MATH5314 Project 3: Queueing System

Junyu Chen, Kewei Zhang, Max Guthrie, Wenzhe Zhang

February 25, 2025

1 Background

Pre-board screening (PBS), overseen by the Borealian Aeronautic Security Agency (BASA) at four main airfields (Auckland, Chebucto, Saint-François, and Queenston), involves a standardized process of main line entry, boarding pass scanning, and security screening of travelers and carry-on bags. PBS wait times are influenced by multiple factors including passenger volume, security server availability, processing speeds, and flight schedule intensity.

The project utilizes four datasets containing passenger and flight characteristics across different time periods to analyze wait time variations affected by seasonal, daily, and weekly patterns, as well as broader trends in passenger traffic and aircraft schedules. The main tasks include creating a data dictionary, analyzing datasets, and developing queueing models to predict wait times, with the ultimate goal of helping BASA improve PBS efficiency and enhance traveler experience.

2 Data Exploration and Visualization

2.1 Data Exploration

Initial analysis showed that dat_F_sub.csv was an aggregated version of dat_P_sub_c.csv at the flight level, combining individual passenger data by flight. Therefore, dat_F_sub.csv was used for analyzing departure intensity patterns, as it grouped data by flight ID, while dat_P_sub_c.csv and years20262030.csv were organized by passenger ID. Using the passenger-level datasets for flight analysis would have resulted in data duplication.

For passenger-related analyses (wait times, server availability, delays), dat_P_sub_c.csv and years20262030.csv were selected due to their unique passenger IDs. BASA_AUC_2028_912.csv was excluded due to limited August flight data and missing C_Start and C_Avg values. Consequently, passenger-level EDA visualizations combine data from dat_P_sub_c.csv and years20262030.csv, while flight-level analyses use dat_F_sub.csv exclusively.

2.2 Data Dictionary

After exploring the datasets, the following variables were deemed necessary for the ensuing analysis.

Column	Data Type	Description
Airfield	String	Code of the airfield
S2	Date	Timestamp for when the boarding pass was scanned at S2
Flight_ID	Integer	Unique identifier for each flight
Sch_Dep	Integer	Scheduled departure time (Unix timestamp)
Act_Dep	Integer	Actual departure time (Unix timestamp)
Wait_Time	Integer	Passenger wait time in minutes
C_Start	Integer	Number of available servers at the start
C0	Integer	Indicates when the passenger arrives at the security line
C_avg	Float	Average number of available servers
BFO_Dest_City	String	Destination city for the flight
BFO_Dest_Country_Code	String	Destination country code for the flight
Time_of_Day	Mixed type	Categorized time of day for the flight
Period_of_Week	Mixed type	weekday or weekend
Day_of_Week	Mixed type	day of the week
Month	Mixed type	Month of the flight
Season	Mixed type	Season of the flight
Year	Integer	Year of the flight
tot_pass	Integer	Total number of passengers of the flight
N	Integer	Number of passengers for whom $WT\_Flag \neq 0$ (i.e. $Wait\_Time \neq 0$
mean	Float	Average wait time for passengers on given flight
median	Integer	Median wait time for passengers on given flight
mean_City_Flag	Float	Flag indicating mean wait time for city-based category
sum_city_mode	Integer	Sum of mode values for city-based data
sum_country_mode	Integer	Sum of mode values for country-based data
N_of_Dest_City	Integer	Number of unique destination cities
N_of_Dest_Country	Integer	Number of unique destination countries
WT_flag	Integer	Equals 0 if $Wait\_Time \neq NA$, 1 if $Wait\_Time = NA$
S2_Sch_Flag	Integer	Equals 0 if $S2 < Sch\_Dep$, 1 if $S2 >= Sch\_Dep$
S2_Act_Flag	Integer	Equals 0 if $S2 < Act\_Dep$, 1 if $S2 >= Act\_Dep$
Sch_Act_Flag	Integer	Equals 0 if $Sch\_Dep = Act\_Dep$, 1 if $Sch\_Dep \neq Act\_Dep$
Delay_in_Seconds	Integer	Flight delay duration in seconds

Table 1: Data Dictionary for dat_sub_F.csv, dat_P_sub_c.csv, years20262030.csv

2.3 Data Cleaning

In our airport operations analysis project, we began by examining four interconnected datasets, implementing several crucial formatting changes. Initial corrections included fixing weekday/weekend classification errors in BASA_AUC2018_912.csv and years20262030.csv, standardizing temporal variables (Month, Season, Time_of_Week) across all datasets, and addressing missing destination information in dat_P_sub_c.csv. We also unified BFO_Dest_City values across datasets, converting variations like 'BORQUE' to 'QUE' for consistency.

A significant discovery was a persistent 10-minute (600-second) offset between departure times in the BASA dataset compared to data_P entries. This synchronization issue required standardization to ensure all subsequent analyses would operate within the same temporal framework. Additionally, we encountered data completeness challenges, particularly with critical fields like C_Start and C_Avg in BASA containing numerous null values. After careful consideration, we selected data_P as our primary source due to its more complete information, resulting in the removal of 26 unique passengers (0.023% of the dataset).

One intriguing finding during the cleaning process was the presence of multiple destinations linked to single flight IDs. Rather than treating these as errors, we recognized them as valuable indicators of connecting flights, demonstrating how thorough data cleaning can reveal meaningful patterns in passenger behavior rather than simply eliminating apparent anomalies.

Our verification approach involved cross-referencing flight-level statistics between dat_F_sub.csv and calculations from dat_P, revealing remarkable consistency with only two minor discrepancies among thousands of records. This validation process identified special cases requiring attention, such as a single-passenger flight (ID 21428), while building confidence in our cleaned dataset's reliability.

The final cleaned dataset, comprising 114,132 passenger records and 3,583 validated flight records, provides a robust foundation for operational insights. The combination of standardized timing data, verified flight statistics, and validated destination patterns enables confident decision-making while acknowledging the minimal limitations introduced by necessary data removal. This meticulous cleaning process transformed raw data into a reliable resource for airport operations analysis.

2.4 Data Visualization

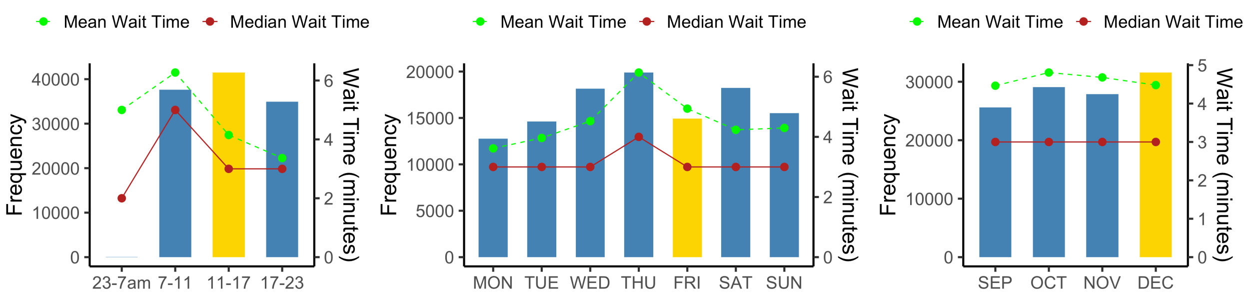

Figure 1: Passengers Level - Schedule Intensity of Passengers with their Avg & Median Wait Time

The daily rhythm of an airport unfolds through distinct patterns of passenger volume and flight scheduling. Looking at Fig. 1, we can see how passenger traffic builds steadily from dawn, reaching its first peak during the morning rush hour when business travelers dominate the terminals. A second, more pronounced surge occurs during the late afternoon and early evening hours as both business and leisure travelers converge. This ebb and flow of human traffic is mirrored in Fig. 2's flight schedule intensity data, with airlines clustering their departures during these peak periods to meet demand.

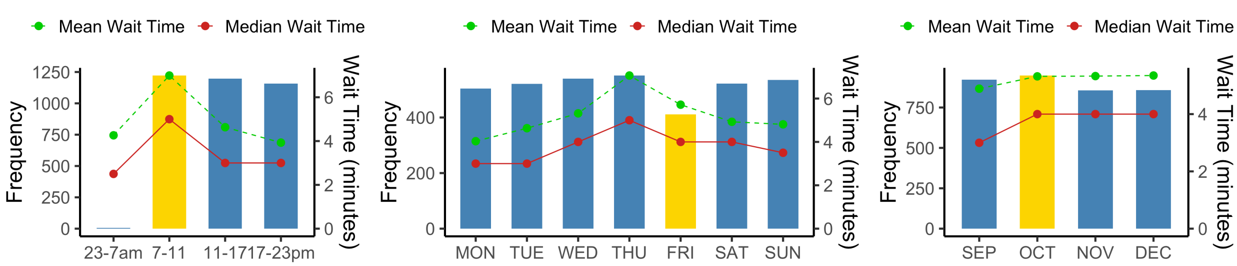

Figure 2: Flight Level - Schedule Intensity of Departing Flights with their Avg & Median Wait Time

Wait times for passengers at service counters and security checkpoints are directly impacted by these fluctuating activity levels. The patterns of wait times at various times of day and week are depicted in Fig. 3, and Fig. 4 shows a clear correlation: wait times tend to increase as passenger volumes rise (left panel), and delays become more noticeable (right panel) when the number of flights surpasses specific thresholds, causing a cascade effect throughout the terminal. According to this statistics, the existing personnel numbers may not be able to effectively manage these spike periods.

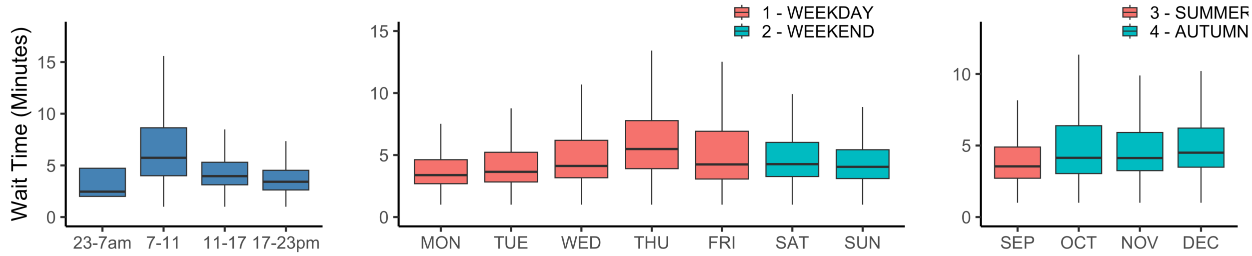

Figure 3: Wait Time Distribution Across Time of the Day, Days of the Week, and Months

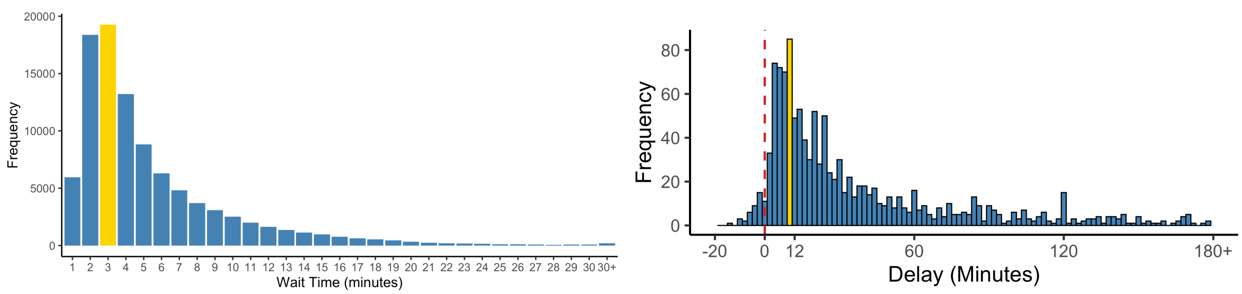

Figure 4: Distribution of Wait Time & Delay Time

In some cities, the average wait time is significantly higher than the median wait time, suggesting that a few passengers with exceptionally high wait times might be skewing the average upward.

Then comes to the distribution of destination city 5. From the result it indicates for most cities, the median wait time remains relatively stable, hovering around 3 to 4 minutes. There are a few cities where the median wait time increases significantly, indicating that some passengers in these cities experience higher wait times.

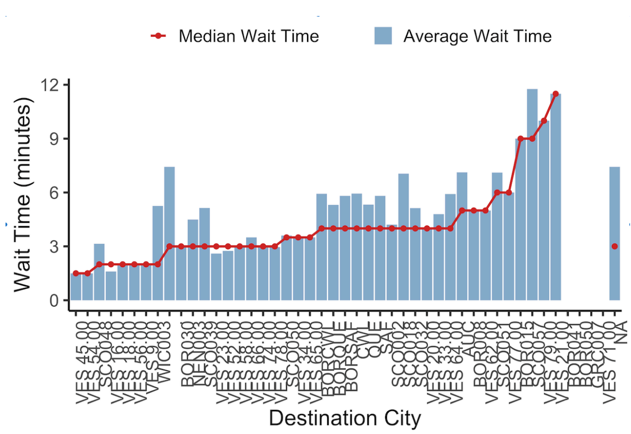

Figure 5: Distribution of Wait Time & Destination City

Finally, Fig. 6 provides a fascinating glimpse into how airports attempt to manage these fluctuations through service rates. There's a clear inverse relationship between wait times and service rates – as more servers are deployed and processing speeds increase, wait times generally decrease. However, the scattered pattern in the data suggests other factors are at play, such as variations in passenger characteristics, time of day, or even seasonal effects. This highlights the complex challenge airports face in optimizing their operations to maintain efficient passenger flow while managing resources effectively throughout different times of the day and week.

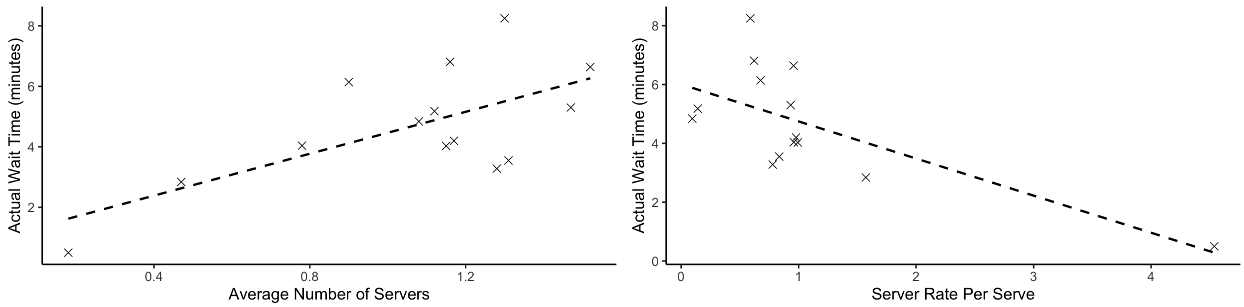

Figure 6: Server Efficiency and Wait Time Correlation

3 Queueing System Model

3.1 Queueing Theory Framework

The fundamental framework of this analysis relies on queuing theory[1, 2, 3], specifically utilizing an M/M/c queuing model. To establish the model parameters, we begin by defining the arrival process. Let N(t) denote the number of arrivals in the cluster by time t, following a Poisson distribution with mean λt:

$$P[N(t+s) - N(s) = n] = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, n = 0, 1, ...$$

For an M/M/1 system, when the arrival rate λ is known and the average wait time $\bar{W}_q$ can be determined through empirical data, we can derive the service rate μ using:

$$\hat{\mu}_M = \frac{\bar{W}_q\lambda + \sqrt{(\bar{W}_q\lambda)^2 + 4\bar{W}_q\lambda}}{2\bar{W}_q}$$

To account for the relationship between service capacity and passenger flow, we introduce a regression model that links the service rate to both the number of active servers and the arrival rate: $\mu = ac + b$ where c represents the number of servers, and a and b are regression parameters estimated from historical data.

In our model, we focus on these three key predictive metrics: System Utilization which using the regression-based service rate:

$$\text{Reg } \rho = \frac{\lambda}{\text{Reg Serv Rate}}$$

Average Queue Time which is the fundamental queue time calculation:

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

Using the regression-predicted service rate:

$$W_q^{\text{Reg}} = \frac{\lambda}{\text{Reg Serv Rate} \cdot (\text{Reg Serv Rate} - \lambda)}$$

Average System Time which is the basic system time equation:

$$W = W_q + \frac{1}{\mu}$$

With regression prediction:

$$W^{\text{Reg}} = W_q^{\text{Reg}} + \frac{1}{\text{Reg Serv Rate}}$$

These predictive metrics provide a comprehensive view of system performance, with the wait time calculations incorporating both queuing and service components. The Quality of Service (QoS) levels can then be determined as:

$$\hat{p}_M(x) = 1 - \frac{\lambda}{\hat{\mu}_M} e^{-(\hat{\mu}_M - \lambda)x} \in (0, 1), \lambda < \hat{\mu}_M$$

3.2 Prediction Results

Arrival Rates The analysis of arrival rates shows distinct operational patterns at AUC and SAF 7a. AUC experiences peak flow during weekday mornings (07:00-11:00) with 1.320 passengers per minute (27,255 total), followed by decreases in afternoon (0.949) and evening periods (0.764). Weekend patterns follow similar distributions at slightly lower rates, with morning peaks at 1.202 passengers per minute (10,383 total). SAF operates at notably lower volumes, peaking at 0.598 passengers per minute during weekday mornings (718 total), with more balanced distribution across periods. These arrival rates λ serve as fundamental inputs for our queuing models, with the volume differences between airfields necessitating distinct operational strategies.

Cluster				# of Hours	Count	Avg Arrival Rate (mins)
AUC	Week Day	NIGHT	23:00 - 07:00	40	82	0.034
		MORNING	07:00 - 11:00	344	27255	1.320
		AFTERNOON	11:00 - 17:00	516	29373	0.949
		EVENING	17:00 - 23:00	516	23661	0.764
	Week-end	NIGHT	23:00 - 07:00	8	1	0.002
		MORNING	07:00 - 11:00	144	10383	1.202
		AFTERNOON	11:00 - 17:00	216	12119	0.935
		EVENING	17:00 - 23:00	216	11262	0.869
SAF	Week Day	NIGHT	23:00 - 07:00	-	-	-
		MORNING	07:00 - 11:00	20	718	0.598
		AFTERNOON	11:00 - 17:00	30	914	0.508
		EVENING	17:00 - 23:00	30	435	0.242
	Week-end	NIGHT	23:00 - 07:00	8	34	0.071
		MORNING	07:00 - 11:00	8	231	0.481
		AFTERNOON	11:00 - 17:00	12	474	0.658
		EVENING	17:00 - 23:00	12	408	0.567

(a) Arrival Rates

Cluster				Avg # of Servers	Distribution of # of Active Servers			
					0	1	2	3
AUC	Week Day	NIGHT	23:00 - 07:00	1.08	3.7%	85.4%	11.0%	-
		MORNING	07:00 - 11:00	1.52	-	50.2%	47.9%	1.9%
		AFTERNOON	11:00 - 17:00	1.17	-	83.7%	15.9%	0.5%
		EVENING	17:00 - 23:00	1.28	-	72.6%	26.6%	0.8%
	Week-end	NIGHT	23:00 - 07:00	1.00	-	100.0%	-	-
		MORNING	07:00 - 11:00	1.47	-	55.3%	43.1%	1.6%
		AFTERNOON	11:00 - 17:00	1.15	-	85.6%	14.4%	
		EVENING	17:00 - 23:00	1.31	-	71.4%	26.9%	1.7%
SAF	Week Day	NIGHT	23:00 - 07:00	0.00	-	-	-	-
		MORNING	07:00 - 11:00	1.16	23.4%	39.3%	37.3%	-
		AFTERNOON	11:00 - 17:00	0.47	54.4%	43.0%	2.6%	-
		EVENING	17:00 - 23:00	0.18	81.8%	18.2%	-	-
	Week-end	NIGHT	23:00 - 07:00	1.12	44.1%	-	55.9%	-
		MORNING	07:00 - 11:00	0.90	20.3%	69.7%	10.0%	-
		AFTERNOON	11:00 - 17:00	1.30	8.9%	55.3%	35.9%	-
		EVENING	17:00 - 23:00	0.78	54.9%	11.8%	33.3%	-

(b) Average Number of Servers

Figure 7: Results Tables

Average Number of Servers The distribution of active servers across different time periods from 7b indicates systematic resource allocation strategies at both checkpoints. At AUC, weekday morning peaks average 1.52 servers, split between one server (50.2%) and two servers (47.9%), with afternoon and evening periods averaging 1.17 and 1.28 servers respectively. Weekend operations maintain similar patterns, with 1.47 servers during morning peaks (55.3% one-server, 43.1% two-server), and slightly lower levels during other periods (1.15-1.31 servers). SAF operates with fewer resources, averaging 0.18-1.16 servers on weekdays, with significant zero-capacity periods (up to 81.8% during evenings). Weekend averages range from 0.78-1.30 servers, reflecting its supplementary role with on-demand resource allocation.

Average Wait Time and Performance Levels The analysis of wait times and performance levels 8a demonstrates varying service efficiencies across different operational periods. At AUC weekday mornings, despite peak volumes (13,179 passengers), average wait time is 6.640 minutes, with 78.2% waiting less than 10 minutes and 90.3% less than 15 minutes. Performance improves in afternoons and evenings (4.196 and 3.276 minutes respectively), with over 93% waiting less than 10 minutes. Weekend performance remains consistent, showing morning wait times of 5.3 minutes (5,314 passengers, 87.9% under 10 minutes) and afternoon/evening times of 4.034 and 3.553 minutes. SAF shows greater variation despite lower volumes, with weekday morning waits averaging 6.812 minutes, improving to 2.844 minutes in afternoons. Weekend wait times range from 4.042 to 8.247 minutes, though with smaller passenger counts (219-269), these variations have limited systemic impact.

Cluster				Count	Avg Wait	Performance					
						5m	10m	15m	20m	25m	30m
AUC	Week Day	NIGHT	23:00 - 07:00	68	4.841	79.3%	87.8%	89.0%	91.5%	95.1%	98.8%
		MORNING	07:00 - 11:00	13179	6.640	55.6%	78.2%	90.3%	96.5%	98.6%	99.5%
		AFTERNOON	11:00 - 17:00	18639	4.196	76.1%	93.2%	97.7%	99.2%	99.6%	99.9%
		EVENING	17:00 - 23:00	13804	3.276	84.3%	96.5%	99.2%	99.8%	99.9%	100.0%
	Week-end	NIGHT	23:00 - 07:00	1	0.000	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
		MORNING	07:00 - 11:00	5314	5.300	62.2%	87.9%	97.0%	99.2%	99.8%	100.0%
		AFTERNOON	11:00 - 17:00	7747	4.034	77.5%	94.5%	99.0%	99.9%	100.0%	100.0%
		EVENING	17:00 - 23:00	6078	3.553	80.8%	95.2%	98.1%	99.0%	99.5%	99.9%
SAF	Week Day	NIGHT	23:00 - 07:00	0	0.000	-	-	-	-	-	-
		MORNING	07:00 - 11:00	361	6.812	51.0%	72.3%	87.7%	96.1%	100.0%	100.0%
		AFTERNOON	11:00 - 17:00	590	2.844	82.6%	91.9%	95.7%	97.9%	99.7%	100.0%
		EVENING	17:00 - 23:00	317	0.503	99.5%	100.0%	100.0%	100.0%	100.0%	100.0%
	Week-end	NIGHT	23:00 - 07:00	23	5.176	52.9%	82.4%	91.2%	100.0%	100.0%	100.0%
		MORNING	07:00 - 11:00	131	6.143	58.9%	81.0%	89.2%	95.7%	1100.0%	100.0%
		AFTERNOON	11:00 - 17:00	269	8.247	45.4%	65.8%	85.7%	94.9%	96.6%	98.9%
		EVENING	17:00 - 23:00	219	4.042	69.1%	89.0%	94.4%	96.6%	98.3%	99.8%

(a) Average Wait Time and Performance Levels

Cluster				Est Serv Rate	Est ρ	Estimated Performance (M/M/1)					
						5m	10m	15m	20m	25m	30m
AUC	Week Day	NIGHT	23:00 - 07:00	0.103	0.332	76.4%	83.3%	88.1%	91.6%	94.0%	95.8%
		MORNING	07:00 - 11:00	1.457	0.906	54.2%	76.8%	88.3%	94.1%	97.0%	98.5%
		AFTERNOON	11:00 - 17:00	1.146	0.828	69.1%	88.5%	95.7%	98.4%	99.4%	99.8%
		EVENING	17:00 - 23:00	0.998	0.766	76.2%	92.6%	97.7%	99.3%	99.8%	99.9%
	Week-end	NIGHT	23:00 - 07:00	-	-	-	-	-	-	-	-
		MORNING	07:00 - 11:00	1.368	0.879	61.6%	83.3%	92.7%	96.8%	98.6%	99.9%
		AFTERNOON	11:00 - 17:00	1.139	0.821	70.3%	89.3%	96.1%	98.6%	99.5%	99.8%
		EVENING	17:00 - 23:00	1.093	0.795	74.0%	91.5%	97.2%	99.1%	99.7%	99.9%
SAF	Week Day	NIGHT	23:00 - 07:00	0.000	0.000	-	-	-	-	-	-
		MORNING	07:00 - 11:00	0.720	0.831	54.9%	75.5%	86.7%	92.8%	96.1%	97.9%
		AFTERNOON	11:00 - 17:00	0.747	0.680	79.4%	93.8%	98.1%	99.4%	99.8%	99.9%
		EVENING	17:00 - 23:00	0.824	0.293	98.4%	99.9%	100.0%	100.0%	100.0%	100.0%
	Week-end	NIGHT	23:00 - 07:00	0.158	0.449	70.9%	81.1%	87.8%	92.1%	94.9%	96.7%
		MORNING	07:00 - 11:00	0.610	0.789	58.5%	78.2%	88.5%	94.0%	96.8%	98.3%
		AFTERNOON	11:00 - 17:00	0.763	0.863	48.9%	69.7%	82.0%	89.4%	93.7%	96.3%
		EVENING	17:00 - 23:00	0.753	0.753	70.3%	88.3%	95.4%	98.2%	99.3%	99.7%

(b) QoS Estimates —M/M/1

Figure 8: Results Tables

QoS Estimates —M/M/1 The M/M/1 queuing model estimates provide theoretical performance predictions 8b that can be applied to compare with the actual performance data. The M/M/1 model at AUC estimates weekday service rates of 0.103-1.457, with utilization rates ρ of 0.332-0.906. During morning peaks (ρ = 0.906), 54.2% wait less than 5 minutes and 83.3% less than 10 minutes. Lower utilization in the afternoon (ρ = 0.828) and evening (ρ = 0.766) improves performance to 88.5% and 92.6% under 10 minutes respectively. Weekend operations show similar patterns (service rates 1.093-1.368, ρ = 0.795-0.879). SAF operates with lower utilization (ρ = 0.293-0.831) and service rates (0.72-0.824), achieving better performance (98.4% under 5 minutes during evenings). The model's estimates align well with actual data, particularly for 15-30 minute wait times, though showing conservative estimates for shorter waits.

Estimates —M/M/1+ Regression The integration of regression analysis with the M/M/1 model 9a provides refined performance estimates based on arrival rates per line. The regression model at AUC with weekday morning arrival rate of 0.87 passengers per line estimates a service rate of 1.47 (Reg ρ = 0.899), predicting 57.4% waiting under 5 minutes and 79.8% under 10 minutes. Lower arrival rates in the afternoon (0.81) and evening (0.60) yield service rates of 1.153 and 0.901 (ρ = 0.823, 0.848), with improved predictions for 15-30 minute waits. SAF shows wider variation in arrival rates (0.51-1.33) and service rates (0.658-0.871). Evening periods notably achieve ρ = 0.367 despite higher arrival rates (1.33), predicting 95.4% under 5 minutes. The regression model aligns well with M/M/1 estimates while providing per-line arrival insights.

Cluster				Class (ArrRate perLine)	Reg Serv Rate	Reg ρ	Estimated Performance (M/M/1)					
							5m	10m	15m	20m	25m	30m
AUC	Week Day	NIGHT	23:00 - 07:00	0.03	0.116	0.295	80.4%	87.0%	91.4%	94.3%	96.2%	97.5%
		MORNING	07:00 - 11:00	0.87	1.470	0.899	57.4%	79.8%	90.4%	95.5%	97.8%	99.0%
		AFTERNOON	11:00 - 17:00	0.81	1.153	0.823	70.4%	89.3%	96.2%	98.6%	99.5%	99.8%
		EVENING	17:00 - 23:00	0.60	0.901	0.848	57.1%	78.4%	89.1%	94.5%	97.2%	98.6%
	Week-end	NIGHT	23:00 - 07:00	0.00	0.105	0.020	98.8%	99.3%	99.6%	99.7%	99.8%	99.9%
		MORNING	07:00 - 11:00	0.82	1.350	0.890	57.5%	79.7%	90.3%	95.4%	97.8%	98.9%
		AFTERNOON	11:00 - 17:00	0.82	1.145	0.816	71.5%	90.0%	96.5%	98.8%	99.6%	99.9%
		EVENING	17:00 - 23:00	0.66	1.014	0.857	58.6%	80.0%	90.3%	95.3%	97.7%	98.9%
SAF	Week Day	NIGHT	23:00 - 07:00	-	-	-	-	-	-	-	-	-
		MORNING	07:00 - 11:00	0.52	0.748	0.800	62.1%	82.0%	91.5%	96.0%	98.1%	99.1%
		AFTERNOON	11:00 - 17:00	1.07	0.871	0.583	90.5%	98.5%	99.7%	100.0%	100.0%	100.0%
		EVENING	17:00 - 23:00	1.33	0.658	0.367	95.4%	99.4%	99.9%	100.0%	100.0%	100.0%
	Week-end	NIGHT	23:00 - 07:00	0.06	0.146	0.485	66.7%	77.1%	84.3%	89.2%	92.6%	94.9%
		MORNING	07:00 - 11:00	0.53	0.697	0.690	76.5%	92.0%	97.3%	99.1%	99.7%	99.9%
		AFTERNOON	11:00 - 17:00	0.51	0.773	0.852	51.9%	72.8%	84.7%	91.3%	95.1%	97.2%
		EVENING	17:00 - 23:00	0.72	0.836	0.678	82.4%	95.4%	98.8%	99.7%	99.9%	100.0%

(a) Estimates —M/M/1+ Regression

Cluster				Avg # of Servers	Arrival Rate	Est Serv Rate	Arr Rate / Server	Serv Rate / Server
AUC	Week Day	NIGHT	23:00 - 07:00	1.08	0.034	0.103	0.032	0.095
		MORNING	07:00 - 11:00	1.52	1.320	1.457	0.867	0.957
		AFTERNOON	11:00 - 17:00	1.17	0.949	1.146	0.810	0.978
		EVENING	17:00 - 23:00	1.28	0.764	0.998	0.596	0.779
	Week-end	NIGHT	23:00 - 07:00	1.00	0.002	0.000	0.002	0.000
		MORNING	07:00 - 11:00	1.47	1.202	1.368	0.819	0.932
		AFTERNOON	11:00 - 17:00	1.15	0.935	1.139	0.815	0.993
		EVENING	17:00 - 23:00	1.31	0.869	1.093	0.664	0.835
SAF	Week Day	NIGHT	23:00 - 07:00	-	-	-	-	-
		MORNING	07:00 - 11:00	1.16	0.598	0.720	0.516	0.621
		AFTERNOON	11:00 - 17:00	0.47	0.508	0.747	1.069	1.572
		EVENING	17:00 - 23:00	0.18	0.242	0.824	1.331	4.538
	Week-end	NIGHT	23:00 - 07:00	1.12	0.071	0.158	0.063	0.141
		MORNING	07:00 - 11:00	0.90	0.481	0.610	0.533	0.676
		AFTERNOON	11:00 - 17:00	1.30	0.658	0.763	0.508	0.589
		EVENING	17:00 - 23:00	0.78	0.567	0.753	0.722	0.959

(b) Predicted Mean Number of Servers

Figure 9: Results Tables

Predicted Mean Number of Servers The final analysis examines the relationships between service rates, arrival rates, and server efficiency on a per-server basis 9b. At AUC weekday mornings, 1.52 servers process 1.32 passengers/minute (service rate 1.457), achieving 0.867 arrivals and 0.957 service rate per server. Afternoons maintain efficiency with 1.17 servers handling 0.949 arrivals/minute (0.81 arrivals, 0.978 service rate per server), while evenings show slight decreases with 1.28 servers processing 0.764 arrivals/minute (0.596 arrivals, 0.779 service rate per server). SAF demonstrates higher variability, with weekday afternoons using 0.47 servers for 0.508 arrivals/minute (1.069 arrivals, 1.572 service rate per server). Evening efficiency peaks with 0.18 servers handling 0.242 arrivals/minute (1.331 arrivals, 4.538 service rate per server). Weekend operations show similar patterns with lower efficiency metrics.

Prediction Results of Waiting Time Then we come to our prediction result of the waiting time 10. At AUC during weekday mornings, with an arrival rate per server of 0.867 and regression service rate of 1.470, the model predicts a wait time in line of 6.02 minutes, closely matching the actual wait time of 6.64 minutes. The system wait time prediction of 6.70 minutes also aligns well with the actual 7.33 minutes. Similar accuracy is observed during afternoon periods, where predicted wait times in line (4.03 minutes) and in system (4.89 minutes) correspond well with actual measurements of 4.20 and 5.07 minutes respectively.

At SAF, the model's predictions show varying degrees of accuracy with more notable differences in some periods. During weekday morning operations, with a lower arrival rate per server of 0.516 and regression service rate of 0.748, the predicted wait time in line (5.35 minutes) underestimates the actual wait time (6.81 minutes). The most significant variations appear during weekend afternoon periods at SAF, where the model predicts 7.45 minutes in line compared to actual 8.25 minutes, and 8.75 minutes in system versus actual 9.56 minutes. These variations might be attributed to the higher variability in operating conditions at SAF, and the data from SAF are not large enough.

Cluster				Arr Rate / Server	Reg Serv Rate	Actual Wait Time in Line	Predicted Wait Time in Line	Actual Wait Time in System	Predicted Wait Time in System
AUC	Week Day	NIGHT	23:00 - 07:00	0.032	0.116	4.84	3.60	14.57	12.22
		MORNING	07:00 - 11:00	0.867	1.470	6.64	6.02	7.33	6.70
		AFTERNOON	11:00 - 17:00	0.810	1.153	4.20	4.03	5.07	4.89
		EVENING	17:00 - 23:00	0.596	0.901	3.28	6.21	4.27	7.32
	Week-end	NIGHT	23:00 - 07:00	0.002	0.105	-	0.19	-	9.71
		MORNING	07:00 - 11:00	0.819	1.350	5.30	6.02	6.03	6.76
		AFTERNOON	11:00 - 17:00	0.815	1.145	4.03	3.88	4.91	4.76
		EVENING	17:00 - 23:00	0.664	1.014	3.55	5.90	4.47	6.88
SAF	Week Day	NIGHT	23:00 - 07:00	-	-	-	-	-	-
		MORNING	07:00 - 11:00	0.516	0.748	6.81	5.35	8.20	6.69
		AFTERNOON	11:00 - 17:00	1.069	0.871	2.84	1.61	4.18	2.75
		EVENING	17:00 - 23:00	1.331	0.658	0.50	0.88	1.72	2.40
	Week-end	NIGHT	23:00 - 07:00	0.063	0.146	5.18	6.46	11.52	13.30
		MORNING	07:00 - 11:00	0.533	0.697	6.14	3.20	7.78	4.63
		AFTERNOON	11:00 - 17:00	0.508	0.773	8.25	7.45	9.56	8.75
		EVENING	17:00 - 23:00	0.722	0.836	4.04	2.51	5.37	3.71

Figure 10: Prediction Results

And in plot 11 depicts the relationship between observed average wait times and predicted average wait times, with the time one passenger spends in line and in system. WE can see in the plot that all the observed points fall relatively close to our prediction line, suggesting the model fits quite well.

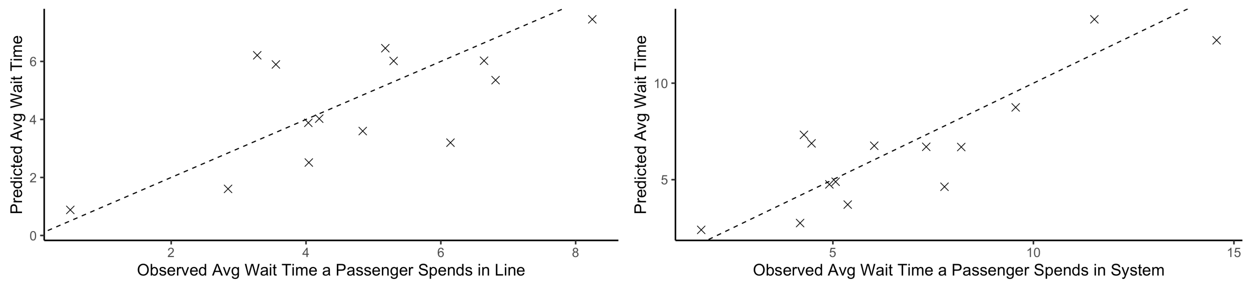

Figure 11: Prediction of Wait Time

4 Conclusion

In this project, we are aiming to comprehensively analyze the operations of the Passenger Boarding System (PBS) implemented at Borealian's four major airports. It focuses on providing the Border Agency Service Authority (BASA) with insights and decision support to enhance screening efficiency and passenger experience.

The background section introduces the implementation of PBS, claiming it will be based on the four given comprehensive datasets for in-depth research. The data dictionary and visualization analysis section thoroughly examines the distribution of passenger wait times, patterns of scheduled flight departures, and trends in wait times by month, time of day, and day of the week. These findings establish a solid foundation for subsequent predictive modeling. Here we made the conclusion that only two datasets with meaningful data will be applied.

Finally, the third part of the project applied a queuing theory framework to forecast key performance indicators such as passenger arrival rates, number of security checkpoints, and passenger wait times. These predictive results will provide BASA with targeted business insights and support to optimize PBS operations and enhance the overall passenger experience. In conclusion, this project offers a comprehensive, data-driven approach to support improvements in the airport PBS system through in-depth analysis and predictive modeling, enabling more efficient operations and greater passenger satisfaction. As the research is based on real-life datasets, the outcomes will assist people in dealing with real-life problems, here including better-allocating screening resources, reducing passenger queuing times, elevating airport service levels, and fostering a positive image for BASA and Borealian airports.

5 Team Contribution

- Kewei Zhang: Text of Section 3
- Junyu Chen: Text of Section 1,2
- Wenzhe Zhang: Code and implementation of Section 2,4,3
- Max Guthrie: Code and implementation of Section 2

References

[1] Patrick Boily. Basics of queueing theory. basics.pdf.

[2] Patrick Boily. Catsa-and-queueing-systems. casta.pdf.

[3] Ehssan Ghashim Patrick Boily. Queueing system. book.pdf.