour basis during the closure. For example, what time will a flight scheduled to leave at 7am actually leave? By way of contrast, what time will a flight scheduled for a 10pm departure actually leave?"

Abstract

This solution begins with an attempt at creating a representative flight arrival/departure schedule for the month of September, for which this data is not yet released. A set of rules for the interaction between planes and runways is then drafted. The problem is then modeled as a queue scheduling problem, and a program is created to simulate it as such. Data from the simulation is then summarized to approximate flight delays on an hour-by-hour basis as requested by the problem.

1. Creating a Representative Schedule

Since the flight schedule for September 2019 is not yet available for SFO, it is necessary to create a representative schedule that approximates the true schedule's important characteristics as closely as possible. We begin with an understanding of the cause of the delays at SFO. Runway 28L was one of two runways used for arrivals at SFO. Intuitively, the number of planes at the airport can be expected to be in dynamic equilibrium. Therefore, if the capacity to host landing aircraft is cut in half, one can expect that the number of flight operations at SFO will be approximately reduced by half. However, this did not occur. Flight activity at SFO was only reduced by 13% (Ho).

In creating a representative flight schedule for September 2019, we will extrapolate flight schedule data relevant to our model from 2013 onwards. In our model, we are concerned by the number of flights and the proportion of international flights to total flights. This data is summarized in the figures that follow.

Figure 1:

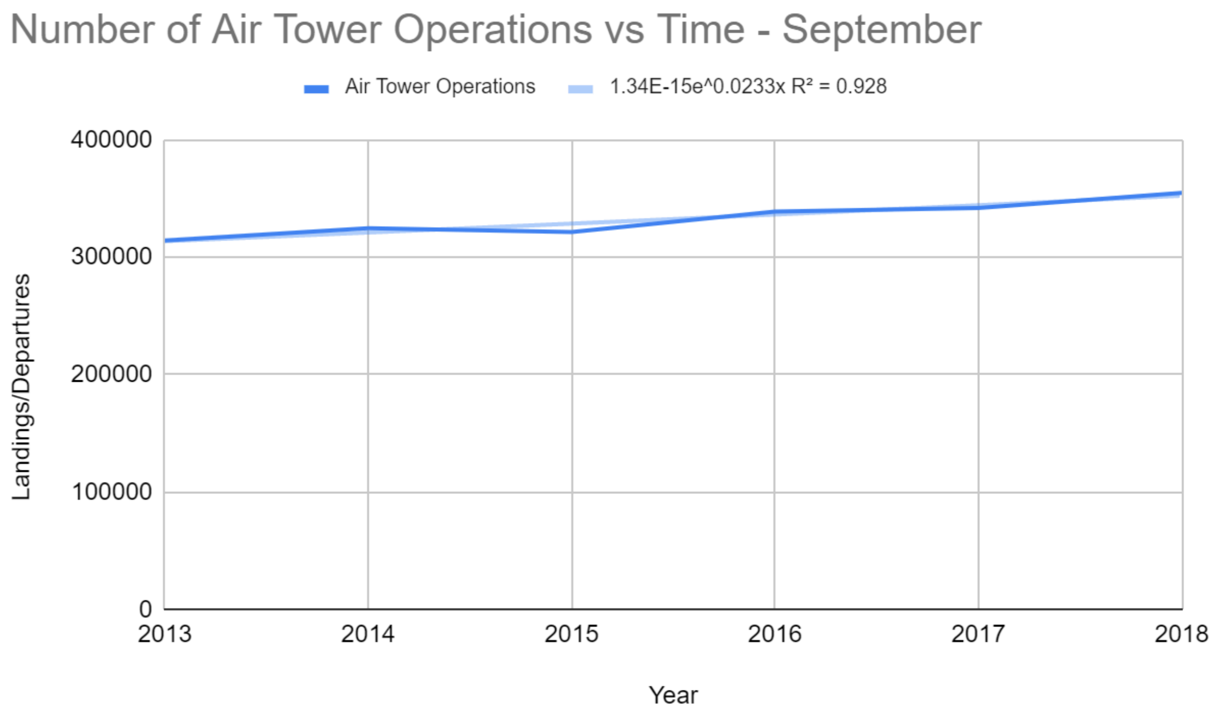
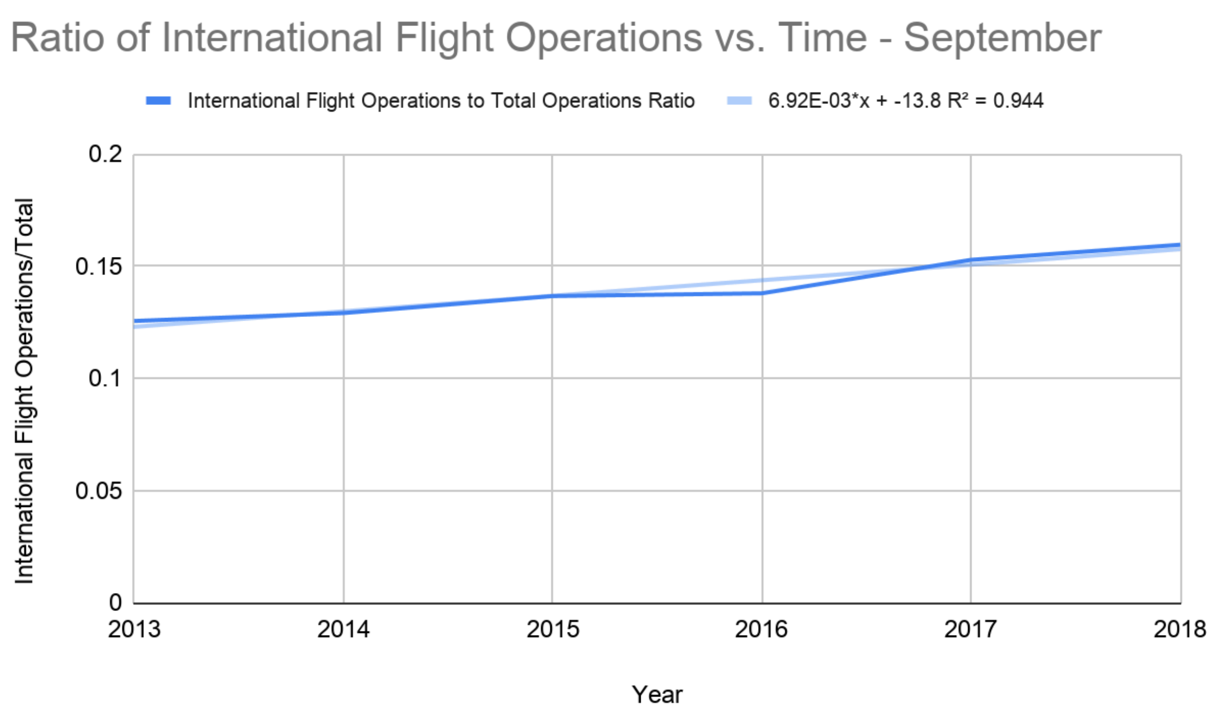


Figure 2:



An exponential regression line is used in Figure 1 since the demand for flight increases with population, which is known to grow exponentially. A quadratic regression line is used in Figure 2 since it fits the data especially well. Both regression lines fit the data relatively well, each with a coefficient of determination close to 1.

Extrapolating the data yields the following critical characteristics for our September 2019 schedule:

Number of Flights/Departures: 359,808

Ratio of International Flights to Total: 0.158

Due to limited time and compute power, a schedule for a single day will be generated, based on the flight schedule for September 17, 2018. September 17 was specifically chosen since it corresponds to the halfway point of SFO's construction project.

We assume the following equation to be approximately true in determining the number of flights in our September 17, 2019 representative schedule.

$$\frac{\#\, September\ 2019\ Flights}{\#\, September\ 2018\ Flights} = \frac{\#\, September\ 17,2019\ Flights \div (1-0.13)}{\#\, September\ 17,2018\ Flights} \quad (1)$$

This equation is based on the assumption that the number of flights in September 17, 2019 will be in the same proportion predicted by the regression line in Figure 1, but cut by 13% (see ¶ 1 of this section). The number of September 2019 flights is 359808, the number of September 2018 flights is 354731, and the number of September 17, 2018 flights is 1011. Plugging the appropriate values into the equation and solving for the number of flights at SFO on September 17, 2019 yields 892 flights.

We assume the following equation to be approximately true in determining the number of international flights in our September 17, 2019 representative schedule.

$$\#\, International\ Flights \ = \ Ratio\ of\ International\ Flights\ to\ Total \ \cdot \#\, Flights \quad (2)$$

The ratio of international flights to total is 0.158 for September 2019, and the number of September 17, 2019 flights is 892. Plugging the appropriate values into the equation and solving for the number of international flights on September 17, 2019 yields 141.

Having gathered this information, we assume that the September 17, 2019 flight schedule is identical to the September 17, 2018 flight schedule, but with flights randomly removed until the number of flights equals the predicted number of flights. Since most flights are domestic, we start with the assumption that all flights are domestic, and randomly selected flights to designate as international until the number of international flights equals the predicted number of international flights.

Links to the data used are available below:
[Prediction Charts]
[September 17, 2018 Arrival Schedule]
[September 17, 2018 Departure Schedule]
[September 17, 2018 Combined Schedule]

The procedure described above to generate a flight schedule representative of that during the SFO construction project is greatly simplified and can be improved with more time. For example, the actual September 17, 2019 schedule is likely recorded, and will be published soon. It may have been possible to get early access to this data. The simulation could also have been run on more than just one day. Rather than randomly assigning flights to international or domestic status, it may have been possible to collect this information from the data.

2. Drafting Rules for the Simulation

While it is possible to create a much more thorough set of rules to govern the simulation, I assume the following rules in order to create a simpler model that I can deal with considering my limited time and experience.

1. Flight delays are additive
    a. Flight delays caused by the SFO construction are added to delays caused by other factors (the "effective scheduled time" is the actual time recorded on September 17, 2018 including delays).
2. Flights (representing a departing or arriving plane) are characterized exclusively by the following variables
    a. Arrival or Departure
        i. Status representing whether a flight is an arrival or departure
    b. Runway type - primary or secondary
        i. primary - 1, secondary - 2
    c. Runway time - 35 sec or 55 sec (initial - other variations tested later)
        i. International - 55 sec, since larger planes take approximately this long to take off (Cox)
        ii. Domestic - 35 sec, since smaller planes take approximately this long to take off (Cox)
    d. Effective scheduled time
        i. Time recorded for landing/take-off for September 17, 2018 used as an estimate of flight arrival/departure times post-usual delays

    e. Scheduled time
        i. Scheduled landing/take-off time for September 17, 2018 used as a prediction for a typical day during the SFO construction
3. Runways are characterized exclusively by type (see rule 2a)
4. Only one plane can interact with a runway at a time
5. As soon as a plane gets off of a runway, the next plane in queue can interact with the runway

This model may be improved by individually tracking each plane and ensuring that no plane departs without having landed. It also may be improved by careful study of the individual factors that cause flight delays and fitting these factors more accurately with the SFO construction project. Including a larger array of runway times based on careful study of airlines' current fleets can also improve this model. These are just a few of the many improvements that can be made to this model.

3. Creating the Simulation Program

In general, the simulation program works as follows.

Flights are queued in order of their effective scheduled times, from least to greatest. A flight can only be assigned to a runway if its effective scheduled time has passed and its type matches the runway type. Every iteration of the main controller assigns at least one flight to a runway and/or "advances time". Time is "advanced" to the next flight's effective scheduled time, or if this is less than the time left for a runway to be available, to the minimum time for a runway to become available. The process continues until all flights have been assigned. While the main controller follows a simple concept, there is a problem that must be addressed before the main controller can be useful.

The data must be presented to the program in a format that it understands. While manually inserting this data into the program's source code is possible, it is very impractical with over a thousand data points per day of flight. Instead, I decided to copy and paste the data into a text file, and to use the Java Scanner class to read said file. It makes sense to create a class to manage the conversion of this text file into data that my main controller can deal with. This class, named "Reader," creates a list of flights for use by the main controller class.

An additional challenge was working around floating-point errors, which often lead to unexpected results. To deal with this problem, all decimals are converted to integers a factor of 100000 times larger, and are converted back when displaying the results.

4. Simulation Results

Arrival and departure flights are grouped by their scheduled time. Flights from 0:00 - 0:30 and 23:30 - 0:00 are considered to be in the 0:00 group. All other flights are rounded to the nearest hour to determine their group. The results of two iterations of the program are shown below.

Hourly Delay Arrivals	Hourly Delay Departures	Hourly Delay Arrivals	Hourly Delay Departures
0: -4	0: -5	0: -4	0: -6
1: -37	1: -11	1: -24	1: -16
2: 0	2: 0	2: 0	2: 0
3: 0	3: 0	3: 0	3: 0
4: -31	4: 0	4: -31	4: 0
5: -23	5: -10	5: -23	5: -10
6: -15	6: -5	6: -15	6: -6
7: -19	7: -5	7: -23	7: -4
8: -13	8: 4	8: -14	8: 3
9: -13	9: -5	9: -13	9: -5
10: -7	10: 4	10: -9	10: -1
11: 1	11: 0	11: 0	11: 6
12: -6	12: 2	12: -8	12: 0
13: -13	13: 0	13: -14	13: -1
14: -3	14: 0	14: -3	14: -1
15: -6	15: -4	15: -6	15: -1
16: -13	16: 5	16: -12	16: 7
17: -11	17: -7	17: -12	17: -1
18: 1	18: -1	18: 0	18: -1
19: -4	19: 1	19: 0	19: 1
20: 25	20: -3	20: 21	20: -3
21: 30	21: 21	21: 35	21: 20
22: 23	22: 6	22: 24	22: 7
23: 16	23: -3	23: 12	23: -4

The numbers on the left of the colons represent the approximate hour of a group of flights. The numbers on the right represent the average delays in hundredths of an hour. The general trend appears to be that delays appear more frequently in larger magnitudes later on in the day. It is possible that this is due to a build up in delayed flights over time as a result of the reduced airport capacity. It would be interesting to see what would occur if the simulation were continued over a larger period of time.

It is worth noting the simulation results predict delays on the order of 15 - 30 minutes, rather than 2 - 3 hours as was the approximate average delay that was observed. I expect this to be largely due to my optimistic estimate of the time that each flight is reserved on a runway. While each plane spends under a minute on a runway, it is likely that airport operations are not perfectly efficient and that some additional delays are enforced for safety reasons. I was unable to source and analyze the relevant data in time, but to get a feel for how the simulation results change with this parameter, I ran it again assuming a uniform 2.7 minutes between flights, regardless of flight type. The results of two iterations of the program are shown below.

Hourly Delay Arrivals	Hourly Delay Departures	Hourly Delay Arrivals	Hourly Delay Departures
0: 344	0: 79	0: 365	0: 130
1: -34	1: -14	1: -34	1: -10
2: 0	2: 0	2: 0	2: 0
3: 0	3: 0	3: 0	3: 0
4: -31	4: 0	4: -31	4: 0
5: -24	5: -10	5: -21	5: -10
6: -13	6: -3	6: -15	6: -5
7: -17	7: -3	7: -17	7: -6
8: -2	8: 9	8: 0	8: 11
9: 34	9: 5	9: 34	9: 4
10: 66	10: 6	10: 81	10: 24
11: 121	11: 28	11: 152	11: 38
12: 175	12: 27	12: 208	12: 29
13: 211	13: 48	13: 242	13: 67
14: 222	14: 0	14: 256	14: 50
15: 255	15: 80	15: 278	15: 41
16: 273	16: 38	16: 309	16: 90
17: 283	17: 28	17: 325	17: 47
18: 309	18: 43	18: 346	18: 32
19: 331	19: 78	19: 367	19: 69
20: 386	20: 31	20: 424	20: 53
21: 427	21: 78	21: 467	21: 103
22: 443	22: 36	22: 480	22: 37
23: 460	23: 54	23: 491	23: 61

The delays predicted by this second set of parameters are much more representative of the observed delays, which indicates that the model might be a reasonable approximation of SFO's actual operations. However, as a whole, this model does not seem to be very accurate, as it depends very heavily on the set of parameters used.

Given more time, I would have done more research into the rules that SFO uses to grant runway access to planes to fine tune my model, and I would have run the simulation on data covering longer stretches of time. I also would have considered different methods of approaching the problem. I expect that a numerical approach focused on extrapolating trends into the future would have been more likely to yield results closer to reality, since such an approach would not have depended so heavily on the nature of the assumptions made.

Resources

"Bureau of Transportation Statistics." *Bureau of Transportation Statistics*, US Department of Transportation, 16 Oct. 2019, www.bts.gov/.

Cox, John. "Ask the Captain: Does Every Takeoff Take the Same Amount of Time?" *USA Today*, Gannett Satellite Information Network, 14 Jan. 2013, www.usatoday.com/story/travel/columnist/cox/2013/01/13/ask-the-captain-takeoff-take-same-amount-of-time/1831525/.

Ho, Vivian. "After Thousands of Delays at One of America's Biggest Airports, Beleaguered Travelers Ask – How?" *The Guardian*, Guardian News and Media, 13 Sept. 2019, www.theguardian.com/us-news/2019/sep/12/sfo-flight-delays-construction-airport.