Peak-Hour Traffic Flow and Bus Stop Optimization:

A Simulation Study of

Pak Shek Kok

By

YIN Xuan Shu Ran

(24416541)

A Thesis Submitted in Partial Fulfilment

of the Requirements for the Degrees of

Master of Sciences in Operational Research and Business Statistics

Hong Kong Baptist University

and

Master of Sciences in Business Analytics

University of Kent (Canterbury)

4 July 2025

(Word Count: 6005Words)

Acknowledgement

I am profoundly grateful to my supervisor, Tian WU. Her well-thought-out insights provided in our weekly meetings have offered crucial guidance for my theoretical framework and modeling methods. Her patient guidance, especially in dealing with the complex modeling issues of data retrieval and analysis, played a significant role in shaping this work. I also want to express my sincere gratitude to Simon. His meticulous observation of bus route patterns and his inspiring discussions have sparked valuable inspiration and enriched my research. I also sincerely thank my family members. Their unwavering encouragement and emotional support gave me the courage to persevere and overcome countless late-night efforts. Additionally, I express my most sincere gratitude to the Operations Research and Business Statistics program. This year's study has benefited me greatly, enabling me to acquire valuable knowledge and broaden my horizons.

Signature: 文本

AI 生成的内容可能不正确。

(Student Name BLOCK LETTER)

Date: 文本

AI 生成的内容可能不正确。

Peak-Hour Traffic Flow and Bus Stop Optimization:

A Simulation Study of

Pak Shek Kok

YIN Xuan Shu Ran

(24416541)

Abstract

This study employed a discrete-event simulation technology based on Python to optimize the average bus waiting time at Hong Kong’s St MARTIN and CHONG SAN ROAD Stations. Four configurations are compared: the current setup (two berths per station), designated routes stopping at one station, a merged four-berth superstation, and a single two-berth station. Grounded in queuing theory, the simulation analyzes peak-hour performance. Results highlight that designated route allocation significantly reduces bus queuing time while maintaining practical feasibility. This study provides data-driven insights for the optimization of bus stations in Hong Kong, proposes practical suggestions, and explores future research directions on dynamic scheduling.

Table of Contents

[List of Figure iv](#_Toc202542460)

[List of Table v](#_Toc202542461)

[1. Introduction 1](#_Toc202542462)

[1.1 Background and research objectives 1](#_Toc202542463)

[2. Literature Review 3](#_Toc202542464)

[3. Data Collection and Analysis 5](#_Toc202542465)

[3.1 Data Collection 5](#_Toc202542466)

[3.2 Determination of Parameters 7](#_Toc202542467)

[4. Methodology 8](#_Toc202542468)

[4.1 Theoretical Framework 8](#_Toc202542469)

[4.2 Implementation of the model 10](#_Toc202542470)

[4.3 Simulation Design 11](#_Toc202542471)

[5. Data visualization and interpretation 14](#_Toc202542472)

[6. Data Insights and conclusions 19](#_Toc202542473)

[6.1 Research limitations and Improvements 21](#_Toc202542474)

[7. References 24](#_Toc202542475)

# List of Figure

[Figure 1 Diagram of Adjacent Bus Stops A and B with Shared Bus Lane 3](#_Toc202542476)

[Figure 2 Merged Four-Berth Superstation 13](#_Toc202542477)

[Figure 3 Merged Two-Berth Station 13](#_Toc202542478)

[Figure 4 Comparison of Average Bus Queuing Time Across Scenarios 14](#_Toc202542479)

[Figure 5 Distribution of Passenger Waiting Time and Bus Queue Time Across Different Scenarios 15](#_Toc202542480)

[Figure 6 Distribution of Queuing Time Across Scenarios 16](#_Toc202542481)

[Figure 7 Average Bus Queuing Time 16](#_Toc202542482)

[Figure 9 Bus Queue Time Distribution Across Different Scenarios 17](#_Toc202542483)

[Figure 10 Violin plots for Bus Queue Time Distribution 18](#_Toc202542484)

[Figure 11 Improvement Rate in Average Bus Queue Time (vs current time) 20](#_Toc202542485)

# List of Table

[Table 1 The Bus Stop IDs of the Two Stations 6](#_Toc202542486)

[Table 2 Bus Routes and Schedules 6](#_Toc202542487)

[Table 3 Simulation Parameters and Values 7](#_Toc202542488)

[Table 4 Improvements of Queue Time 19](#_Toc202542489)

# Introduction

## Background and research objectives

St. Martin Road Station and Chong San Road Station serve as essential components of this network, each marked by its distinct geographical and operational features. These stations handle significant daily passenger traffic, especially during peak hours, due to their strategic placement near economic centers and residential neighborhoods. Their proximity to housing areas results in exceptionally high passenger volume, particularly during the morning rush on weekdays. At present, congestion during these early peak periods is notably severe. This analysis aims to reflect the current state and identify potential areas for improvement to address these challenges.

Hong Kong’s St. Martin Road Station and Chong San Road Station are situated approximately 52 meters apart. During the morning peak hours, approximately 800 passengers per hour arrive at each station, with buses arriving sequentially based on their scheduled departure times. In instances where both berths at a station are fully occupied, buses are required to queue and wait for an available space. Most bus routes overlap and stop at both stations, causing traffic congestion during morning peak hours. This affects passenger travel efficiency and waste resources. This study aims to optimize bus routes through simulation, proposing three improvement schemes compared against the current setup. These methods focus on bus queuing time at St. Martin and Chong San Road Stations. Four configurations are evaluated: the existing setup (two berths per station), designed routes stopping at only one station, a consolidated four berth superstation, and a single two-berth station.

The study examines the operational dynamics of Kowloon Motor Bus (KMB) services, focusing on 12-meter buses operating at a designated bus stop with a capacity to accommodate two buses simultaneously. Unlike real-world scenarios where marked lines or designated passenger zones constrain boarding, this analysis assumes a simplified framework where buses dock at predefined positions without such restrictions. The investigation centers on managing a total of eight hundred passengers, initially distributed evenly across available bus routes, with the potential for passenger numbers to grow overtime due to ongoing arrivals. Each newly arriving passenger is systematically assigned to the waiting queue of a specific route, identified by its unique route identifier (e.g., route["id"]). This approach allows for a detailed exploration of queue accumulation and passenger flow under varying operational conditions.

This stochastic process results in fluctuating queue lengths for each route over time. As passengers continue to arrive, the waiting queue for a given route may either grow significantly or remain relatively stable, depending on the arrival rate and the frequency of bus departures. The accumulation of passengers in these queues is directly influenced by the interval between consecutive bus arrivals. When buses are scheduled with long intervals between services, the number of passengers queuing tends to increase, leading to a higher volume of individuals boarding each bus as it arrives. This scenario can strain the system, particularly if the stop’s capacity to handle two buses is fully utilized, potentially causing delays or overcrowding. Conversely, shorter intervals between bus arrivals reduce the buildup of waiting passengers, thereby decreasing the number of individuals boarding per bus and alleviating pressure on the stop’s infrastructure.

The relationship between bus arrival frequency and queue dynamics serves as a pivotal element in this analysis. When the interval between bus arrivals is extended, it provides greater opportunity for passenger arrivals to build up, particularly during peak hours. This accumulation is accentuated by the exponential distribution, which can produce clusters of arrivals, leading to a surge in waiting passengers. Furthermore, the concentration of passengers exacerbates boarding times, as more individuals seek to enter each bus. This extended dwell time at the stop, in turn, prolongs the overall duration of bus occupancy, contributing to increased congestion on the roadway. Such dynamics underscore the necessity of optimizing arrival schedules to mitigate both queue growth and traffic disruptions, especially under high-demand conditions.

The image below is a schematic diagram illustrating the layout of two adjacent bus stops, labeled Bus Stop A and Bus Stop B. Each bus stop is equipped with two berths (Berth 1 and Berth 2) for buses docking. A bus lane runs horizontally, with the arrows indicating the direction of bus traffic.

图形用户界面

AI 生成的内容可能不正确。

Figure Diagram of Adjacent Bus Stops A and B with Shared Bus Lane

The objectives are to reduce bus queuing time, enhance road traffic efficiency, and lower the average passenger waiting time. Attaining these objectives holds considerable importance, as it plays a key role in alleviating traffic congestion during morning peak hours. This improvement can enhance satisfaction and boost the reliability of the public transportation system.

# Literature Review

The study of bus stop interactions has evolved significantly over the years, driven by the need to enhance mobility and efficiency in public transport systems.

Early research, such as that by Frye (1966) and Brown (1996), laid the foundation by identifying the journey chain in public transport, which includes walking to a stop, waiting, boarding, traveling, alighting, and walking to the destination. [1][2]This framework, further refined by Tyler and Brown (1997), emphasizes mobility as a broader concept encompassing accessibility, access, and movement.[3]Accessibility, as defined, involves the ease of reaching stop locations and includes factors like pedestrian infrastructure and service reliability (Tyler, 1996a). [4]Access pertains to the physical process of boarding and alighting, heavily influenced by bus stop design, while movement addresses in-vehicle travel efficiency (Tyler and Brown, 1997). [3] This underscores the critical importance of bus stop design, as poorly conceived layouts and ineffective allocation of bus routes can impede the overall system performance, leading to congestion and diminished passenger satisfaction.

Operational insights reveal that bus stops significantly affect journey times. Lobo (1997) noted that stop times constitute approximately half of the total journey time between termini, [5]a finding echoed by Gardner et al. (1991), who reported mean delays ranging from 45 to 90 seconds at stops compared to 15 seconds at junctions. [6] This suggests that bus stops act as primary bottlenecks, necessitating targeted design improvements to facilitate passenger transfers and reduce delays (Tyler, 1992). [7]

Previous studies have applied simulation-based optimization methods to address various challenges in public transport systems. Obiora A. Nnene, Johan W. Joubert, and Mark H. P. Zuidgeest (2023) proposed a method combining activity-based simulation with multi-objective metaheuristic optimization to design transit networks while balancing conflicting stakeholder objectives.[8]Haodong Yin, Jianjun Wu (2019), and colleagues developed an agent-based simulation and genetic algorithm model to optimize passenger flow guidance in urban rail networks, reducing congestion and improving efficiency. [9]Mingye Zhang et al. (2023) introduced a bus scheduling method integrating full-stop and skip-stop services to minimize total passenger travel time on overlapping routes.[10] These investigations offer significant perspectives on enhancing public transportation networks by leveraging simulation and optimization strategies.

In conclusion, the literature reveals an evolution from foundational explorations of journey chains to the development of intricate simulation models, with bus stop design emerging as a critical element in enhancing public transportation efficiency. Although significant progress has been made, this field requires continuous empirical verification and model expansion to address the complex dynamics of urban traffic.

The matter under consideration involves devising simulations that accurately reflect the present circumstances to mirror the most genuine scenario, a task that holds significant importance for advancing research. This approach also furnishes essential theoretical grounding to enhance the existing conditions.

# Data Collection and Analysis

This study employs a combined approach of real-time data collection and simulation to evaluate the performance of various bus stop configurations at St MARTIN ROAD and CHONG SAN ROAD stations in Hong Kong. This method mainly collects real data regarding the locations of bus stops and arrival times and then simulates the model to obtain the waiting times for passengers and the queueing times for buses in four different operational scenarios. This approach aims to develop a simulation that closely mirrors real-world conditions, providing a foundation for optimizing bus stop operations.

## Data Collection

These two stations each have four Bus Stop IDs, corresponding to different bus routes. Through the query, we can obtain their Bus Stop IDs as shown in the following table. The main source for obtaining the ID of the bus stop is the KMB Open Data Platform, which provides comprehensive information about bus stops and their related routes throughout Hong Kong. Accuracy of the Bus Stop ID has laid a solid foundation for our subsequent research.

Table The Bus Stop IDs of the Two Stations

|  |  |
| --- | --- |
| St MARTIN ROAD Station | CHONG SAN ROAD Station |
| "3F24CFF9046300D9" | "3A7AC3A5F9530786" |
| "33ABA49F0E91A247" | "023E9E5A9E073E1A" |
| "07AB149DAD888683" | "B34F59A0270AEDA4" |
| "4A0ECA0D5AA4CB7E" | "437CE05BCFE6248C" |

Gathering precise and detailed insights into the geographical placement of bus stops and the precise scheduling of bus arrivals is a critical step in this research. Main data source is the Open API of KMB (The Kowloon Motor Bus). The locations of bus stops and the estimated arrival times can be obtained from the public API of KMB, and the data source is processed and visualized.

Table Bus Routes and Schedules

|  |  |  |  |
| --- | --- | --- | --- |
| St. Martin Station | | Chong San Road Station | |
| Bus routes | Bus schedule | Bus routes | Bus schedule |
| 73D | 60min | 73D | 60min |
| 74 | 8:00 | 74 | 8:00 |
| 74D | 60min | 74D | 60min |
| 74P | 8:00,8:15 | 74P | 8:00,8:15 |
| 82D | 7:15 | | |
| 96 | 7:00-8:00 20min | 96 | 7:00-8:00 20min |
| 263C | 7:15 | 263C | 7:15 |
| 271B | 7:10,7:50 | 271B | 7:10,7:50 |
| 272A | 8-10min | 272A | 8-10min |
| 272P | 7:00-7.45 15min | 272P | 7:00-7.45 15min |
| 272X | 25min-30min | 272X | 25min-30min |
| 274P | 8:00-9:00 15min | 274P | 8:00-9:00 15min |
| 907D | 7:05 | 907D | 7:05 |
| A47X | 30min | A47X | 30min |
| 900 | 30min | 900 | 30min |
|  | | 271R | no data |
| 82D | 7:15 |

The above table presents the timetables and routes for these two bus stations, covering data up to 10:00 AM. It is obvious that only a few bus routes at these stations do not overlap, indicating that most bus operations in this area suffer from low efficiency. This redundancy highlights the significant importance of current study.

## Determination of Parameters

Based on the data collected, I have determined that a one-hour period represents the optimal timeframe for conducting simulations. This duration prevents the exclusion of certain bus routes due to an excessively short interval thereby preserving the diversity of bus services. Consequently, the study period is established as the morning hours from 8:30 to 9:30 on a weekday. Conducting one hundred simulations was chosen to ensure a robust analysis, allowing for a comprehensive evaluation of the system’s performance across varied conditions while minimizing the risk of overlooking critical patterns.

Parameter configuration is as follows:

Table Simulation Parameters and Values

|  |  |  |
| --- | --- | --- |
| Parameters | Value | Description |
| SIM\_TIME | 3600 | One-hour simulation, close to the morning rush hour |
| BOARDING\_TIME | 4 | The boarding time for each passenger (in seconds) |
| CLEARANCE\_TIME | 3 | Departure preparation time for each vehicle (in seconds) |
| BERTHS\_PER\_STOP | 2/4 | The number of parking Spaces |
| PASSENGER\_RATE | 800 | The number of passengers arriving per hour during the morning rush hour |
| NUM\_SIMULATIONS | 100 | number of simulations |

The model initiates with the configuration of essential parameters tailored to reflect a one-hour simulation duration, aligning with early peak traffic conditions. A fixed boarding time of 40 seconds per passenger and a clearance time of 30 seconds per bus departure are established to account for the time required for passenger embarkation and vehicle preparation, respectively. The bus stop is designed to accommodate two berths simultaneously, reflecting the operational capacity of 12-meter Kowloon Motor Bus (KMB) vehicles. An initial passenger arrival rate of 800 individuals per hour is set, distributed evenly across the active routes at each stop, with the simulation iterated 100 times to ensure robust statistical outcomes.

# Methodology

This study examines four distinct scenarios to optimize average bus queuing and passenger waiting times at St MARTIN ROAD and CHONG SAN ROAD Stations in Hong Kong, utilizing Python-based simulation.

## Theoretical Framework

A key assumption in this research is that passenger arrivals conform to a Poisson distribution, reflecting the seemingly random yet statistically consistent behavior of commuter flows during peak hours.

Based on the established rate of 800 arrivals per hour, it is estimated that approximately 13 passengers will arrive per minute during peak hours.

: the simulation time when bus secures a berth

: bus arrive time

：The total number of bus that arrive at bus stop

The bus queue time is the duration a bus waits to secure a berth at the stop when all berths are occupied. This is calculated as the difference between the time a bus arrives and the time it secures a berth.

A Queuing Model is a mathematical model used to analyze the performance of queuing systems.

= Passenger arrival rate

= Bus service rate

= The number of buses operating during the morning rush hour

= The average waiting time for passengers

= The average number of queues in the system

= System load ratio

In the case of multiple servers (i.e. multiple buses), the M/M/c model can be used to calculate the average waiting time:

The formula for (average number of passengers in the queue) is given as:

After calculating , we obtain ​, which represents the average waiting time in the queue for passengers.

Then, the total system time for a passenger (including both waiting and boarding time) is:

: The average waiting time for a passenger in the queue before boarding.

: The service time for a passenger once they board, which can be approximated as half of the bus departure interval.

If buses depart at fixed intervals of minutes and passengers arrive randomly, the average waiting time can be calculated using a simple mean formula:

Explanation:

* Since passengers arrive randomly, their waiting time is uniformly distributed between [0, T]
* The expected (average) waiting time is the mean of this uniform distribution, which is
* This formula applies to fixed-interval bus departures, assuming no peak-hour congestion or variations in bus arrivals.

Longer dwell times increase berth occupancy, leading to higher queue times for subsequent buses.

## Implementation of the model

The implementation of the simulation model for analyzing bus stop operations at St. Martin and Chong San Road stations is structured around a discrete-event simulation framework using the SimPy library, integrated with real-time data retrieval and statistical analysis.

The implementation of the simulation model relies on SimPy, a discrete-event simulation framework written in Python, which facilitates the modeling of complex systems such as bus stop operations. The advantage of SimPy lies in its ability to handle dynamic processes, including the arrivals and departures of buses and passengers. Therefore, it is highly suitable for capturing the interaction at the stations of St. Martin’s Road and Chong San Road stations. The framework enables the definition of processes like passenger boarding, vehicle clearance, and queue management, which are executed based on predefined parameters such as simulation time, boarding time, and berth availability.

To integrate real-time data into the model, SimPy is coupled with an interface to the Open API provided by KMB (Kowloon Motor Bus), a key public transport operator in Hong Kong. This API serves as the primary source for obtaining current information on bus stop locations and estimated arrival times. The process involves periodically querying the API to fetch live data, which is then processed and fed into the simulation environment. For instance, the model retrieves data on bus identifiers and schedules, aligning them with the one-hour study period from 8:30 to 9:30 AM on weekdays. This real-time input ensures that the simulation reflects actual operational conditions, enhancing accuracy in assessing performance metrics such as passenger waiting times and bus delays. The data is structured within SimPy’s event-driven framework, allowing for dynamic adjustments to the simulation as new arrivals are detected, thereby supporting a realistic representation of the system.

Bus operations are simulated through a process that requests access to a berth, records queuing delays, and handles passenger boarding. Upon arrival, a bus occupies a berth and services the accumulated passengers in its route’s queue, with the dwell time calculated as the sum of the clearance time and the boarding time multiplied by the number of waiting passengers. The waiting time for each passenger is calculated as the duration from arrival to boarding and is adjusted based on the average boarding delay. The queuing time for the bus is recorded as the time interval from arrival to obtaining a parking space. These processes run concurrently in the SimPy environment, and the arrival of the bus is triggered by an exponential interval based on the route frequency.

## Simulation Design

The study encompasses a total of four scenarios. One scenario depicts the current operational situation, serving as a baseline for comparison, while the other three are designed for optimization purposes, aiming to reduce the average bus waiting time. These scenarios are compared to evaluate their effectiveness in enhancing operational efficiency, analyzing metrics such as queue lengths, passenger wait times, and overall system performance.

The first scenario leverages actual operational data to assess the current state. Real-time data over a one-hour period was collected, encompassing estimated bus arrival times, expected route deployments. It calculates the average passenger waiting time and bus queuing time at both stations, reflecting their existing two-berth configuration. This serves as the benchmark for evaluating subsequent optimizations.

The second scenario addresses the issue of most buses stopping at both stations by reallocating certain routes. The allocation principle retains the current stopping pattern for unique bus routes at each station. For overlapping routes, their arrival frequencies are aggregated and ranked. Subsequently, these routes are assigned alternately to the two stations based on odd and even frequency rankings, aiming to reduce congestion and streamline service.

Given the proximity of St. Martin Road and Chong San Road (approximately 52 meters), the third scenario proposes consolidating them into a single large station. This reconfiguration increases the berth capacity from two per station to a unified platform with four berths. The simulation explores how this expansion impacts queuing times and berth utilization.

The fourth scenario involves merging St. Martin Road and Chong San Road into a single standard station with only two berths. This setup tests the effects of centralizing all bus services into a constrained platform, evaluating its influence on queuing efficiency under high demand. This is evaluated based on whether it is possible to reduce the waste of resources and eliminate duplicate stations.

These scenarios reveal unique operational dynamics and their influence on queue times, offering diverse perspectives for future efficiency assessments.

We will name the four scenarios as follows to clearly delineate their operational configurations and facilitate comparative analysis within the study.

1. Scenario 1: Baseline with Real data
2. Scenario 2: Route Allocation for Overlapping Services
3. Scenario 3: Merged Four-Berth Superstation
4. Scenario4: Merged Two-Berth Station

表格

AI 生成的内容可能不正确。

Figure Merged Four-Berth Superstation

表格

AI 生成的内容可能不正确。

Figure Merged Two-Berth Station

Preliminary sketches for Scenarios 3 and 4 are provided above to illustrate the proposed station layouts.

# Data visualization and interpretation

图表

AI 生成的内容可能不正确。

Figure Comparison of Average Bus Queuing Time Across Scenarios

The figure illustrates the average queuing times across various scenarios, with the bar heights representing the duration of queuing times. Additionally, error bars are incorporated into the graph, indicating the range or variability of queuing times across different conditions, thus highlighting the fluctuation in performance.

From Figure 4, we can observe that all three methods can to some extent alleviate the problem of bus queueing. Allocating shared routes to specific stops can effectively reduce the congestion of bus queues. If these two stops are combined into a single large stop, it can significantly reduce the bus queueing phenomenon, making the waiting time almost zero. Moreover, merging them into a standard single stop has little impact on the average bus queueing time, perhaps because the bus routes during the morning rush hour are limited and the stop at both stations have no significant effect on the outcome.

When exploring potential strategies to alleviate bus congestion, the adoption of a four-berth superstation, as exemplified in Scenario 3 (Merged Four-Berth Superstation), offers a viable option. This configuration contributes to more stable and reduced bus queue durations, potentially streamlining traffic flow by allowing buses to stop once and utilize four berths efficiently. As can be seen from the figure below, the waiting time for buses is not high in any of these scenarios. Scenario 4 (Merged Two-Berth Station) has a very similar queue time density to scenario 3 (Merged Four-Berth Superstation). And methods three and four might effectively reduce the waiting time for passengers. However, it should be noted that Scenario 2 (Route Allocation for Overlapping Services) might increase the waiting time for passengers.

图表, 散点图

AI 生成的内容可能不正确。

Figure Distribution of Passenger Waiting Time and Bus Queue Time Across Different Scenarios

However, due to the significant difference in the waiting time for buses and passengers, it is difficult to observe the detailed aspects of the waiting time. Therefore, in Figure 6, we only observe the bus queuing time. This means that the height of each box represents the probability density of the values within the corresponding range. Based on the waiting time for buses, we can see that all three methods have improved the situation. Among them, scenario three is the best, with the bus waiting time being almost zero.

图表

AI 生成的内容可能不正确。

Figure Distribution of Queuing Time Across Scenarios

The violin plot shows the range and density of the distribution of queuing times. Due to the small number of 100 samples, it is difficult to discern any patterns. So, 1000 samples were selected to draw the distribution.The median value in Scenario 3 (Merged Four-Berth Superstation) is significantly lower than that in the other scenarios, indicating that this configuration is more effective in reducing queuing time. The differences between Scenario 1 (Baseline with Real data)cand Scenario 2 (Route Allocation for Overlapping Services) are relatively small, indicating that the allocation strategy has not significantly improved the queuing time. In Scenario 2 (Route Allocation for Overlapping Services), the longer tail of the violin plot for Chong San Road indicates that there may be extreme queuing situations.

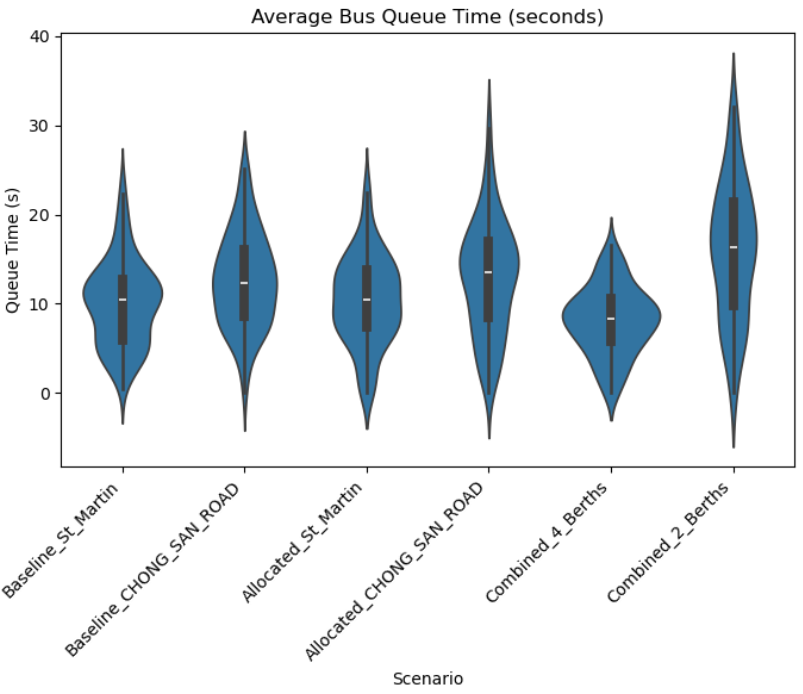


Figure Average Bus Queuing Time

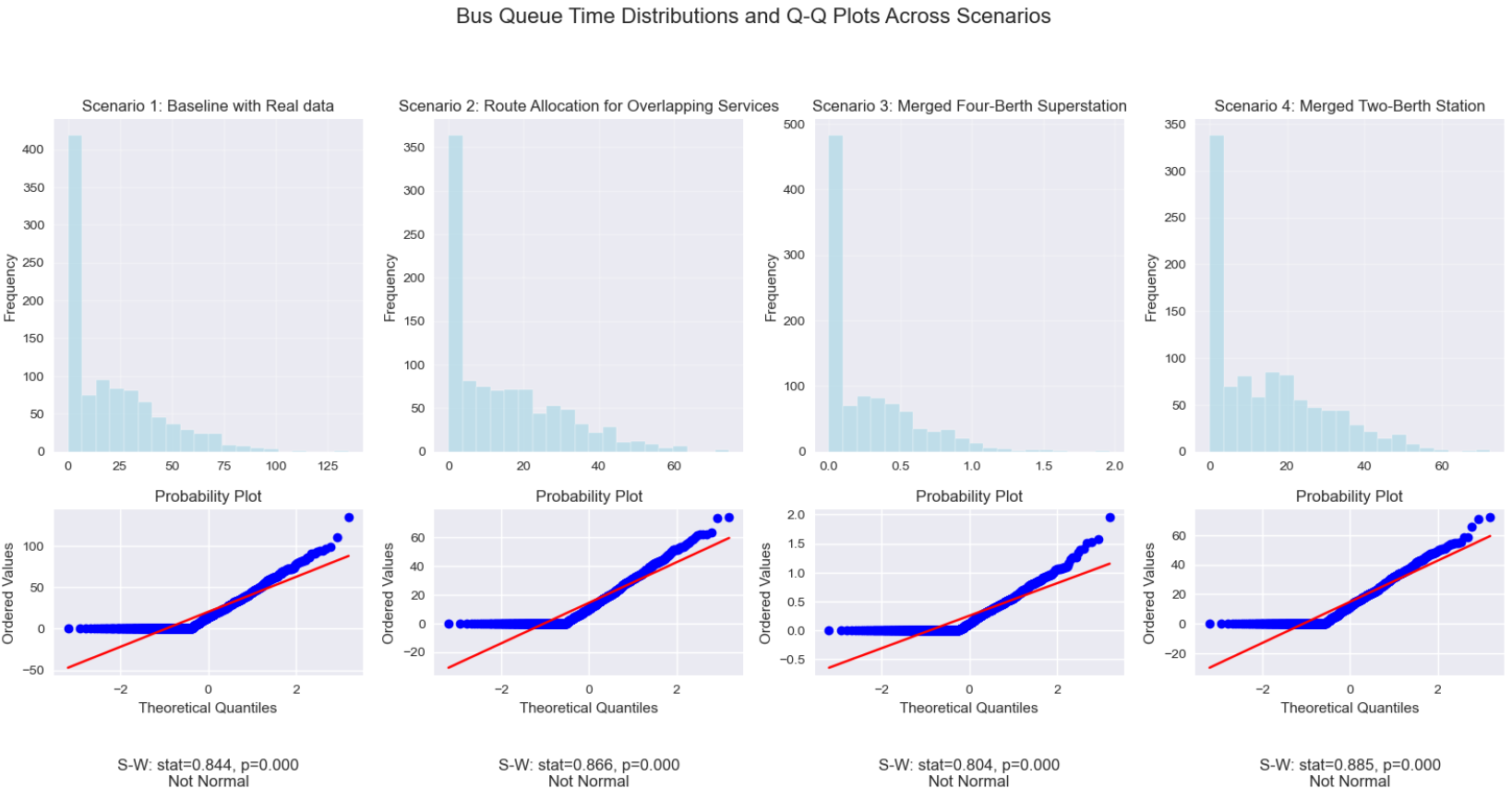


Figure Bus Queue Time Distribution Across Different Scenarios

In Figure 9, the total sample size is 1000, and the frequency refers to the proportion of the sample quantity within a certain period of time to the total sample size.

This graph has chosen different ranges for the horizontal and vertical axes to enable a clearer view of the details. Moreover, we can observe that the queuing time in each scenario shows a right-skewed trend, indicating that most of the queuing times are concentrated in a shorter time interval. And there are very few long queuing times. The queuing times in scenarios 1 and 2 are generally longer, while the queuing times in scenarios 3 and 4 are relatively shorter.

I also conducted the Shapiro-Wilk test to check the normality of the data. In the QQ plot, the data do not follow the 45-degree line, and the P-value is less than 0.05. Therefore, we reject the null hypothesis and there is no evidence to suggest that the data conform to a normal distribution. From this, we can draw the conclusion that the waiting time for buses does not follow a normal distribution.

图表, 图示

AI 生成的内容可能不正确。

Figure Violin plots for Bus Queue Time Distribution

Violin plots are a hybrid visualization that combines elements of box plots and kernel density plots. Violin plots are created to provide a more comprehensive view of data distribution compared to box plots or histograms alone. Additionally, the violin plots incorporate box plot elements (displaying median, quartiles, and whiskers), which enhance the visualization by highlighting key summary statistics alongside density information. These plots deliver a thorough depiction of data density, revealing not only aggregated statistics but also the shape and spread of the distribution. They prove valuable in spotting skewness, multiple peaks, or outliers that might go unnoticed in less detailed charts, providing a deeper insight into the dynamics of queue time variations across the bus network.

Figure 8 displays violin plots, effectively illustrating the distribution of queue times for various bus routes under different scenarios. In Scenario 1 (Baseline with Real Data), it is evident that the 272A bus experiences the highest queue times, likely due to its highest departure frequency. The 272A's average queue time distribution appears the most symmetrical compared to other routes, with the greatest variation in wait times. The distribution range of queuing times in scenarios 1 and 2 is wide, indicating an uneven distribution of resources. In Scene 3, almost all the lines do not require queuing. The waiting time for the buses is almost zero, indicating that the optimization result is the best. The improvement measures in Scenes 3 and 4 effectively reduced the waiting time and enhanced the operational efficiency of the public transportation system.

# Data Insights and conclusions

The line graph below, illustrating the percentage change in average bus queue time compared to Scenario 1, provides several notable observations. All these three methods have to some extent changed the queuing situation. Scenario 2 shows a moderate improvement of approximately 30%, suggesting that the adjustments made in this scenario effectively reduced wait times compared to the baseline.

Table Improvements of Queue Time

|  |  |
| --- | --- |
| Scenarios | Improvements of queue time |
| Route Allocation for Overlapping Services | 30% |
| Merged Four-Berth Superstation | 100% |
| Merged Two-Berth Station | 10% |

Scenario 3 demonstrates the highest improvement, peaking at around 100%, indicating a significant improvement in queue time efficiency, due to successful operational changes. It may have eliminated the queue or significantly reduced the waiting time. If there is sufficient budget, the best approach would be to merge into a single large station. This not only resolves the issue of bus queues but also enables the residents of both areas to avoid traveling a longer distance to catch their desired buses. However, if we take into account the issue of funds, and without spending extra money to improve the situation, I recommend using the other two methods.

图表, 折线图

AI 生成的内容可能不正确。

Figure Improvement Rate in Average Bus Queue Time (vs current time)

Based on the analysis of various situations, both scenario 2 and scenario 4 significantly reduced the average bus queue time. Scenario 2 provides a relatively stable and feasible solution to improve public transportation efficiency. Compared to the baseline scenario 1, scenario 2 has a moderate improvement of approximately 30%, indicating that it is a balanced strategy that can effectively simplify the operation process without causing significant disruptions. This scenario may involve route adjustments or resource reallocation to optimize the service process. Therefore, it is a reliable choice for enhancing passenger experience during peak hours. The steady progress achieved by this scenario indicates that it is a feasible strategy with relatively low implementation risks and can be achieved quickly. However, it requires professional personnel for debugging and has a relatively higher operational level.

In conclusion, prioritizing Scenario 2 offers a safer, more consistent improvement with fewer risks, making it an appealing option for immediate action. Scenario 3, while highly effective in reducing queue times, requires a detailed evaluation of its impact on waiting periods to ensure it aligns with passenger needs.

In a certain sense, merging bus stops can reduce the number of stops for buses, thereby improving the operational efficiency of the entire route. However, it may not be applicable in areas with a large population base. Therefore, using this method might require simultaneously increasing the frequency of bus services. However, an increase in frequency might lead to a further aggravation of congestion on public transportation routes. How to achieve a balance still requires more detailed research for verification. Therefore, it cannot be determined whether this method has any effect on alleviating traffic congestion in this area.

Finally, I suggest redistributing the routes or converting them into a large station with four platforms, which would facilitate the flow of traffic. Scenario 4 requires considering the resident population base and the vote to determine which station to cancel. It should be set up as a regular station with only two stops. Compared to upgrading the stations as proposed in Scenario 3, Scenario 2 may be implemented the fastest because it does not require additional urban planning and construction.

Finally, I suggest redistributing the routes or converting them into a large station with four platforms, which would facilitate the flow of traffic. Scenario 4 requires considering the resident population base and the vote to determine which station to cancel. It should be set as a regular station with only two stops. Compared to upgrading the stations as in Scenario 3 and cancelling one station as in Scenario 4, Scenario 2 can be implemented the fastest because no additional urban planning and construction is required.

## Research limitations and Improvements

The recent analysis of bus queue time distributions and operational efficiency across various scenarios provides valuable insights, however there are several areas where improvements can be made to enhance the accuracy and applicability of the model.

One significant enhancement would involve incorporating a more realistic representation of passenger arrival patterns. Currently, the model assumes a uniform or unspecified distribution of passenger arrivals. A more robust approach would be to model passenger arrivals as following a Poisson distribution, which is commonly used to represent random arrivals in queueing systems like bus stops. To implement this, we could collect actual passenger arrival data from Hong Kong bus stops during peak hours and perform a statistical fitting process to determine the most suitable Poisson parameter , representing the average arrival rate. This could be achieved by aggregating timestamp data from bus stop sensors or ticketing systems, calculating inter-arrival times, and using maximum likelihood estimation or a chi-square test to fit the data.

Another critical aspect to address is the variation in passenger boarding numbers across different routes. In reality, each bus route in Hong Kong serves distinct passenger volumes, influenced by factors such as route popularity, destination, and time of day. The current model treats all routes uniformly, which oversimplifies this diversity. To improve this, we could integrate route-specific passenger load data, potentially sourced from operator records or smart card transactions, to adjust the synthetic data generation process. This would allow the model to reflect realistic boarding demands, enhancing the fidelity of queue time predictions.

Additionally, the model should consider the fixed stopping positions of buses in Hong Kong, where buses halt at designated berths. This spatial constraint may lead to interference, where one bus’s occupancy of a berth delays others’ access. The current analysis does not account for this dynamic, and future iterations could simulate berth allocation and potential blockages, perhaps using a spatial queueing model, to better capture real-world interactions.

Furthermore, the instability in calculated passenger average queue times poses a challenge. These times are heavily dependent on passenger numbers, which fluctuate significantly during peak hours. The model’s reliance on repeated experiments to approximate real values is a reasonable strategy, but the large standard deviations observed in the data suggest high variability. To mitigate this, we could increase the number of simulations run (e.g., from 100 to 1000) to smooth out fluctuations, though this would require more computational resources. Alternatively, incorporating a confidence interval analysis could provide a clearer picture of the reliability of the mean queue times, acknowledging the inherent randomness

Lastly, validating the model against historical operational data from Hong Kong’s transport authority could help calibrate it further, reducing the gap between simulated and actual performance. These enhancements would collectively strengthen the model’s predictive power and practical utility.

# References

|  |  |
| --- | --- |
| [1] | N. Brown, "Bus System and Accessibility: A Qualitative Approach," *MSc Thesis, University of London,* 1996. |
| [2] | A. Frye, "Bus Travel - A Vital Link in the Chain of Accessible Transport," *Proceedings of Bus and Coach,* 1996. |
| [3] | N. a. B. N. Tyler, "On Mobility," *University of London Centre for Transport Studies Working Paper,* 1997. |
| [4] | N. Tyler, "On Accessibility," *University of London Centre for Transport Studies Working Paper,* 1996a. |
| [5] | A. Lobo, "Automatic Vehicle Location Technology: Application for Buses," *PhD Thesis, University of London,* 1997. |
| [6] | G. C. P. a. C. J. Gardner, "The Performance of Busway Transit in Developing Cities," *Transport and Road Research Laboratory,* 1991. |
| [7] | N. Tyler, "High Capacity Bus Systems: A Design Methodology Using Artificial Intelligence," *PhD Thesis, University of London,* 1992. |
| [8] | O. A. J. W. J. a. M. H. P. Z. Nnene, "Integrating Activity-Based Simulation with Multi-Objective Metaheuristic Optimization for Transit Network Design," *Transportation Research Part C: Emerging Technologies,* vol. 45 (July), p. 87–104, 2014. |
| [9] | H. J. W. e. a. Yin, "Agent-Based Simulation and Genetic Algorithm for Passenger Flow Guidance in Urban Rail Networks," *Journal of Advanced Transportation,* vol. 50, no. 3 (May–June), p. 345–362, 3 (May–June) 2016. |
| [10] | M. e. a. Zhang, "Bus Scheduling with Full-Stop and Skip-Stop Services to Minimize Passenger Travel Time on Overlapping Routes," *Transportation Research Record,* vol. 2672, no. 1 (January), p. 12–23, 2020. |