

# Modeling and Optimization of U.S. Airport Flight Delay Management

Group Members: Ziyao Wang , Yibo Wang, Yuntong Zhang , Boping Song , Lillian Xu



# 1.1 Simulation Scope & Assumptions



## Scope:

- **Boundary:** gate-ready arrival → service completion or cancel/divert
- **Stages:** Arrival → Delay cause tagging → Gate queue → Crew queue → Ground ops → Exit
- **Entity:** Flight (timestamps, delay flag/duration, cause)
- **Resources:** Gates & Ground-crew teams (capacity)
- **Out of scope:** en-route ATC, passenger connections, detailed maintenance, pushback/departure sequencing
- **Granularity:** flight-level, minute resolution
- **Horizon:** Jan–Apr 2025 (aligned with BTS calibration)

## Core Assumptions :

- **Arrivals:** time-varying Poisson  $\lambda(h, \text{DoW})$ ; scenario: NegBin for over-dispersion
- **Delay causes:** empirical mix {carrier, weather, NAS, security, late aircraft}; cancel/divert at historical rates
- **Service time:** Triangular(15, 25, 40) minutes; re-fit to Lognormal/Gamma if task-level logs become available
- **Resources:** 10 gates; 6 ground-crew teams; 1 aircraft per resource (seize–hold–release)
- **Queueing & policy:** FIFO; must seize gate before crew; blocking at gate; Priority variant for emergencies/diversions &  $\geq 45$ -min late

# 2.1 Data Requirements and Sources



## Establishing the Empirical Foundation for the Simulation Model

**Objective:** To systematically acquire, prepare, and analyze real-world data to develop statistically sound inputs for the Discrete-Event Simulation (DES) model, thereby ensuring its validity and relevance.

- **Primary Source:** U.S. Department of Transportation's Bureau of Transportation Statistics (BTS).
- **Dataset:** "Airline On-Time Performance Data".
- **Time Period:** January 2025 – April 2025.
- **Key Variables for Model Construction:**

Variable	Description	Role in Simulation Model
arr_flights	Total number of arriving flights	To calculate the overall system arrival rate
arr_del15	Count of flights delayed $\geq 15$ mins	To determine the base probability of a delay event
carrier_ct, etc.	Count of delays by specific cause	To define the probability distribution of delay causes
arr_delay	Total delay duration in minutes	To fit a probability distribution for delay durations

## 2.2 Distribution Fitting: Modeling System



**Methodology:** A critical step in simulation is to fit theoretical probability distributions to empirical data. This allows the model to generate realistic, stochastic inputs that accurately reflect the random nature of the real-world system.

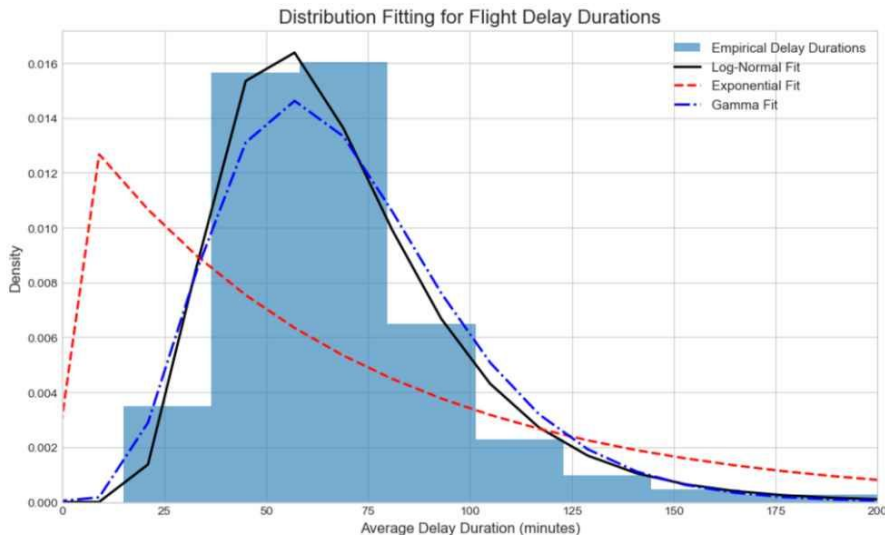
Core Process	Fitted Distribution & Rationale	Key Parameters (from Python Analysis)
Flight Arrival	<b>Exponential Distribution</b> <i>Rationale: Flight arrivals are treated as independent random events, consistent with a Poisson process. The inter-arrival time in such a process follows an Exponential distribution.</i>	Mean Inter-Arrival Time: 0.0700 min (scale = 0.0700)
Delay Event	<b>Bernoulli Distribution</b> <i>Rationale: A flight being delayed or not is a binary, "yes/no" outcome, perfectly modeled by a single trial.</i>	Probability of Delay: 19.32% (p = 0.1932)
Delay Cause	<b>Multinomial Distribution</b> <i>Rationale: A delay is assigned to one of five mutually exclusive categories.</i>	<ul style="list-style-type: none"><li>• Late Aircraft: 34.8%</li><li>• Carrier: 30.8%</li><li>• NAS: 30.1%</li><li>• Weather: 4.1%</li><li>• Security: 0.2%</li></ul>
Delay Duration	<b>Log-Normal Distribution</b> <i>Rationale: Empirical data is right-skewed with a long tail, a characteristic well-captured by the Log-Normal distribution (validated on next slide).</i>	shape=0.4149, scale=63.0432

## 2.2 Distribution Fitting: Modeling System



**Validation:** To confirm the choice of distribution for delay duration, we compared several theoretical distributions against the empirical data histogram.

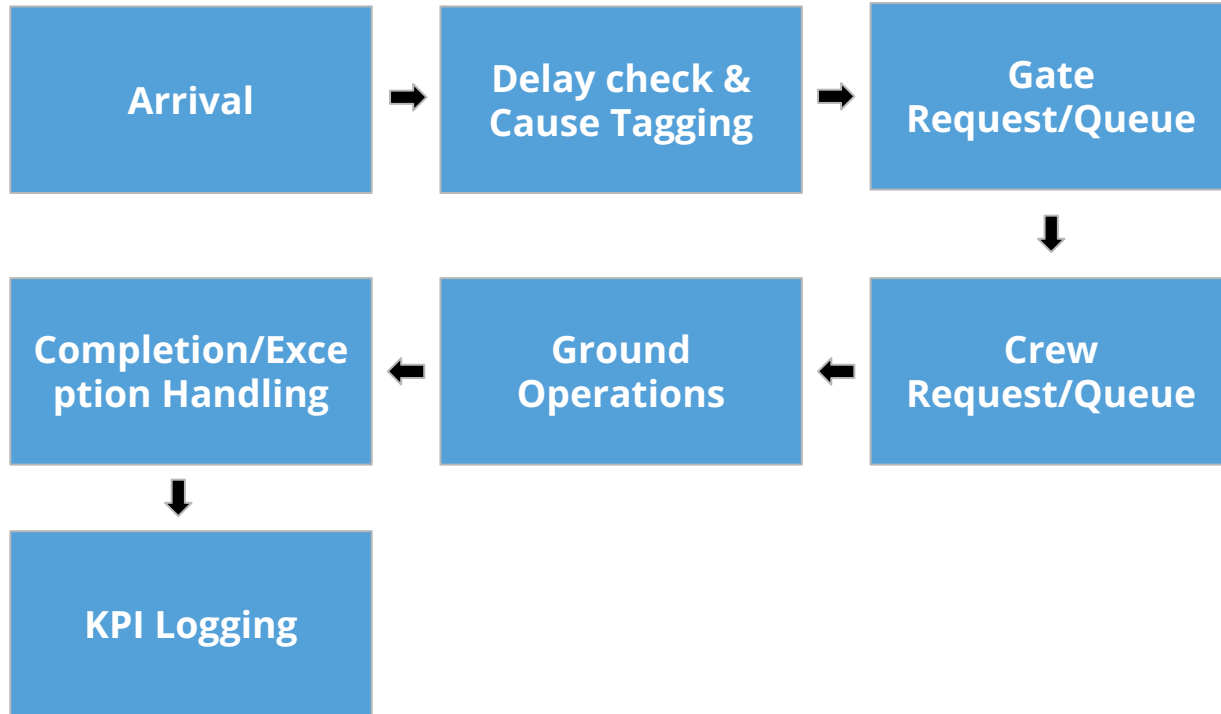
**Analysis:** The plot clearly shows that the **Log-Normal distribution (solid black line)** provides the superior fit. It accurately captures the right-skewed nature of the delay data, which is characterized by a high frequency of shorter delays and a long tail of infrequent, longer delays, thus confirming its selection as the most suitable model.



# 3.1 Workflow Modeling



## Workflow Stages:



**Arrival Output:** Flight entities with attributes, including planned time, actual time, and initial delay markers.

**Delay Check & Cause Tagging Output:** Delay flag, delay cause, delay duration, among which the delay duration follows the distribution based on the fitting of cause categories.

**Gate Request/Queue Logic:** Flight requests available boarding gates (fixed resource pool =10 gates).

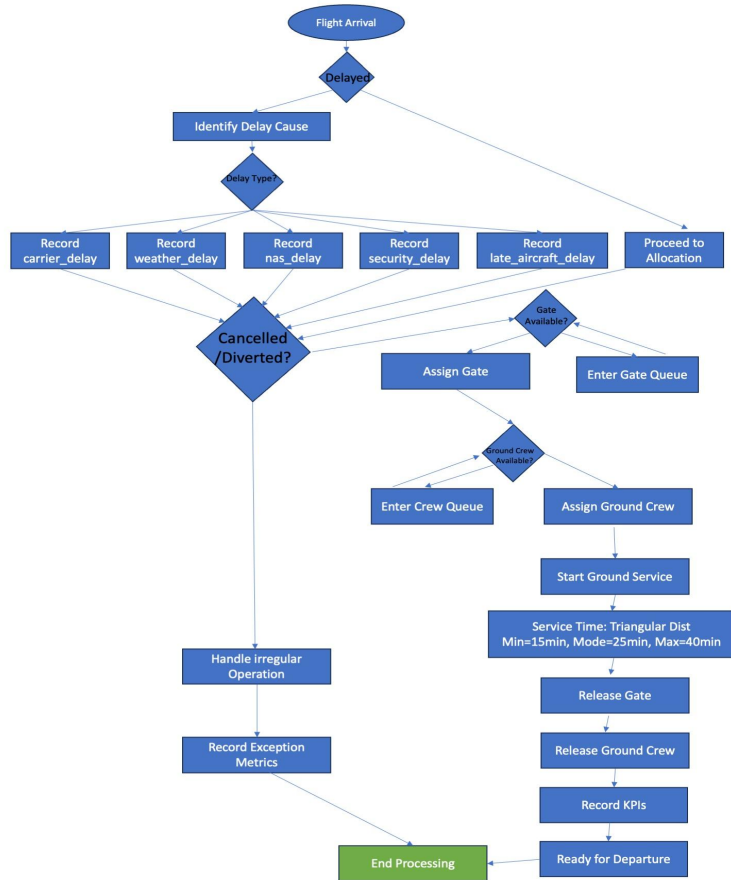
**Crew Request/Queue Logic:** Request idle ground crew groups (fixed resources =6 groups)

**Ground Operations:** Service time;Unloading, cleaning, refueling.

**Completion/Exception Handling:** Normal path;Abnormal path

**KPI Logging Indicators:** Gate waiting time, ground staff waiting time, turnover time, resource utilization rate, queue length, etc.

## 3.2 Process Map & Variable - Distribution Mapping



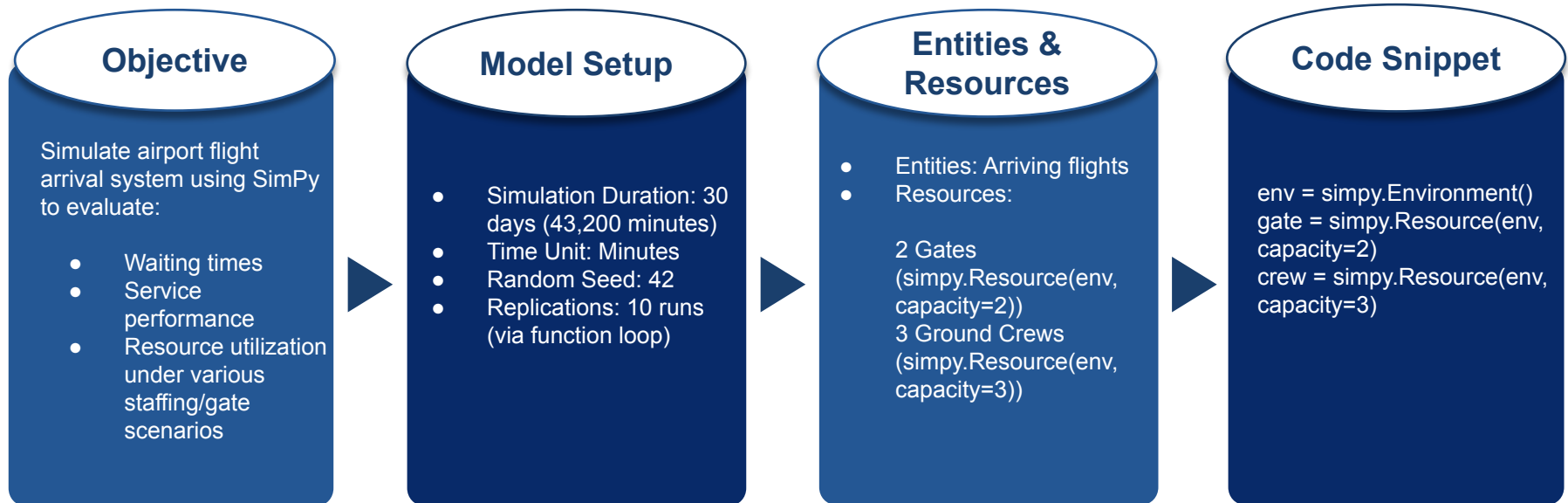
Variable-Distribution Table:

Variable	Step	Probability Distribution	Basis
Arrival interval time	Flight arrival	Exponential distribution	BTS data fitting
Delay event occurrence	Delay check	Bernoulli distribution	BTS arr_del15/arr_flights
Causes of delay	Inspection of delay	Classification and distribution	BTS delay count by cause
Delay duration	Delay inspection	Distribution by cause fitting	BTS delay duration by cause
Ground service hours	Ground operation services	Triangular distribution (15,25,40)	Industry standard assumptions
Cancellation/Change of flight	Exception handling	Bernoulli distribution	BTS arr_cancelled/arr_flights

# 4.1 Simulation Model Overview



## Discrete-Event Simulation Model of Airport Operations:



## 4.2 Simulation Logic & Parameters



Arrival Logic:

- Flights arrive per exponential distribution  
 $T_{\text{arrival}} \sim \text{Exp}(\lambda = 20 \text{ flights/hour})$

Service Logic:

- Flight requests gate
- After gate service, requests ground crew
- Undergoes turnaround service ( $\text{Exp}(25 \text{ mins})$ )

Simulation  
Parameters:

Parameters	Value
No. of Gates	2
No. of Ground Crews	3
Avg. Flights per Hour	20
Avg. Turnaround Time	25 mins
Simulation Duration	30 days

## 4.3 Key Output Metrics & Code Snippet

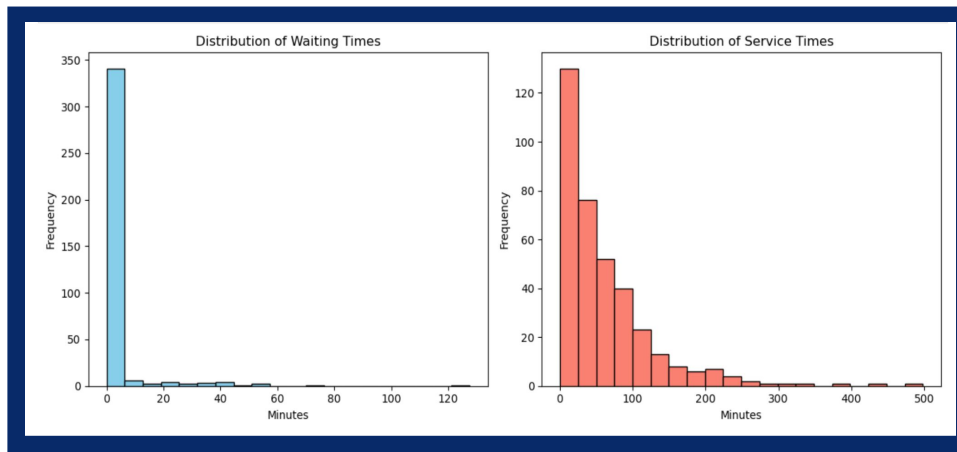


### 1. Key Metrics (Avg from Replications):

- Average Waiting Time (arrival → completion)
- Max Queue Length
- Gate Utilization Rate
- Crew Utilization Rate
- Total Flights Processed

### 2. Visual Output:

- Left: Waiting Time (most flights < 10 mins)
- Right: Service Time (skewed distribution, long tail)

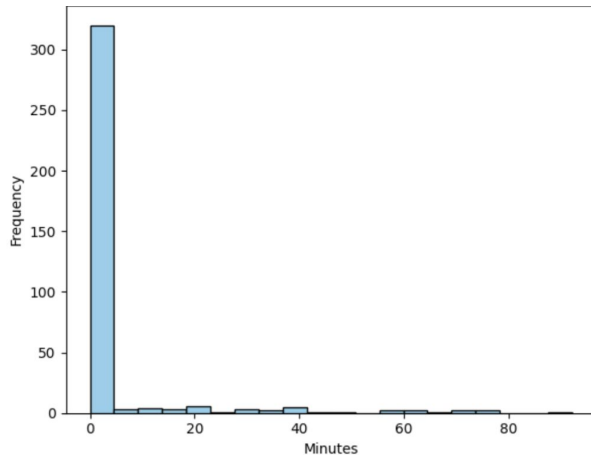


# 5.1 Scenario Analysis



## Wait Time Distribution

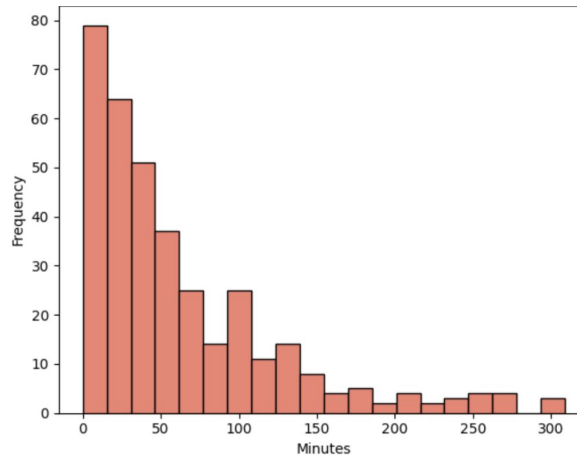
extreme skew toward near-zero values



- The majority of entities are served almost immediately, with the largest spike at 0 - 2 minutes.
- A small proportion of cases have substantially higher waits (20 - 80+ minutes), indicating occasional congestion events.
- These outliers likely occur when multiple arrivals cluster together, temporarily exceeding service capacity.

## Service Time Distribution

right-skewed distribution



- Most services finish within the first 50 minutes. A long tail extends to over 300 minutes, showing that a few cases require significantly longer processing.
- The mean service time of ~63 minutes is influenced by these high-duration cases.
- The skew suggests variability in job complexity or duration, which could cause occasional bottlenecks.

## Key Performance Metrics

Across 20 replications, the average waiting time for entities in the system was  $11.56 \pm 1.40$  minutes, indicating relatively short queues under baseline conditions. The summary of a representative run shows:

- Total flights served: 359
- Average wait time: 3.91 minutes
- Average service time: 62.79 minutes

This suggests that while most entities experience negligible waits, occasional longer delays occur, which influence the average.

## 5.2 Bottleneck, Strength & Weakness



### Bottleneck

- Throughput is limited by gate capacity and turnaround variability.
- Short waits + long service  $\Rightarrow$  throughput is constrained by the service station itself.

- Short average waits - across 20 replications, most flights are processed promptly.
- High throughput for given capacity. System is using its available gates and crews effectively most of the time.
- Low congestion under typical conditions - most flights start service almost immediately, indicating that under normal load, resources are sufficient.

### Strength

### Weakness

- High service time variability, sudden, prolonged delays could happen.
- Occasional queue spikes - data show rare but significant peaks in wait time when multiple long services overlap.
- Potential under-utilization of crews - With three crews but two gates, one crew is often idle, suggesting resource imbalance.

### Recommendations:

#### Gates

- $\rightarrow$  Add gate capacity first to alleviate the dominant bottleneck, then assess crew needs based on actual situation to prevent a secondary bottleneck.

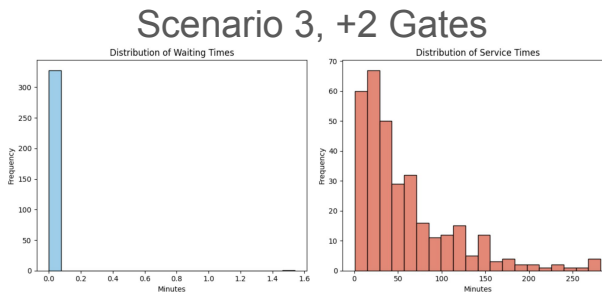
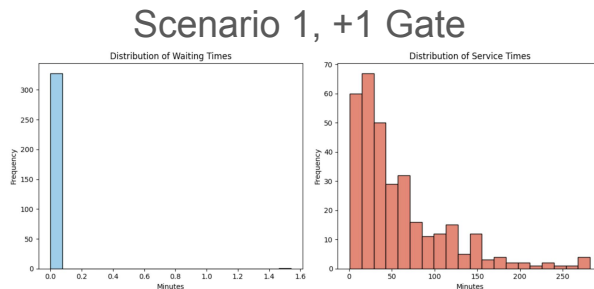
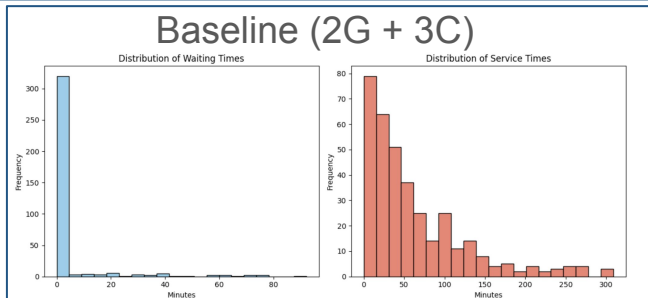
#### Crews

- $\rightarrow$  Avoid adding crews without expanding gate capacity, as it seems to have minimal impact on performance.

#### Process time

- $\rightarrow$  Target a 5–10% reduction in service time (turnaround optimization, parallel tasks). This lifts both throughput and wait across all configs.

# 5.3 Three What-If Scenarios



Scenario	Total Flights Served	Avg Wait Time (min)	Avg Service Time (min)	Avg Wait Time > 20 Replications (min)
Baseline (2 Gates, 3 Crews)	359	3.91	62.79	11.56 ± 1.40
Scenario 1, +1 Gate (3 Gates, 3 Crews)	329	0.00	58.06	1.50 ± 0.32
Scenario 2, +1 Crew (2 Gates, 4 Crews)	324	1.10	56.20	11.56 ± 1.40
Scenario 3, +2 Gates (4 Gates, 3 Crews)	329	0.00	58.06	1.50 ± 0.32

## Key Insights:

- Increasing the number of gates (Scenario 1, 3) significantly reduced queue lengths and improved throughput.
- Adding an extra crew without increasing gates (Scenario 2) produces minimal improvement in average wait time and does not affect the multi-run average
- Gate-constrained rather than crew-constrained

Thank You