Train Dwell Time Models for Urban Rail Transit - Investigation on Impact of Station Design and Passenger Flow Dynamics

Buchunde, Suryakant

https://hdl.handle.net/1880/118540
Downloaded from PRISM Repository, University of Calgary
Train Dwell Time Models for Urban Rail Transit - Investigation on Impact of Station Design and Passenger Flow Dynamics

by

Suryakant Buchunde

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN CIVIL ENGINEERING

CALGARY, ALBERTA

APRIL, 2024

© Suryakant Buchunde 2024
Abstract

Dwell time is an important part of the total travel time in urban rail transit, which directly impacts the system’s reliability and line capacity. In this thesis the impact of station design, train load, and passenger flow in terms of boarding fraction on train dwell times is investigated through a system-wide automated data sources including Automatic Passenger Counter (APC) and Automatic Vehicle Locating (AVL) data in the Light Rail Transit System in Calgary. Moreover, a simulation model of train station in Calgary in PTV-Vissim is developed. Methods to improve simulation results are proposed by including passengers’ walking behaviour and their interaction within the transit system. Alternative dwelling strategies are then explored.

Regression analyses were conducted to achieve accurate dwell time estimation by calculating passenger load per car, boarding, and alighting passengers per door, and also identifying the critical door with the highest boarding and alighting. Observations for each critical door were divided based on the fraction of boarding passengers with respect to the sum of boarding and alighting. Six stations with distinct geometric designs were selected for comparison to assess their impact on dwell time. The results indicate that for dominant boarding or alighting, a longer time is needed per passenger to alight or board, respectively. The findings in this thesis also indicate that a station with a middle platform and two entrances positioned in the middle performed better in terms of dwell time in the case of alighting-dominant and mixed passenger flow. For stations experiencing boarding-dominant passenger flows, side platforms with multiple entrances at ends and middle outperformed. Narrower platforms experienced significantly longer dwell times than other selected stations under similar demand.

The desired speed factor is used in Vissim to adjust the walking speed of pedestrians in the simulation. Moreover, the social force parameters are tuned to improve the accuracy of the
simulation by adjusting the forces around pedestrians and their interactions with their surroundings. As such, the Vissim models were improved by accessing Vissim objects through their Component Object Model (COM) interface, which allows communication and interaction between software components across different programming languages and environments, in Python to enhance the simulation performance. The simulation results indicate that a dedicated boarding and alighting door with uniform distribution of boarding passengers on the platform outperformed in handling mixed passenger flow. On the other hand, uniform distribution of boarding passengers on the platform outperformed in the cases of boarding-dominant and alighting-dominant passenger flows.

While the conclusion regarding the station design may not be generic, the proposed model provides a consistent and adaptable approach to study the impact of station design and passenger flows on urban rail dwell times facilitating more informed decision-making for station design or modification and enhancement of the overall rail system performance.
Acknowledgement

I am immensely grateful to my supervisor, Dr. Saeid Saidi, for his immense support, guidance, and mentorship. His faith in my abilities and commitment to cultivating a supportive academic atmosphere have been indispensable factors in my path to achievement. His compassionate mentorship style has left a lasting impact on me, and I consider myself fortunate to have had such an exceptional supervisor. I am confident that this relationship will continue to flourish throughout my academic career in the years to come.

I extend my heartfelt gratitude to Shervin Ataeian whose diverse perspectives and dedicated time, despite her many commitments, have greatly enriched my research. The constructive feedback and suggestions provided by her have significantly enhanced the quality of this thesis. I am thankful for always being ready to answer my queries.

This research could not have been possible without the invaluable data provided by Calgary Transit, City of Calgary. Their cooperation and support were imperative for the successful execution of this study. I extend my sincere gratitude to Dr. Asim Mohammad, Nicholas Band, Andy Chu, and Janice Lau for their invaluable support and insightful comments, which significantly contributed to this endeavor.

The attainment of my master’s degree would have been impossible without the financial assistance I received from various sources. I express my sincere gratitude to the Ministry of Social Justice and Empowerment, Government of India, for sponsoring my studies through the National Oversea Scholarship Scheme. I am particularly thankful to Sanjay Singh, Rajinder Kumar, and Sanjay Malhotra for their efficient management of my scholarship and their efforts to extend it beyond the approved period. Furthermore, I am deeply appreciative of the support provided by Alberta Innovates and the Natural Sciences and Engineering Research Council of Canada.
This work would not have been possible without the generous support from these organizations.

I express my gratitude to James Murphy, the librarian, for his assistance in developing effective strategies for literature search in databases pertaining to my research topic. Additionally, I extend my thanks to Kertz Tobias for providing technical support in developing the microsimulation model for this study. I am grateful to Dr. Merkebe Dimissie for giving me the opportunity to work on Curb Space Management, which expanded my technical expertise beyond my main research focus. I also wish to acknowledge the mentorship and support of Dr. Pouya Zangeneh and Jacob Lamb throughout this journey. I am thankful to my labmates for engaging in numerous interesting discussions and providing peer support.

I owe a profound debt of gratitude to my mentor, Anoop Kumar, whose insights into the significance of pursuing higher education have been truly enlightening. I am also immensely thankful to my friends, who have provided financial, emotional, and mental support, as well as to the mentors at Bluebird School for Foreign Education, Nalanda Academy, Wardha. Finally, and most importantly, my deepest appreciation goes to my family, particularly my brother, Chandrakant Buchunde. Your constant presence, encouragement, and continuous support have been my pillars of strength throughout this academic pursuit. To my father, Keshao Buchunde, mother, Kalawati Buchunde, and sister, Swati Buchunde, your endless love and encouragement have been my greatest source of inspiration. Without your unconditional support, I would not have been able to embark on this academic journey without your unconditional support.
Dedication

DR. BABASAHEB AMBEDKAR

Renowned as a visionary leader, social reformer, and the chief architect of the Indian Constitution. His commitment to social justice, equality, and the empowerment of marginalized communities, alongside his tireless efforts to dismantle the entrenched caste system, has left a memorable mark on generations. From my grandparents being denied access to education to me becoming the first generation to attend graduate school. This journey has been unimaginable for me. It was Dr. Ambedkar’s efforts to ensure social justice that changed the lives of me as well as many individuals who aspired to live a life of dignity.

&

LATE ABHIYAN HUMANE

He was a teacher, artist, storyteller, mentor, scientist, and changemaker. After studying and working in United States in Arts-Science-Technology for over decade, he left a lucrative job to join Nalanda Academy, Wardha. There, he established a Science Lab to train school-going students in Art, Design, Robotics, and Artificial Intelligence. He traveled across India, conducting workshops at both private and government-owned schools in rural areas. He believed in promoting scientific temper and rational thinking to empower the marginalized.

I dedicate this thesis for their lifelong efforts to empower the marginalized and establish social equalities.

“JAI BHIM”
Table of Contents

Abstract .............................................................................................................................................. ii

Acknowledgement ................................................................................................................................. iv

Dedication .............................................................................................................................................. vi

List of Tables ........................................................................................................................................ x

List of Figures ....................................................................................................................................... xi

CHAPTER 1. INTRODUCTION ........................................................................................................ 1

1.1 Background ................................................................................................................................ 1

1.2 Objective of Study ......................................................................................................................... 4

1.3 Proposed Methodology and Research Contributions ................................................................. 5

1.4 Research Limitations .................................................................................................................... 8

1.5 Organization of Thesis .................................................................................................................. 9

CHAPTER 2. LITERATURE REVIEW ............................................................................................... 12

2.1 Significance of Dwell Time Analysis and the Influential Factors ............................................. 12

2.1.1 Significance of Dwell Time Analysis ....................................................................................... 13

2.1.2 Influential Factors on Dwell Time ......................................................................................... 15

2.2 Dwell Time Estimation Models .................................................................................................... 17

2.2.1 Linear Regression Models ...................................................................................................... 17

2.2.2 Nonlinear Regression Models ............................................................................................... 20

2.2.3 Simulation Models ................................................................................................................. 21
2.2.4 Machine Learning Models ................................................................. 23

2.3 Data Collection Methods for Dwell Time-Analyses .............................................. 24

2.4 Passengers Walking Behaviour and Modeling ...................................................... 26

2.4.1 Pedestrian Walking Behaviour ................................................................. 26

2.4.2 Pedestrian Flow Modeling and Calibration .................................................. 28

CHAPTER 3. METHODOLOGY ............................................................................. 31

3.1 Regression Analysis ......................................................................................... 31

3.1.1 Data Collection .......................................................................................... 31

3.1.2 Data Processing ......................................................................................... 32

3.1.3 The Dwell Time Estimation Model .............................................................. 34

3.2 Microsimulation Model .................................................................................... 37

3.2.1 Dwell time model and relevant parameters in Vissim .................................. 37

3.2.2 Calibrating the simulation model .................................................................. 38

3.2.3 Passenger Flow Modeling ............................................................................ 39

CHAPTER 4. DWELL TIME MODELING ............................................................... 46

4.1 Station Selection ............................................................................................... 47

4.2 Descriptive Analysis ....................................................................................... 50

4.3 Results ............................................................................................................ 53

4.4 Identifying Outliers in the Data ....................................................................... 55

4.4.1 Test for Nonlinearity ................................................................................... 56
4.4.2 Residuals’ Distribution ................................................................. 58
4.4.3 Date of the Month ........................................................................... 60
4.4.4 Day of the Week .............................................................................. 61
4.4.5 Time of the Day .............................................................................. 62
4.4.6 Passenger Load ............................................................................... 64

CHAPTER 5. IMPACT OF STATION DESIGN AND PASSENGER FLOW .......... 68

5.1 Alighting-Dominant Passenger Flow .................................................. 68
5.2 Mixed Passenger Flow ........................................................................ 70
5.3 Boarding-Dominant Passenger Flow ................................................... 71
5.4 Discussion ........................................................................................ 75

CHAPTER 6. MICROSIMULATION MODEL ............................................... 78

6.1 Simulation Model for Erlton Station .................................................... 78
6.2 Alternatively Dwelling Strategies ......................................................... 81

CHAPTER 7. CONCLUSIONS AND RECOMMENDATIONS ...................... 84

7.1 Overview ........................................................................................ 84
7.2 Summary of the Work and Results ...................................................... 85
7.3 Contributions .................................................................................... 87
7.4 Recommendations .............................................................................. 88
7.5 Future Research ................................................................................ 89

REFERENCES ....................................................................................... 91
List of Tables

Table 3.1: Description of Groups Based on the Fraction of Boarding Passengers ..................... 36
Table 3.2: Dwell Time Components in Vissim and Their Definition and Units. ......................... 43
Table 4.1 Description of Platform Design. .................................................................................. 49
Table 4.2: Results for Without Considering Passenger Flow. ...................................................... 53
Table 4.3: MSE for Linear, and Polynomial Data Fits for Inbound and Outbound Directions. ... 57
Table 5.1 Results for Alighting-Dominant Passenger Flow. ...................................................... 69
Table 5.2 Results for Mixed Passenger Flow. .............................................................................. 70
Table 5.3 Results for Boarding-Dominant Passenger Flow. ....................................................... 71
Table 5.4 Results for Alighting-Dominant Passenger Flow. ....................................................... 73
Table 5.5 Results for Mixed Passenger Flow. .............................................................................. 74
Table 5.6 Results for Boarding-Dominant Passenger Flow. ....................................................... 74
List of Figures

Figure 3.1: Data processing steps. .......................................................... 33
Figure 3.2: Dwell time components in Vissim. .................................................. 38
Figure 4.1 Calgary’s light rail transit map (source: Calgary Transit). ....................... 47
Figure 4.2 Station layout, (a) Erlton, (b) Chinook, (c) University (d) Sunnyside (e) 39th Avenue and (f) Lions Park, red and yellow boxes to show the entrances and platform locations, respectively (source: Google Map). .......................................................... 48
Figure 4.3: Passenger demand distribution for (a) northbound, and (b) southbound directions. . 51
Figure 4.4: Box plots for (a) alighting, (b) boarding, (c) passenger load, and (d) dwell time..... 52
Figure 4.5: MSE comparison for linear, and polynomial data fits (a) inbound and (b) outbound directions for Erlton station. .................................................................................. 56
Figure 4.6: Residuals plots for Erlton, Chinook, University, Sunnyside, 39th Avenue, and Lions Park stations for (a) inbound and (b) outbound directions. ....................................................... 59
Figure 4.7: Dwell time comparison with passenger demand for the day of July 2022 and 2023 for inbound direction of Erlton station. ................................................................. 61
Figure 4.8: Dwell time comparison with passenger demand for day of the week for inbound direction of Erlton station. ............................................................................. 62
Figure 4.9: Dwell time comparison with passenger demand for Time of the day for inbound direction of Erlton station. .................................................................................. 63
Figure 4.10: Dwell time comparison with passenger demand for passenger load inbound direction of Erlton station.................................................................................................. 65
Figure 5.1: Dwell time comparison for boarding, and alighting passengers for (a) inbound and (b) outbound direction. ..................................................................................... 76
Figure 6.1: Platform layout of the simulation model in Vissim. ............................. 79

Figure 6.2: Results for simulation model calibration.................................................. 80

Figure 6.3: Comparison of alternative dwelling strategies for 50 passenger load................. 82
CHAPTER 1. INTRODUCTION

1.1 Background

The total travel time for passengers in any public transit system consists of different components, such as waiting time, riding time, and dwell time. Dwell time is an important component of rail transit operations that directly affects the rail line capacity and reliability of the service. Dwell time in a railway network is defined as the time required for passengers to board and alight at a station and the time needed to close and open doors. Dwell time is a relatively small amount of time compared to the ride time, but for a high-frequency transit system and near-capacity operations, accumulation of these insignificant periods of dwell time can significantly impact the total travel time and the system capacity.

There should be some minimum possible dwell time, ensuring all the passengers board or alight, given there is enough space on the train. On the one hand, a longer dwell time benefits passengers to board/alight comfortably without any rush; on the other, onboard passengers would experience a longer riding time, resulting in a longer total travel time. Although having a longer scheduled dwell time constrains the rail line capacity, it can improve reliability as a longer dwell time cuts back on possible delays. Therefore, while designing a railway schedule, specifically the dwell time component, one should consider all these factors to make a beneficial timetable for both onboard as well as boarding and alighting passengers without compromising the line capacity. In order to create a timetable that optimally serves the needs of passengers, who are either boarding, alighting, or onboard, it is important for transit planners to possess a comprehensive understanding of several critical factors that affect average boarding and alighting time per passenger. Moreover, it is equally crucial for transit planners to recognize that boarding and
alighting time per passenger is not the same in all situations. They can exhibit considerable variation under different circumstances, such as the presence of obstacles on a platform, location of entrances to the platform, width of a platform, and scenarios such as alighting dominant, boarding dominant, or mixed passenger flow.

The minimum required dwell time depends upon the passenger demand for boarding and alighting. However, it is important to acknowledge that several other variables play a significant role in determining the average time for passengers to board and alight. Few studies have reported the factors that influence the dwell time such as the distribution of passengers on the platform, width of a platform, width of the train car door, vertical height between a platform surface and vehicle floor, etc. However, the current literature lacks the recognition of the impact of the geometrical design of a platform and passenger flow on dwell time.

Many studies on dwell time rely on manual data collection, video camera recordings, or Automatic Fare Collection (AFC) data. These methods have limitations when it comes to collecting large datasets, accurately counting onboard passengers, or counting boarding and alighting passengers at each door. However, Automatic Passenger Counter (APC) systems have the ability to count the number of alighting and boarding passengers at each door. In this study, the critical door refers to the door with the highest sum of boarding and alighting passengers and is also associated with the car, which has the highest number of onboard passengers. The dwell time is influenced by the critical door, which has the longest passenger service time. We considered that the door with the highest passenger demand has the longest passenger service time. Since the distribution of boarding and alighting passengers on the platform is not uniform (Szplett & Wirasinghe, 1984; Wirasinghe & Szplet, 1984), identifying the critical door becomes crucial for accurately estimating the dwell time.
Advanced computational methods and simulation models have significantly facilitated the development of realistic visual models of passenger walking behaviour to study different strategies and test their efficiency. These powerful tools can be applied to develop robust solutions and reduce dwell time or improve the dwelling process. Simulation models can be classified into microscopic models, in which individual passengers and their interaction are modelled (Elmitiny et al., 2007; Han & Yuan, 2005; Kattan et al., 2012; Shiwakoti et al., 2011; Zhou et al., 2020), macroscopic models in which all pedestrians are considered as a single flow (Chan, 2010; Klunder et al., 2009) and mesoscopic models, which are a combination of microscopic and macroscopic models (Abdelgawad & Abdulhai, 2010; Saidi et al., 2019; Wang Li & Hong Yuliang, 2010). Experimental studies are relatively time-consuming and expensive to conduct. However, the simulation models provide an opportunity to test the efficiency of each or a combination of the proposed strategies. Many strategies can be used to improve the dwelling process, including but not limited to uniform distribution of boarding passengers over all doors to minimize the boarding time at exclusive doors, and dedicated doors for boarding and alighting passengers to reduce the interactions between passengers. The critical step in simulating transit operations is to develop a model which can represent the studied case well and then calibrate the model to test its efficiency. Model calibration can be done manually by tuning the parameters to match the observed data with the obtained results or using automated computational methods in which the objective function is to minimize an error term with simulation parameters as dependent variables. On the other hand, instead of adjusting relevant parameters or performing automated calibration methods, simulation results can be enhanced to capture the real trends by considering the walking behaviour and interactions of passengers within the transit system. Although the microsimulation approach allows users to develop effective strategies to improve the dwelling process, it is essential to calibrate a
dwell time model to test the efficiency of the developed model before starting with evaluating different strategies. In the literature, it is known that the process of calibrating simulation models is unique to each case study and requires a significant amount of effort to select appropriate parameters. While automated calibration processes offer a generalized approach, they often result in overfitting parameters to the observed trend.

Although many studies have focused on dwell time for urban rail, they failed to utilize automatic data collection methods and identify the impact of station geometric design and passenger flow on dwell time. This fundamental literature gap is the motivation for this thesis, which presents a step-by-step process to utilize data collected by AVL and APC systems and later use mathematical model. Moreover, this research aims to create a generalized framework that can improve the dwell time simulation performance. The proposed framework will accomplish this by introducing an equation that allows for the selection of the desired speed factor in Vissim. The goal is to provide a more straightforward and effective way to calibrate simulation models while avoiding overfitting followed by testing the alternative dwell strategies to minimize the dwell time.

1.2 Objective of Study

The primary goal of this thesis is to develop a systematic methodology for utilizing AVL and APC data to conduct an in-depth study on train dwell time and investigate the impact of station design and passenger flow on dwell time. Additionally, the research aims to develop a comprehensive framework to enhance the performance of dwell time simulations. This framework will be achieved by introducing an equation in Vissim that enables the selection of an optimal speed factor. The methodology for utilizing APC and AVL data to analyze dwell time involves initial data processing and subsequent calculation of passenger loads by considering both boarding
and alighting passengers for each trip. The subsequent objectives of this study revolve around employing mathematical techniques. In this phase, the data for each station is categorized based on direction, and further, categorized into three passenger flow types: alighting-dominant, boarding-dominant, and mixed passenger flow. Subsequently, regression models are developed for each passenger flow category and direction. The simulation approach complements the mathematical methods, addressing comparisons that may be challenging through mathematical means. This involves the creation of a dwell time model in the Vissim simulation software.

Using the developed methodology, this thesis addresses the aforementioned literature gap, by using AVL and APC data to identify the critical door and accurately estimate passenger load for dwell time estimation. Additionally, the study considers the friction between boarding and alighting passengers and divides the data into multiple groups based on the fraction of boarding passengers, allowing for a comparison of changes in average boarding and alighting time per passenger. By comparing the average time that takes for passengers to board and alight at various station designs, we can distinguish the effects of station design on dwell time and make recommendations for improving the platform geometry. Additionally, the methodology to enhance the performance of the simulation model is developed, considering passenger flow. Subsequently, alternative dwelling processes, such as uniform distribution of passengers on the platform and dedicated boarding and alighting doors, are tested.

1.3 Proposed Methodology and Research Contributions

Recent advancements in technology have brought a new era of data collection, wherein the acquisition of precise and extensive datasets, facilitated by sensors, has become remarkably effortless. The focus of this thesis revolves around the utilization of data sourced from the Calgary
Transit’s Light Rail Transit (LRT) system, a dataset covers the entire months of July 2022 and July 2023, as well as 15 days each from January, April, and September 2023. The methodology started with data preprocessing, employing the programming language, Python. This was aimed at extracting meaningful insights from the data, primarily involving the calculation of critical metrics such as dwell time, the number of boarding passengers, alighting passengers, and onboard passengers for each station situated along both Calgary’s red and blue LRT lines. The next step was to select stations from both lines with unique platform designs. These designs encompassed a variety of platform layouts, including side and middle platforms, as well as a combination of entrance configurations. This strategic selection was undertaken with a specific objective in mind to investigate and draw comparisons regarding the impact of station design on dwell time. Subsequently, regression models were developed for each station, while considering various passenger flows and directions. This exercise allowed to gain a more profound understanding of how station-specific factors, including design and passenger flow, influence the critical operational aspects of the transit system.

Advancements in microsimulation software packages allow us to develop a detailed model of rail transit stations and test various strategies and their efficiency to improve the dwelling process in urban rail operations. In this thesis, we developed dwell time models in PTV-Vissim and then proposed a method to improve simulation results by including passengers’ walking behaviour and their interaction within the transit system. The desired speed factor is used in Vissim to adjust the walking speed of pedestrians in the simulation. Moreover, the social force parameters are tuned to improve the accuracy of the simulation by adjusting the forces around pedestrians and their interactions with their surroundings. These adjustments help to better reflect real-world pedestrian behaviour in simulations. The Vissim’s user-friendly graphical interface (GUI) was used
to develop the dwell time models and afterwards, the Vissim models were improved by accessing Vissim objects through the COM interface in Python to enhance the simulation performance. Later, alternative dwelling strategies, including uniform distribution of passengers on the platform and dedicated boarding and alighting doors, are tested.

This thesis makes the following contributions to the dwell time of urban rail operation:

1. **Utilization of APC and AVL data**: The thesis presents a pioneering approach by demonstrating the effective utilization of APC and AVL data for in-depth investigations related to dwell time. This innovative utilization of such data sources enhances our understanding of the factors influencing dwell time within urban rail systems.

2. **Passenger flow dynamics**: An essential aspect of this research involves an insightful comparison of the fraction of boarding passengers, while also considering the precise passenger load. This nuanced analysis enriches our comprehension of how passenger dynamics impact dwell time, going beyond conventional metrics.

3. **Station-Specific Geometric Characteristics**: The thesis offers a detailed examination of how station-specific geometric characteristics influence dwell time. This comparative assessment sheds light on the varying effects of station layouts and designs, emphasizing the need for tailored solutions in optimizing urban rail operations.

4. **Enhancing Microsimulation Performance**: The study presents a methodological framework aimed at enhancing simulation model performance by incorporating realistic passenger walking behavior, especially during crowded platform
situations. This approach ensures a more accurate representation of real-world scenarios.

5. **Alternative Dwelling Process:** Additionally, the research explores alternative dwelling strategies such as uniform passenger distribution on platforms and dedicated boarding and alighting doors. This enables policymakers to consider various strategies, either individually or in combination, to minimize dwell time effectively.

### 1.4 Research Limitations

This study relies on data obtained through APC and AVL systems. It is important to note that within the Calgary Transit system, there is a mix of older and newer trains currently in operation. However, not all of these trains are equipped with APC and AVL systems, which are integral to this research. The vehicles equipped with these sensors belong to the same model and have identical interior designs. This model features a longitudinal seating configuration, which provides more space for standing passengers. This train model has a capacity of 200 passengers per car. It is a high-floor vehicle and platforms in Calgary are also raised, which allows for level boarding. It is also noteworthy that Calgary Transit does not have a specific strategy or plan in place to phase out trains that lack these data-capturing systems. The absence of a comprehensive strategy regarding the deployment of trains with and without APC and AVL systems imposes certain limitations on this study. Data is available only for the trains equipped with APC and AVL systems, which limits this study from calculating headways and identifying the impact of longer headways on dwell time. Also, collected data does not include the information about passengers’ characteristics like gender, age, and accessibility requirement.
Moreover, it is crucial to understand the methodology behind the data collection process. The data is gathered during the moments when passengers either board or alight the train. This is done by recording the number of doors, through which passengers pass. However, it is important to acknowledge that this data collection method does not provide insights into the choices made by passengers regarding their entry and exit locations for the platform. Additionally, it does not reveal passengers' preferences when it comes to selecting waiting areas on the platform. These intricacies remain undisclosed within the scope of this study and limit the analysis of passenger distribution on the platform and dwell time.

1.5 Organization of Thesis

This thesis consists of chapters as follows:

Chapter two focuses on the literature review which delves into an extensive examination of existing literature in the field. It explores various methodologies used in analyzing dwell time within urban rail systems, categorizing them into linear regression models, nonlinear regression models, simulation models, and machine learning models. The chapter also provides an in-depth analysis of data collection methods, from manual techniques to sophisticated systems like AFC and APC. Furthermore, it explores the role of simulation models, specifically in Vissim, for transit analysis. Lastly, the chapter delves into the behavior of passengers, including their walking behavior and pedestrian flow modeling and calibration in Vissim.

Chapter three presents the methodology employed in this research, which is divided into two main sections: regression analysis and microsimulation model. The regression analysis section begins with an explanation of the data collection method utilized for this study. It then outlines the process of data processing to identify independent variables, followed by the procedure for
constructing a regression model. The microsimulation model section begins by elucidating the dwell time model and pertinent parameters within Vissim. Subsequently, it progresses to the model calibration process, where the selection of relevant parameters and their corresponding values are explained. Following this, the methodology for passenger flow modeling in Vissim is established.

Chapter four focuses on presenting the results obtained from the regression analysis. It starts by selecting multiple stations with unique platform designs to ensure a diverse representation in the study. Following this, a descriptive analysis of the collected data is conducted to provide an overview of its characteristics. The chapter then delves into discussing the results of the dwell time model, examining any outliers observed in the data to gain insights into their impact on the findings.

Chapter five examines the impact of station design and passenger flow on dwell time. In this chapter, six stations with unique platform designs, such as stations with middle platforms or side platforms and single or multiple entrances to the platform, are considered. For each station, data is divided into two directions and three passenger flows, resulting in six regression models for each station. These models are carried out to make a comparison of dwell time across stations' geometric designs and passenger flows.

Chapter six is dedicated to showcasing the results derived from the simulation model, specifically when implemented at the Erlton Station. This station, situated on Calgary's red line, is highlighted due to its high traffic volume and unique layout, featuring a middle platform and entrances at both ends. The chapter extensively examines the calibration process of the microsimulation model, which involves considerations such as passenger flow modeling. Additionally, various strategies aimed at reducing dwell time are tested and evaluated within the context of the station's operations.
Chapter seven summarizes the key findings and conclusions drawn from the study. Importantly, this chapter explains the significance of the research within the broader context of urban rail operations. In addition to the conclusions, the chapter offers practical recommendations based on research outcomes, serving as valuable guidance for transit system planners and decision-makers.
CHAPTER 2. LITERATURE REVIEW

The literature review for this thesis involves a comprehensive exploration of existing research within the field of urban rail. This chapter is divided into four sections: dwell time importance and influential factors, methodologies for dwell time modeling, and data collection methods, and finally, passengers’ behavior. The first section focuses on the importance of conducting studies related to dwell time and the relevant factors considered in the literature. The second section investigates different methodologies used for analyzing dwell time in urban rail systems, organizing them into different categories such as linear regression models, nonlinear regression models, simulation models, and machine learning models. Additionally, the chapter extensively discusses various data collection techniques, ranging from manual methods to more advanced systems such as AFC and APC systems. Lastly, the chapter explores passenger behavior, including their walking patterns and the modeling and calibration of pedestrian flow using Vissim, which is a simulation software package for multi-model traffic flow simulation.

2.1 Significance of Dwell Time Analysis and the Influential Factors

In terms of rail transportation, dwell time is the duration from wheel stop to wheel start when a train arrives at and departs from a station platform (Karekla & Tyler, 2012). There are different definitions for train dwell time in the literature; Koffman et al., (1984) considered the total train stopping time as dwell time which included delays due to traffic signal timings. However, in a study by Wirasinghe & Szplett, (1984) dwell time was defined as the train standing time, which is divided into two components: a fixed time and a passenger service time. Moreover, most authors used the same definition for dwell time as the time spent by a train at a station for
passengers’ boarding and alighting (Buchmueller et al., 2008; Chu et al., 2015; Li et al., 2018; Lin & Wilson, 1992; Thoreau et al., 2016). Dwell time can consist of five processes: door-unblocking, door opening, passenger alighting/boarding, door closing, and train departing (Buchmueller et al., 2008). Door-unblocking refers to the duration between the train's arrival and the initiation of the door opening process. Although dwell time is a relatively small duration of time compared to the total travel time, variations in dwell time can influence the overall journey time and passenger experience significantly (Chu et al., 2015; Li et al., 2018; Thoreau et al., 2016; Wirasinghe & Szplett, 1984).

2.1.1 Significance of Dwell Time Analysis

The importance of analyzing dwell times has increased as transportation authorities seek to improve the continuity and efficiency of their railway service (Buchmueller et al., 2008). The purpose of a mass transit system is to move as many people as possible in a fast, safe, and inexpensive manner. Extended or inconsistent dwell times lead to longer trips, delays, and irregular headways. If dwell time is left unchecked, there is a reduction in total capacity and customer satisfaction (Barron, 2015). Becker & Schreckenberg, (2018) studied the effect of irregular dwell times on rail line capacity and found that increased dwell times constrain system capacity in high-frequency transit systems. While these problems are more common where passenger volumes and train frequencies are higher, passenger time can also be wasted during less crowded times if dwell times are not well managed (Hyun et al., 2016).

The consequences of inefficient dwell time have a significant negative effect on the service quality of a rail system. An increase in dwell time not only extends the total journey time, but also leads to crowding at subsequent stations (Saidi et al., 2023; Yoo et al., 2022). Considering a
uniform boarding passenger arrival rate over time, an increase in the actual dwell time not only increases the total journey time but also the crowdedness at subsequent stations, which can result in personal and social problems such as stress, anxiety, security, and safety issues (Yoo et al., 2022).

Although it is observed that small delays in railway operations are not well reported, there are studies available to understand the causes of dwell time variations. Harris et al., (2013) conducted a study on the Norway transit network to understand the railway performance and reasons for delays at stations. The authors concluded that railway management paid attention to significant incidents of train delays, while small delays were inadequately recorded and not well understood. Nevertheless, short delays could become important in a busy railway network. It was mentioned that reasons to ignore small delays were the inability to measure the train stop time by an automatic system such as signaling equipment, lack of understanding of passenger behaviors, and lack of ability to measure delays of less than 1 minute by the railway management (Harris et al., 2013).

In summary, the analysis of dwell times within railway systems is crucial for transportation authorities attempting to enhance the efficiency and reliability of the systems. As highlighted by various studies, extended or irregular dwell times can lead to numerous challenges, including longer trips, delays, and reduced capacity, ultimately impacting customer satisfaction and service quality. The implications of inefficient dwell time management extend beyond mere inconvenience, affecting passenger experiences and even safety. While small delays may often go unnoticed, they can accumulate to significant disruptions, particularly in busy railway networks. Therefore, there is a pressing need for further research and attention to understanding and addressing the dwell time to ensure seamless operations of rail systems and the satisfaction of passengers' needs.
2.1.2 Influential Factors on Dwell Time

The dwell time process represents the period, during which interactions between passengers and the physical railway environment occur. Consequently, the physical environment directly impacts dwell times as passengers engage with it. Upon a train's arrival at the platform, the initial factors contributing to increased dwell time include door unblocking, which depends on the train operator's action to initiate door opening for the subsequent processes (Martínez et al., 2007). Harris et al., (2022) illustrated the importance of the number of doors in minimizing total dwell time. For instance, with 50 passengers moving through two doors at a rate of one passenger per second, it would take 25 seconds. However, with five doors, this duration reduces to 10 seconds. Furthermore, besides the quantity of doors, the configuration of the doorway has been recognized as a significant factor impacting passenger flow rates, thus influencing the duration of dwell time. Harris et al., (2014) found that door width, spacing between doors along the train, and any steps between the platform and the train all influence the time taken for passengers to board and alight. Yang et al., (2023) expanded the parameters considered for dwell time analysis to include factors such as passenger gender, presence of luggage, waiting area width, and door width. They discovered that passenger gender emerged as the most influential factor affecting dwell time among those examined in their model.

Szplett & Wirasinghe, (1984) investigated the spatial characteristics of passengers on an urban rail platform specifically examining how station design influenced characteristics related to passenger waiting, boarding, and alighting. Their research revealed that a station featuring a middle platform with a single entrance at one end displayed a waiting passenger distribution following a log-normal pattern. However, stations with a side platform and multiple entrances exhibited a
waiting passenger distribution conforming to the normal distribution. Also, they found that a few passengers, who were denied boarding could have successfully boarded if they had chosen to move toward or wait near a door experiencing lower demand. The significant finding of their research is that the distribution of passengers on the platform is not uniform. Utilizing uniform demand for all doors or average demand at each door in dwell time modeling leads to an underestimation of dwell time. This highlights the significance of identifying the critical door for accurate dwell time estimation. Another study conducted in Calgary aimed to estimate the dwell time by accounting for both boarding and alighting passengers (Wirasinghe & Szplett, 1984). However, the study overlooked the consideration of passenger load of train cars with the critical door. Passenger load refers to the number of passengers on a transit vehicle at a particular station or location (Parkinson & Fisher, 1996). The interaction between boarding and alighting passengers is characterized as friction between boarding and alighting passengers (Wirasinghe & Szplett, 1984). To capture this effect, Wirasinghe & Szplett, (1984) calculated the ratio of boarding passengers to the total of boarding and alighting passengers at the critical door. This ratio is referred to as the fraction of boarding passengers. Passenger flow refers to the fraction of boarding passengers. For instance, a passenger group with a higher fraction of boarding passengers, is referred to as a boarding-dominant passenger flow. Average boarding and alighting time per passenger changes with the fraction of boarding passengers (Parkinson & Fisher, 1996; Wirasinghe & Szplett, 1984). This underlines the importance of considering passenger flow for an accurate estimation of dwell time.

In summary, the dwell time process in railway operations is influenced by various factors, including the number and configuration of doors, spatial characteristics of station design, and the interaction between boarding and alighting passengers. Studies have shown that increasing the number of doors reduces dwell time, with considerations such as door width and spacing also
impacting passenger flow rates. Additionally, spatial characteristics of stations contribute to non-uniform distributions of waiting passengers, emphasizing the need for accurate modeling. Accounting for passenger load and the ratio of boarding to the total of boarding and alighting passengers at critical doors is crucial for precise dwell time estimation. However, existing literature on dwell time estimation lacks an investigation of the impact of the fraction of boarding passengers considering the exact passenger load for the car with critical door and station-specific geometric characteristics on dwell time.

2.2 Dwell Time Estimation Models

Numerous dwell time estimation models can be found in the literature. This section is divided into four subsections: linear regression models, nonlinear regression models, simulation models, and machine learning models for a better understanding of dwell time models in the literature.

2.2.1 Linear Regression Models

The simplest form of a linear regression model can be obtained by considering the total time required for the boarding and alighting process. A study conducted for Hong Kong LRT system estimated dwell time as the total time needed for boarding passengers to board and alighting passengers to alight (Lam et al., 1998, 1999). In this approach, passenger demand for each door is not considered. This is the simplest form of a linear regression model where dwell time is the dependent variable with only two independent variables, total number of alighting and boarding passengers. The linear regression model can be expressed as shown in Equation 2-1,
\[ DT_n = DF + \beta_1 \mu_n + \beta_2 \lambda_n \]  

where,

\( DT_n = \) total dwell time of train \( n \),

\( DF = \) fixed time,

\( \mu_n = \) total number of alighting passengers of train \( n \),

\( \lambda_n = \) total number of boarding passengers of train \( n \), and

\( \beta_1, \beta_2 = \) estimated coefficients.

Similar forms of linear regression models can be found in the literature, where authors replace the total boarding and alighting demand with the passenger demand for each door, considering the uniform distribution of passengers on the platform (Gysin, 2018; Koffman et al., 1984). However, this argument does not hold true in the real world. Passenger distribution on the platform depends on the location of entrances (Szplett & Wirasinghe, 1984). To tackle this problem, a study by Wirasinghe & Szplett, (1984) considered the passenger demand at the critical door, defined as the door with the maximum number of boarding and alighting passengers. By considering the fraction of boarding demand to the total boarding and alighting demand at the critical door (shown in Equation 2-3), this study divided passenger flow into five distinct groups and modeled the dwell time as presented in Equation 2-2. Their model can be divided into two parts, a fixed time, which represents the time required to open and close doors, and passenger service time, which refers to the time required for passengers’ movement.

\[ DT_n = DF + \max_{i \in I} (\beta_0(\psi_i) + \beta_1(\psi_i) \mu_n^i + \beta_2(\psi_i) \lambda_n^i) \]  

2-2
\[
\psi_i = \frac{\lambda_n^i}{(\lambda_n^i + \mu_n^i)}
\]

where,
\[
\beta_0(\psi_i), \ \beta_1(\psi_i), \ \text{and} \ \beta_2(\psi_i) = \text{estimated coefficients based on the ratio of } \psi_i.
\]

Yoo et al., (2022) developed a linear regression model that considered the train car’s weight as a parameter related to passenger load presented in **Equation 2-4**. The weight of a train car is divided by the average weight of a passenger to estimate the total number of onboard passengers. Weighing scale is installed at each metro station to measure the weight of the train after the doors have closed and before the train departs. Study by Yoo et al., (2022) revealed that passenger load is directly correlated with the average boarding time, but not with the average alighting time. A similar approach can be found in the study by Lin & Wilson, (1992), where they included the number of boarding and alighting passengers in the last term, which is multiplied by the number of standing passengers. They remarked that passenger load add extra time to both the boarding and alighting processes. The major difference in their modeling approach compared to **Equations 2-4** is that they only considered standing passengers instead of total passenger load.

\[
DT_n = DF + \max_{i \in I}(\beta_0 + \beta_1 \mu_n^i + \beta_2 \lambda_n^i + \beta_3 \lambda_n^i \omega_n^i)
\]

In recent advancements within dwell time research, scholars have explored additional independent variables pertinent to dwell time analysis. For instance, Palmqvist et al., (2020) incorporated factors like arrival delay and scheduled dwell time into their study. Similarly, Coulaud
et al., (2023) extended their analysis to include scheduled dwell time, arrival time, and departure time, alongside actual dwell time, arrival time, and departure time.

2.2.2 Nonlinear Regression Models

Linear regression models consider constant average alighting and boarding times; however, in practice, the average alighting and boarding rate decrease with an increase in the alighting and boarding crowds. However, nonlinear regression models allow us to consider varying alighting and boarding rates by considering power terms. Weston (1989) developed a complex nonlinear regression model for the London Underground that accounts for many important factors. The original model developed by Weston, (1989) is shown in Equation 2-5 and the coefficients are subject to specific site conditions, although they can be changed to improve the prediction performance. Harris & Anderson, (2007) tested the model’s validity on other metro systems; however, they opposed on the site-specific nature of the passenger boarding and alighting time per passenger.

\[
DT_n = DF + \left[ \left( F \times \frac{A}{D} \right)^{\beta_1} + \left( F \times \frac{B}{D} \right)^{\beta_2} + \left( \beta_4 \times \left( F \times \frac{A}{D} \right) \times \left( F \times \frac{B}{D} \right) \right) \right] \times \left[ \beta_3 \times \left( 1 + \frac{F}{F^*} \right) \times \left( \frac{T-S}{D} \right) \right]
\]

where,

\( A \) = number of alighting passengers,
\( B \) = number of boarding passengers,
\( D \) = number of doors,
\( F \) = peak door/average door factor,
\( S = \) number of seats,
\( T = \) number of through passengers,
\( \beta_4 = \) estimated coefficient.

Factor F represents the proportion of passenger usage at the busiest door of the train compared to the average usage across all doors. This factor essentially quantifies the degree, to which the critical door is used more than the other doors on average.

A simplified nonlinear regression model was developed by Hor & Mohd Idrus, (2017), wherein they accounted for the decrease in average boarding and alighting rates per passenger with increasing passenger demand. Additionally, they included the extra time added for boarding passengers due to passenger load In this study, passenger load was considered in terms of crowding level, similar to the concept of level of service, which calculated using the number of passengers per square meter inside the train car. Kim et al., (2015) also employed a similar approach for the dwell time model, considering linear relationships for boarding and alighting time per passenger and a nonlinear relationship for the extra time added for boarding passengers due to onboard crowd. This study also factored in the degree of onboard crowd, which depended on the number of passengers standing near the boarding door.

2.2.3 Simulation Models

Simulation software such as Vissim, RailSys, and Nexus are utilized to explore alternative dwelling strategies aimed at minimizing dwell time or adhering to scheduled dwell times. While some studies have explored optimizing total travel time, incorporating dwell time as a factor within simulated models, this section specifically focuses on dwell time models and excludes studies that
do not directly address dwell time optimization. Dwell time simulation models play a crucial role in understanding and optimizing railway station operations. Several studies have utilized various simulation software to investigate dwell time dynamics and their impact on system performance. Lindfeldt, (2017) employed RailSys simulation to evaluate the capacity of a new railway line in Stockholm equipped with platform screen doors. The study revealed that extended dwell times caused by these doors limited capacity, highlighting the importance of considering dwell time distributions for effective capacity planning. Pu et al., (2022) utilized the integrated crowd and transit simulation platform Nexus to study the impact of pedestrian and train movements on system performance at urban railway stations. The study identified dwell time length and variation due to pedestrian movement as significant factors affecting system performance and platform crowding levels. Zhang et al., (2008) developed a cellular automata-based micro-simulation model for passenger behavior in Beijing metro stations using the STARLOGO programming language. The study highlighted the importance of factors such as individual desire and passenger cooperation in modeling alighting and boarding movements. Yoo et al., (2022) proposed a cost-effective strategy to reduce train delays by controlling passenger flow at station entry using simulation software Viswalk. The study demonstrated that passenger flow control could reduce scheduled dwell time delays, although causing congestion in the gate area, thereby improving rail capacity and reliability. Srikukenthiran & Shalaby, (2017) presented a proof-of-concept case study of the Greater Toronto transit network using Nexus simulation platform, integrating crowd dynamics and transit network simulation. The study illustrated how the platform could be used for disruption management by analyzing the impact of disruptions and implementing response strategies.
2.2.4 Machine Learning Models

In this section, literature on dwell time prediction using machine learning techniques is explored, which includes Extreme Learning Machine (ELM) models, fuzzy logic-based models, dynamic updating methods, and the Hunter Prey optimizer (HPO) algorithm. Several studies have explored machine learning approaches to estimate dwell time. One notable approach, as explored by Alvarez et al., (2014), introduced a novel approach to estimate dwelling time, combining origin-destination matrices with fuzzy logic. Their method, employing three levels of passenger flow in a train's coach, aimed to provide a more accurate prediction of real-world effects, offering a practical application for mass transit planning. Using the similar approach but leveraging the principle of maximum entropy, Alvarez et al., (2015), estimated dwelling times by emphasizing the use of an artificial-intelligence technique. Yang et al., (2020) investigated the relationship between the width of alighting areas and the overall efficiency of alighting and boarding processes. Their study utilized an enhanced social force model, which incorporated fuzzy logic principles to dynamically adjust passenger speeds based on factors such as train dwell time and passenger volume. Through this approach, they explained how passenger behavior and infrastructure design influence overall system efficiency. Chu et al., (2015) proposed a novel method for estimating urban rail dwell times using ELM neural networks. This approach, validated using real-world data from the Beijing subway, aimed to provide a robust and accurate means of estimating dwell times. Meanwhile, Yang et al., (2023) introduced the HPO Back Progression Neural Network prediction model, integrating advanced optimization algorithms to improve the accuracy of alighting and boarding time predictions. Drawing insights from field investigations, their model sought to address the complexities of passenger behavior and system dynamics, thereby enhancing prediction accuracy. Furthermore, Pang et al., (2023) contributed to this body of research by
developing a comprehensive train dwell time model that considers real-time passenger flow fluctuations. Their approach, incorporating an averaging mechanism and dynamic updating method, aimed to provide more accurate predictions by adapting to changing passenger volumes and behavior patterns.

Overall, various approaches to estimating dwell time have been explored, including linear regression models, nonlinear regression models, and simulation models. While simulation models offer greater accuracy in predicting dwell time, they often lack the ability to estimate coefficients for independent variables within the model. In contrast, regression models allow for the estimation of coefficients for independent variables, facilitating comparisons of the influence of each variable on the dependent variable. Despite the existence of diverse dwell time modeling approaches, the literature failed to consider passenger flow with exact number of onboard passengers in the same model.

2.3 Data Collection Methods for Dwell Time-Analyses

Data required for dwell time estimation includes the number of boarding, alighting and onboard passengers and train stopping time. Collecting sufficient and accurate data is crucial to support the accuracy of the dwell time prediction using the developed model. In the literature, various data collection methods have been explored for passenger demand and dwell time estimation. Some studies used manual data collection method (Christoforou et al., 2017; Chu et al., 2015; M. Yang et al., 2018; X. Yang et al., 2020), but this method has limitations, such as small sample sizes and the need for multiple surveyors when recording demand for each door. Video cameras have also been used (Adachi et al., 2019; Hayes et al., 2022; Yoo et al., 2022), however, extracting passenger counts from the recorded videos converting the collected data from cameras
to passenger counts requires additional effort and may pose privacy issues for detailed investigations external to the transit agency. Additionally, both methods mentioned above have limitations in selecting multiple stations for comparing boarding and alighting time per passenger. Moreover, they often overlook platform characteristics, such as entrance locations, presence of obstacles, and train stopping positions.

Another approach involves using automatic data collection, such as AFC data, but they mostly assume that all boarding passengers will board the first arriving train (Wolofsky et al., 2019). In addition, a major drawback with this approach is that it does not distinguish the door with the highest passenger demand, instead uses the total demand for the entire train to estimate the dwell time. This may lead to an inaccurate estimation of the actual dwell time at the critical door which can impact the overall accuracy of the dwell time model. Another use of AFC data is to calculate the percentage of denied boarding passengers by combining it with the train schedule or AVL data in real-time or offline (Chen et al., 2022). In a recent study, AVL and APC data were used for dwell time modeling, but the study also considered total number of boarding and alighting passengers per train instead of per door (Coulaud et al., 2023).

The literature on dwell time estimation in public transit systems emphasizes the need for accurate data collection methods, including manual surveys, video cameras, and automatic sensors like AFC, APC, and AVL systems. However, existing methods have limitations such as small sample sizes, privacy concerns, and overlooking passenger demand at each door. Notably, emerging technologies using AVL and APC system have ability to collect data for each door with large sample size; however, recent studies focused on total passenger counts per train rather than per door.
2.4 Passengers Walking Behaviour and Modeling

This section reviews the literature on pedestrian walking behavior and its modeling in microsimulation models extensively.

2.4.1 Pedestrian Walking Behaviour

Pedestrian walking behaviours follow fundamental traffic flow theories based on fluid mechanics, where basic elements are flow, density, and speed (Fruin, 1971; Seneviratne, 1983). However, the degree of freedom for pedestrians is higher than the degree of freedom for the traffic flow (Seneviratne, 1983). Moreover, observations show that the average speed and flow capacity of passenger movements has significantly higher variations compared to vehicles (Seneviratne, 1983). Nonetheless, the theory of vehicle traffic flow can be directly used in pedestrian flow simulations (Daamen, 2002). Galpin et al., (2018) presented a methodology to use agent-based simulations for individuals and their flow based on the partial differential equations of density using Collective Adaptive Resource-sharing Markovian Agents (CARMA) modeling language for mesoscopic simulation models. Tordeux et al., (2018) concluded that mesoscopic models are easy to build and faster to simulate than microscopic models; moreover, this model is more effective and accurate than macroscopic models.

Research on passenger behavior on platforms has identified two primary categories of passengers: those who select their waiting location based on their origin station and those who choose it based on their destination (Fang et al., 2019). Passengers, who select their waiting location based on their origin station are influenced by the physical layout of the platform, often opting to wait near pillars or ticket machines (Davidich et al., 2013) and utilizing available roof coverage, particularly during unpleasant weather conditions (Bosina et al., 2015; Nash et al.,
Additionally, studies have observed that passengers tend to cluster near platform entrances (Krstanoski, 2014; Lee et al., 2018; Peftitsi et al., 2020; Van Den Heuvel, 2016; Wiggenraad, 2001). This behavior varies depending on the time gap between passenger arrivals on the platform and the train's departure time, with passenger circulation decreasing as this time gap narrows (Bosina et al., 2015; Fang et al., 2019; Fox et al., 2017; Wu & Ma, 2013). The second group of passengers, who choose their waiting position based on their destination, tend to position themselves closest to the exit at their destination station (Fang et al., 2019; Zheng, 2018). This tendency is especially pronounced when passengers are familiar with the surroundings at their destination and when exits are oriented towards specific directions, such as the city center or connecting modes of transportation (Bosina et al., 2015; Nash et al., 2006).

Although there are numerous studies on pedestrian walking behaviour and pedestrian simulation modeling to capture passengers walking behaviour more realistically, the literature on studies related to walking behaviour in integration with public transit is still limited. In the literature, three types of walking behaviour modeling approaches can be found (a) the social force model, in which the motivation of a pedestrian to make a certain decision, is considered as a force (Helbing & Molnár, 1995; Lakoba et al., 2005; Lewin, 1951), (b) cellular automata, in which each cell has a predefined local rule for the behaviour of each pedestrian (Blue & Adler, 2001; Schadschneider, 2001), and (c) discrete choice models, which assume that the route choice decision of pedestrians depends on the utility of routes (Antonini et al., 2006; Venegas et al., 2004). Pedestrians normally take the shortest path to their destination without detours, and their walking speed depends on the distance from other pedestrians, the density and speed of other pedestrians (Helbing & Molnár, 1995).
2.4.2 Pedestrian Flow Modeling and Calibration

Simulating passenger movements allows us to model passengers’ flow and their interactions within public transit system. Daamen, (2002) developed a pedestrian flow simulation model considering pedestrian behaviour, route choice, and activity performance. Another study on the Beijing transit network by Zhang et al., (2008) presented a microsimulation model for boarding and alighting passengers based on a cellular automata approach using the STARLOGO modeling tool. Since it can model the interactions between pedestrians, automobiles, Heavy Goods Vehicles (HGVs), and any other forms of transportation, Vissim has become one of the most well-known traffic flow microsimulation software (Choa et al., 2004; Perrone et al., 2006). VDOT Vissim User Guide, (2020) describes a stepwise procedure to develop a model in Vissim and then, debug and calibrate it. Debugging is a distinct process from model calibration, which involves reviewing the input data and correcting error logs. Patange & Bhakhtyapuri, (2017) developed a pedestrian network for a railway station in Vissim to compare simulated passenger flow on the ramp and platform with the observed data from both locations. Another study created a network to investigate the behaviour of various transportation modes and their speed and travel time when a bicycle lane is included (Rios et al., 2021).

It is important to understand the behaviour of passengers on a transit stop platform to calibrate the model. Daamen, (2002) explained events in the passenger boarding and alighting process to model passenger interactions within the transit system. Daamen, (2002) divided walking activities on platform into three main parts as the “walk link”, “wait link”, and “service link”. The type of each link depends on the destination of the passenger. For example, when modeling the boarding process for waiting passengers on the platform, the walk link will be the time required to walk from the standing point to the vehicle door. The wait link will be when passengers wait for
the alighting process or the boarding process for the passengers who came first to finish, and the service link will be the time required for the passenger to board.

Overall, due to the inefficiencies in measuring minor delays in urban rail operations, dwell time-related delays are poorly recorded in the literature. These small variations in dwell time can significantly affect the total travel time, which can be reduced by developing effective passenger management strategies. It has been found that identifying the critical door and considering passenger flow is required for an accurate estimation of dwell time in urban rail systems. Literature is lacking in supporting automated large data collection methods for identifying the critical door for dwell time analyses. Moreover, to the best of our knowledge, there is not any literature available for comparison of fraction of boarding passengers with consideration of exact passenger load for the car with critical door and comparison of impact of station-specific geometric characteristics on dwell time. Advanced computational algorithms and microsimulation methods enable us to develop strategies for the dwelling process and test them visually. However, the first and most crucial step in studies on dwell time using microsimulation is to develop an accurate model, which can replicate the observed passenger interactions and then, calibrate the developed model to improve its efficiency. The next step is developing new strategies to improve the dwelling process and passengers’ travel experience. This research addresses these limitations found in the literature by using AVL and APC data to identify critical doors for dwell time modeling. Additionally, the study considers the friction between boarding and alighting passengers and divides the data into multiple groups based on the fraction of boarding passengers, allowing for a comparison of changes in average boarding and alighting time per passenger. Moreover, this research develops a microsimulation model for the dwelling process, calibrates it to improve simulation performance, and tests alternative dwelling strategies to minimize dwell time. The
consideration of the fraction of boarding passengers in this study is influenced by the work of Wirasinghe & Szplett, (1984), wherein they categorized passenger flow into five distinct groups. While this study adopts a similar approach by dividing passenger flow into three groups, we also incorporate the passenger load, a factor not addressed in their research.

Chapter 3 describes the methodology developed for this thesis. It comprises two parts: the methodology for developing a regression model for dwell time considering passenger demand at critical doors and onboard passenger load, and the latter part focuses on the methodology for the microsimulation model which is used to test alternative dwelling strategies.
CHAPTER 3. METHODOLOGY

The methodology chapter consists of two parts: the first part involves developing a regression model, referred as regression analysis, to construct the dwell time model and investigate the impact of station design and passenger flow on dwell time. The second part focuses on developing a microsimulation model for studies related to dwelling process, referred to as the microsimulation modeling. Furthermore, a methodology is developed for passenger flow modeling to enhance the performance of simulation models.

3.1 Regression Analysis

This section is further divided into three parts: data collection, data filtering, and dwell time model development. In the data collection step, essential information is gathered using APC sensors installed in the train cars and GPS data for determining vehicle locations and the stopping times. The data processing step follows to ensure accuracy by removing anomalies, handling missing values, and estimating passenger load. Lastly, the dwell time model is developed to analyze the impact of station-specific geometric design, passenger flow, and passenger load on dwell time.

3.1.1 Data Collection

Data collection for the Light Rail Vehicle (LRV) includes using APC sensors for recording passenger counts and GPS system for determining vehicle locations. These sensors are installed on train cars and play a crucial role in capturing valuable information about both train operations and passenger movements. The vehicle location is continuously recorded for every 5 seconds by
the AVL system, and the collected data also includes whether the doors are closed or open at each interval. The APC sensors trace passenger movements (i.e. boarding and alighting) for each door. Typically, an APC system uses two sensors (e.g., infrared or ultrasonic) to detect passenger movements (Furth et al., 2006). The infrared sensor counts passengers when passengers pass through a beam of infrared light. While the ultrasonic sensor detects the direction of motion using high-frequency sound waves. The data collected by these sensors is stored as a file with CSV format and consists of three types of records: cyclic records, stop records, and passenger records. Cyclic records are reported at regular 5-second intervals and provide essential data points that help analyze various aspects of train operations. Stop records, as the name suggests, capture arrival and departure time, when the trains halt at stations to facilitate passenger movements. Finally, passenger records contain information about passenger movements for each door, enabling a comprehensive understanding of passenger flow.

3.1.2 Data Processing

The data processing steps presented in Figure 3.1 involves several important procedures in organizing and extracting meaningful insights from the collected datasets. Firstly, the data was separated into two distinct categories: stop records and passenger records. Passenger records capture data related to boarding and alighting passengers at each stop, including door numbers and passenger counts. Stop records report the arrival and departure times along with the coordinates for the location of each stop. As we know each train car has four doors. The data for each door in the raw data was stored in separate row with a column name “door number”. To facilitate the analysis, the door data for each car was consolidated into a single row by adding four columns for boarding passengers' data and four columns for alighting passengers' data. This allows for a more
comprehensive understanding of passenger movements within each train car and facilitates subsequent calculations of passenger load separated for each car. Next, the stop and passenger datasets were combined by comparing the date, time, and vehicle number fields. By matching these key variables, the two datasets were merged, providing a holistic view of passenger activity at specific stops. In order to enhance the analysis with meaningful station names, coordinates from the dataset were matched with station location data. A least-distance matrix approach (Yuan et al., 2021) was utilized to determine the closest station to each recorded location, and the corresponding station name was added to the dataset.

![Data processing steps](image)

**Figure 3.1: Data processing steps.**

The data was then further processed by separating individual trips, identifying consecutive stations, and removing incomplete trips from the dataset. To analyze passenger occupancy, the passenger load was calculated by subtracting cumulative alighting passengers from cumulative boarding passengers at each stop for each train trip and each train car. Dwell time, which represents the time, during which the train is fully stopped at a station, was calculated by comparing the departure and arrival times for each station. The dataset was further segregated for each station,
allowing for statistical analysis of key metrics such as minimum, maximum, standard deviation, and percentile values for boarding, alighting, passenger load, and dwell time information. These statistical measures provide a comprehensive understanding of passenger behaviour, station activity, and variations in performance indicators across different locations.

3.1.3 The Dwell Time Estimation Model

In this study, three independent variables were considered for the dwell time model: the number of boarding passengers at each door \((\lambda_{ni})\), the number of alighting passengers at each door \((\mu_{ni})\), and the passenger load per car \((\omega_{ni})\). The total dwell time \(DT_n\) of a train \((n)\), as presented in Equation 3-1, was then modeled as a function of these three variables:

\[
DT_n = f(\lambda_{ni}, \mu_{ni}, \omega_{ni})
\]

where,

\(DT_n\) = the total dwell time of train \(n\),
\(\lambda_{ni}\) = the number of boarding passengers at door \(i\) of train \(n\),
\(\mu_{ni}\) = the number of alighting passengers at door \(i\) of train \(n\), and
\(\omega_{ni}\) = the passenger load of the car, which has door \(i\) on train \(n\).

The critical door number \(M_n\) for train \(n\) is the argmax of the sum of alighting passengers, boarding passengers, and the product of boarding passengers and passenger load, as presented in Equation 3-2. Using this equation, we want to identify the door \(i\) which was heavy used for the passenger movement from all doors. It should be noted that four car train has 16 doors, and we are interested in finding the door with the highest passenger demand.
\[ M_n = \arg\max_{\text{ref}} (\lambda_n^l + \mu_n^l + \lambda_n^l \omega_n^l) \forall n \in N \] 3-2

It should be noted that since each car has four doors, \( \omega_n^{M_n} \) for the four doors on the car that has the critical door would be the same. According to the existing literature, it is noted that passenger load affects the boarding time per passenger but does not have a significant impact on the alighting time per passenger. A stepwise regression analysis was conducted to identify the most suitable model by eliminating statistically insignificant variables. Different combinations, encompassing both linear and nonlinear forms, were explored to establish a regression model. After meticulous comparison of the results for all the selected models, the regression model \textbf{Equation 3-3} was selected as the optimal model featuring the statistically significant variables and higher \( R^2 \) value. It should be acknowledged that advanced machine learning techniques allow for higher prediction accuracy. However, these techniques have limitations in predicting coefficients for independent variables. On the other hand, regression models allow for easy estimation of coefficients. Our aim in this thesis is to compare coefficients for each station to assess the impact of station design on dwell time. In this model, the second and third terms account for the combined time required for both boarding and alighting processes. The fourth term within this equation is additional time that is added as a direct result of passenger load. It should be noted that this added time increases with increasing passenger load or the number boarding passengers.

\[ DT_n = \beta_0 + \beta_1 \lambda_n^{M_n} + \beta_2 \mu_n^{M_n} + \beta_3 \lambda_n^{M_n} \omega_n^{M_n} \] 3-3

where,
As mentioned by Parkinson & Fisher, (1996) and Wirasinghe & Szplett, (1984), the average boarding and alighting time per passenger can vary based on the fraction of boarding passengers and the overall demand. To capture this relationship, the fraction of boarding passengers (Equation 3-4) was defined as the ratio of boarding passengers at door $i$ of train $n$ to the total demand for the same door.

\[
\psi_n^i = \frac{\lambda_n^i}{(\lambda_n^i + \mu_n^i)}
\]

Equation 3-4 allows us to account for the friction between boarding and alighting passengers and compare the changes in boarding and alighting time per passenger. Table 3.1 presents passenger groups based on the fraction of boarding passengers.

Table 3.1: Description of Groups Based on the Fraction of Boarding Passengers

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Fraction of boarding passengers</th>
<th>Group description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\psi_n^i \leq 0.32$</td>
<td>Alighting-dominant</td>
</tr>
<tr>
<td>2</td>
<td>$0.33 \leq \psi_n^i \leq 0.66$</td>
<td>Mixed passenger flow</td>
</tr>
<tr>
<td>3</td>
<td>$0.67 \leq \psi_n^i \leq 1$</td>
<td>Boarding-dominant</td>
</tr>
</tbody>
</table>

Finally, the coefficients for each group were estimated using the Ordinary Least Square (OLS) method. The estimated coefficients were then compared with selected stations for both peak and off-peak periods. This comparison allowed for the evaluation of the dwell time model’s performance and its ability to accurately capture the variations in dwell time at different stations during different times of the day.
3.2 Microsimulation Model

This section is focused on developing a dwell time model and then enhancing the simulation model considering passenger flow modeling and testing alternative dwelling strategies to minimize dwell time. The aim is to develop a generalized methodology that can be applied to any cases in the future. This section presents dwell time modeling followed by discussing the methodology to enhance simulation models by considering passenger flow modeling.

3.2.1 Dwell time model and relevant parameters in Vissim

Vissim can integrate with programming languages such as Python, MATLAB, and C++ in order to run simulations through a COM server. This enables the user to modify the predefined functions in Vissim according to users’ requirements. In this thesis, the Vissim model is accessed through the Python language using the COM interface to enhance the simulation model developed in the Vissim environment. Figure 3.2 represents the critical parameters related to dwell time in Vissim. As discussed in the previous chapter, the dwell time model has two components, fixed time and passenger service time.
3.2.2 Calibrating the simulation model

As discussed in the Literature Review chapter, model calibration is an important step in the simulation process to measure the efficiency of the developed model. Calibration can be done using automated or manual methods. This research used a manual calibration method, and the
required parameters were set to match the simulated dwell time with the observed dwell time. The calibration process aimed to minimize the mean of the squared errors (MSE).

The calibration process starts with setting values for parameters based on the observations and the average boarding and alighting rates are used based on estimations from the regression model. The first step in the model calibration is to identify appropriate parameters in simulation models and assign the observed values to it. In this case, the door lock duration before departure, door closure delay, and stop clearance time are obtained from the field observation. Along with it, the passenger service time that is boarding and alighting time per passenger are obtained from the regression analysis. Afterward, simulations are performed to collect results for the boarding, alighting, and passenger load and the total dwell time. Then, the simulation results are compared with the observed dwell time data to calculate the MSE value. Parameter values related to the passenger walking behaviour are adjusted during the calibration process, and results are compared to select the best combination of parameter values. Although calibration of dwell time parameters can minimize the MSE value, it was observed that the trend of simulated dwell time was not always following the trend of observed dwell time. Therefore, passenger flow modeling was considered to deal with this problem and enhance the simulation results to capture the actual trend, and the methodology is discussed in the following subsection.

3.2.3 Passenger Flow Modeling

This section presents some key pedestrian flow models that are implemented in this research to enhance the dwell time in our simulation model. Predicting pedestrian walking behaviour is useful in a variety of contexts. However, collective behavioural patterns such as clustering, queues, and lanes make pedestrian modeling more complicated. In macroscopic
modeling, pedestrians are described as a flow that follows the same properties as traffic flow. Fruin, (1971) used a similar idea for the pedestrian flow movement and developed an equation for the pedestrian flow movement. The author defined the flow volume ($P$) in pedestrians per meter per second equal to the average speed ($S$) in meters per second and passenger density ($D$) in pedestrians per square meter as shown in Equation 3-5.

$$P = S \times D$$  

3-5

As pedestrian density increases, the available free space to walk and the probability of passing slow-moving pedestrians decreases, resulting in a decreasing average walking speed. According to Equation 3-5, when density on the platform increases, boarding passengers' walking speed decreases, leading to an increase in the inter-arrival time of boarding passengers at the boarding door. Passengers’ walking speed can be estimated from the fundamental traffic flow diagram using a linear relationship of speed and density. However, the speed of passengers varies depending on passenger characteristics such as age, gender, any baggage, and the geometric design of infrastructures like the width of a walkway, and obstructions. As found in the literature, pedestrian flow follows the same traffic flow theory; however, their speeds at free flow and density at jamming are different from those of traffic flow. Also, pedestrians have more degrees of freedom, which means they have more flexibility in making turns and changing lanes. Greenshields et al., (1933) expressed a linear relationship between the maximum speed and density for the traffic flow as shown in Equation 3-6. The same equation can be used for the pedestrian flow where $D_{jam}$ and $S_{free}$ are the density at jam when passengers cannot walk or the speed is zero (pedestrians per square meter) and free flow speed (meters per second), respectively.
\[ S = S_{\text{free}} - \frac{s_{\text{free}}}{D_{\text{jam}}} \times D \]

The density of a platform can be calculated by considering the number of passengers on the platform and the area of the platform and thus, can be obtained from Equation 3-7.

\[
\text{Actual speed}(S_1) = S_{\text{free}} - \frac{s_{\text{free}}}{D_{\text{jam}}} \times \left( \frac{\text{Number of passengers on platform}}{\text{Area of platform}} \right)
\]

Figure 3.2 shows the events in simulating a public transit station. As soon as the train arrives, it stops at the platform, and the next event is the door opening. The time required for door opening relies solely on the vehicle type and technology used for the vehicle design. Irrespective of passenger distribution, alighting passengers choose the closest door to them. At the same time, boarding passengers on the platform start clustering near the door and wait to finish the alighting process. Once the alighting process is finished, boarding passengers start boarding based on the First In, First Out (FIFO) rule. In some instances, passengers start boarding just after the doors open, this process is represented by the dotted line in Figure 3.2. The time needed for all passengers to board and alight is passenger service time. The next event is door closing and the time required to close the doors also relies on the technology and vehicle type. In some cases, if a train is earlier than the scheduled departure time, it waits until the scheduled departure time. Otherwise, the train departs just after closing all doors.
Boarding passengers cluster during the alighting process or before the boarding process begins. Based on the longest distance needed to walk on the platform and the total alighting time, the speed of boarding passengers can be calculated as,

\[
\text{Required speed (}\, S_2) = \frac{\text{Longest walking distance on platform}}{\text{Alighting time}}
\]

The longest walking distance on the platform depends on the distance between the rail door and the entrance of the platform when the entrance is located at the edge of the platform. However, this longest walking distance on the platform is case-specific and depends on the geometric design of the platform.

The alighting time can be obtained by multiplying the total number of alighting passengers by the average alighting time per passenger. Suppose the required speed (\( S_2 \)) is greater than the actual speed (\( S_1 \)); boarding passengers cannot cluster before the boarding process begins due to the crowding on the platform. As the difference between the required and actual speed increases, there could be major discrepancies between the simulated and the observed dwell times, which can be enhanced by considering passenger flow and interactions within the transit system. As discussed earlier, boarding passengers start clustering around a door during the alighting process; therefore, when alighting passenger demand is relatively lower than boarding passengers, all boarding passengers do not get a chance to cluster around a door, leading to higher inter-arrival time between boarding passengers which results in a higher MSE. Therefore, in order to improve the simulation model and match the observed dwell times with simulated dwell times, the inter-arrival time between boarding passengers should be reduced by considering passengers walking behaviours, e.g., increasing the walking speed of passengers on a platform and increasing the reaction time of
passengers in the simulation model. This can minimize the differences in simulated and calculated
dwell times and make better predictions. In Vissim, the desired speed factor represents the walking
speed of passengers in percentage which is a very important parameter to capture the realistic
behaviour of passengers depending on the type of area, e.g., decreasing walking speed on
stairways, ramps, escalators, etc., according to the observed trend. In this study, we modify the
walking speed of passengers in the waiting area based on the difference between the actual speed
and the required speed obtained by Equation 3-9, respectively. To adjust the walking speed in
simulation, the desired speed factor needs to be changed according to Equation 3-9, only when
the required speed is greater than the actual speed; otherwise, the desired speed factor is considered
1, which represents the actual walking speed.

\[ \text{Desired speed factor} = 1 + \frac{S_2 + S_1}{S_1} \] 3-9

Moreover, we adjusted values of React to n pedestrians and Tau parameters in the
simulation model for boarding passengers and observed that a reduction in the value of these
parameters results in a closer simulated dwell time to the observed dwell time. This is because
when a train arrives at a platform, the boarding passengers react quickly and start getting closer to
the boarding door. Table 3.2 includes definitions related to the dwell time model in Vissim.

Table 3.2: Dwell Time Components in Vissim and Their Definition and Units.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Door lock duration before departure</td>
<td>The duration of time, after which the doors are fully closed until the actual departure of the vehicle.</td>
<td>Seconds</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
<td>Unit</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Door closure delay</td>
<td>The duration of time, when the passengers have cleared all doors, and the time, at which the doors start closing.</td>
<td>Seconds</td>
</tr>
<tr>
<td>Clearance time</td>
<td>The time needed for a train to stop including any other possible delays.</td>
<td>Seconds</td>
</tr>
<tr>
<td>Door lock duration</td>
<td>The time needed to open or close train doors.</td>
<td>Seconds</td>
</tr>
<tr>
<td>Alighting time</td>
<td>The average time needed for a passenger to alight.</td>
<td>Seconds per passenger</td>
</tr>
<tr>
<td>Boarding time</td>
<td>The average time needed for a passenger to board.</td>
<td>Seconds per passenger</td>
</tr>
<tr>
<td>Desired speed</td>
<td>The percentage of walking speed in an area which is used to capture the realistic movement of passengers on stairs, escalators, walkways, etc.</td>
<td>Percentage</td>
</tr>
<tr>
<td>Actual speed</td>
<td>Walking speed of passengers at free flow capacity on a platform.</td>
<td>Meter per second</td>
</tr>
<tr>
<td>Required speed</td>
<td>The walking speed of a boarding passenger to cluster near the closest door before the alighting process is finished.</td>
<td>Meter per second</td>
</tr>
</tbody>
</table>

In this chapter, passenger flow is categorized into three distinct groups: alighting-dominant, boarding-dominant, and mixed passenger flows, based on the fraction of boarding passengers at critical doors. A regression model is proposed for estimating dwell time, which is then utilized to assess the impact of passenger flow and geometric design of station platforms on dwell time. Additionally, a methodology for enhancing the performance of the microsimulation model is
provided, followed by a procedure for developing a microsimulation model for dwell time estimation.

In Chapter 4, stations with unique platform designs are selected to compare the impact of station design on dwell time. Furthermore, the results for the regression model, considering passenger flow, and the discussion on outliers in the selected data are presented.
CHAPTER 4. DWELL TIME MODELING

In this thesis, we have obtained LRV data recorded by Calgary Transit. The data covers the entire months of July 2022 and July 2023, as well as 15 days each from January, April, and September 2023. Calgary Transit has data available for all days of operation since the APC and AVL sensors are installed on vehicles. However, not all vehicles are equipped with these sensors. The vehicles equipped with these sensors belong to the same model and have identical interior designs. This model features a longitudinal seating configuration, which provides more space for standing passengers. This train model has a capacity of 200 passengers per car (Calgary Transit launches, 2015). It is a high-floor vehicle and platforms in Calgary are also raised, which allows for level boarding. Usually, Calgary Transit runs three-car trains, and sometimes four-car trains during the peak hours (Calgary Transit launches, 2015). The month of July was intentionally chosen as the annual outdoor festival of Stampede lasting for 10 days is held during this month. During this festival many tourists and local visitors are attracted to major locations in the downtown core and the LRT system operates 24 hours a day. In 2022, a total of 1.2 million people visited the Stampede (Back in the Saddle, 2022). On the first day, there were 130 thousand attendees. In 2023, 1.3 million people visited during 10 days of Stampede event (Fantastic year, 2023). During this period, the demand often exceeds the system’s capacity, leading to challenges in managing passenger flow. It is important to note that the obtained data includes 85 regular days and 20 Stampede days. The blue line and the red line of Calgary Transit LRT network share the same track in part of the Central Business District (CBD), which has road grade crossings. The remainder of this chapter explains more details of the case study and is divided into four parts: station selection, descriptive analysis, results, and identifying outliers in the data.
4.1 Station Selection

Calgary’s LRT Network is comprised of two lines: the blue inbound and the red line, as shown in Figure 4.1 (Calgary Transit, 2023). The orange dots in the Figure 4.1 represents the selected stations for this study. The blue line operates at a headway of 5 minutes during peak hours, while the red line operates at a headway of 3 minutes during peak hours (Red line, 2023).

Figure 4.1 Calgary’s light rail transit map (source: Calgary Transit).

For the purpose of this study, we chose stations with unique platform designs, such as those featuring side and middle platforms, as well as single and multiple entrances. Consequently, six stations were carefully selected for analysis, each showcasing unique platform layouts. Stations within the CBD are intentionally avoided because of road grade crossings. These crossings complicate the differentiation between train stopping by passenger boarding or alighting and those stopping by the traffic signal at junctions. Often, trains at these stations need to wait for traffic
signals, influencing the dwell time more significantly through the traffic signal rather than passenger movements. Among the chosen stations (Figure 4.2(a), (b), and (c)), three featured a middle platform design, while the remaining three stations (Figure 4.2(d), (e), and (f)) were featured with side platforms. We have annotated yellow boxes to indicate the platform locations and red boxes to highlight where the platform entrances are situated. No stations on the blue line were selected because of the line’s lower frequency in comparison to the red line which greatly affects the sample size collected within the same duration of data collection.

Figure 4.2 Station layout, (a) Erlton, (b) Chinook, (c) University (d) Sunnyside (e) 39th Avenue and (f) Lions Park, red and yellow boxes to show the entrances and platform locations, respectively (source: Google Map).

Specifically, the stations in question are as follows: Erlton Station (Figure 4.2(a)) with a middle platform and two entrances located at both ends; Chinook Station (Figure 4.2(b)), also a middle platform, but with only one entrance situated at end of the platform; University Station (Figure 4.2(c), featuring a middle platform and two entrances positioned in the middle of the platform; Sunnyside Station (Figure 4.2(d)), has side platforms with multiple entrances at end of a platform and in the middle; 39th Avenue Station (Figure 4.2(e)), offering side platforms and two entrances positioned at end and middle of the platform; and finally, Lions Park Station (Figure
4.2(f), has side platforms and multiple entrances at end of a platform and in the middle. All selected stations with middle platforms are at least 8 meters wide except the Erlton Station, while stations with side platforms are at least 4.5 meters wide. To supplement the description of these stations, Table 4.1 is provided, furnishing a concise summary of the platform and entrance configurations for each station. Lions Park Station features two entrances at both ends of the platform and an additional two entrances in the middle of the platform dedicated to the inbound direction. For the outbound direction, there are only two entrances situated at both ends of the platform. However, on this side of the platform, a sidewalk at a lower grade is connected to the platform, enabling passengers to enter or exit the platform at the midpoint. This sidewalk shares the same entrances as illustrated in Figure 4.2(e), while on the opposite side of the sidewalk, a highway is located, preventing pedestrians from engaging in jaywalking. This sidewalk is curial for alighting passengers to exit the platform, especially when higher number of boarding passengers are waiting on the platform.

Table 4.1 Description of Platform Design.

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Station Name</th>
<th>Platform Location</th>
<th>Entrance Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Erlton</td>
<td>Middle platform</td>
<td>Two entrances at both ends of platform</td>
</tr>
<tr>
<td>2</td>
<td>Chinook</td>
<td>Middle platform</td>
<td>One entrances at end of platform</td>
</tr>
<tr>
<td>3</td>
<td>University</td>
<td>Middle platform</td>
<td>Two entrances at middle of platform</td>
</tr>
<tr>
<td>4</td>
<td>Sunnyside</td>
<td>Side platforms</td>
<td>Multiple entrances at end and middle</td>
</tr>
<tr>
<td>5</td>
<td>39th Avenue</td>
<td>Side platforms</td>
<td>Two entrances at end and middle</td>
</tr>
<tr>
<td>6</td>
<td>Lions Park</td>
<td>Side platforms</td>
<td>Multiple entrances at end and middle</td>
</tr>
</tbody>
</table>
4.2 Descriptive Analysis

The first crucial task was to identify the independent variables from the dataset. In the methodology section, a step-by-step process for data processing was explained and followed to extract the relevant independent variables required for the analysis. The raw dataset was processed to extract key variables, including the number of alighting and boarding passengers for each door, the passenger load per train car, and the dwell time per train. Subsequently, statistical calculations were performed on the extracted variables. This involved analyzing the data to obtain various measures such as averages, medians, standard deviations, etc., to gain insights into the patterns and trends related to alighting, boarding, passenger load, and dwell time. The plot in Figure 4.3 represents the fluctuations in the sum of passenger demand for all doors at different stations throughout the day, categorized by the two directions: (a) northbound and (b) southbound. The y-axis represents the passenger demand, while the x-axis consists of hourly data, starting at midnight. Calgary Transit provides LRT services 24 hours a day during the Stampede event, but there is no service between 1:00 AM and 4:00 AM on regular days. The study identified a morning peak period, which spans from as early as 7:30 AM to 9:30 AM, and an evening peak period from 4:00 PM to 8:00 PM. The analysis did not segregate the data between peak and off-peak periods. It is important to note that there were instances with exactly zero alighting or boarding demand, contributing to a lower average demand per train car.
Figure 4.3: Passenger demand distribution for (a) northbound, and (b) southbound directions.

Figure 4.4 depicts box plots for (a) alighting passengers, (b) boarding passengers, (c) passenger load, and (d) dwell time to gain more insights from the selected stations on passenger demand along with the passenger load and dwell time for each direction. In this figure, the y-axis represents the passenger demand, measured as the sum of passengers at all doors for each train car or dwell time in seconds, with station names plotted along the x-axis. It was not feasible to analyze the passenger demand for each individual train, as Calgary Transit operates both 3-car and 4-car trains. The comparison of passenger demand and dwell time between the inbound and outbound revealed interesting patterns. In general, the median of the number of alighting passengers was higher for the outbound direction, while the median of the number of boarding passengers was higher for the inbound direction. This symmetry is a result of the nature of passenger trend travel to the CBD, where during morning peak hours, there is a high demand for boarding in the inbound direction, and during the evening peak hours, there is a high demand for alighting in the outbound direction. Chinook station exhibited a higher median value for boarding passengers in the inbound direction and alighting passengers in the outbound direction. Furthermore, Erlton Station, located near the Stampede Park, exhibited a higher demand on Stampede days. Conversely, the station
experienced a lower demand on regular days. As a result, this lower demand contributes to the observed decrease in medians for both boarding and alighting passengers in the presented data plot.

Figure 4.4: Box plots for (a) alighting, (b) boarding, (c) passenger load, and (d) dwell time.

The median of the passenger load ranged from 45 to 65 passengers per car. University station, situated far from the downtown area, exhibited the lowest passenger load per train car. This trend suggests that passenger load tends to be higher when trains enter or exit the CBD. For both inbound and outbound directions, the median of the dwell times ranged from 28 to 32 seconds. A consistent observation in dwell time distribution was that the median of the dwell times was higher for the inbound direction compared to the outbound direction.
4.3 Results

The data for each station was categorized into two groups based on the selected direction. Subsequently, a regression analysis was conducted individually for each station and each direction using the OLS method available in the Statsmodels library in Python. Equation 3-3 is used to conduct regression analysis for each case mentioned above, and the results are presented in Table 4.2. The primary goal of this exercise was to investigate into the relationship between station design and dwell time. Additionally, another objective was to estimate the average boarding and alighting time per passenger, a valuable metric for potential future research endeavors aimed at predicting dwell time under varying passenger demand scenarios at selected stations, facilitating the implementation of new policies. Furthermore, we aimed to assess whether the estimated parameters remain consistent across all stations or exhibit variations influenced by platform geometry.

Table 4.2: Results for Without Considering Passenger Flow.

<table>
<thead>
<tr>
<th>Station Name</th>
<th>Direction</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3 \times 100$</th>
<th>$R^2$</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erlton</td>
<td>inbound</td>
<td>24.63***</td>
<td>0.94***</td>
<td>0.85***</td>
<td>0.05</td>
<td>0.46</td>
<td>12.69%</td>
</tr>
<tr>
<td>Chinook</td>
<td></td>
<td>28.31***</td>
<td>0.34***</td>
<td>0.16***</td>
<td>0.33***</td>
<td>0.20</td>
<td>13.64%</td>
</tr>
<tr>
<td>University</td>
<td></td>
<td>23.93***</td>
<td>0.59***</td>
<td>0.42***</td>
<td>0.46***</td>
<td>0.25</td>
<td>10.59%</td>
</tr>
<tr>
<td>Sunnyside</td>
<td></td>
<td>25.54***</td>
<td>0.45***</td>
<td>0.44***</td>
<td>0.55***</td>
<td>0.15</td>
<td>11.81%</td>
</tr>
<tr>
<td>39th Avenue</td>
<td></td>
<td>25.68***</td>
<td>0.53***</td>
<td>0.48***</td>
<td>0.50***</td>
<td>0.14</td>
<td>12.64%</td>
</tr>
<tr>
<td>Lions Park</td>
<td></td>
<td>24.88***</td>
<td>0.66***</td>
<td>0.65***</td>
<td>0.48***</td>
<td>0.15</td>
<td>11.80%</td>
</tr>
<tr>
<td>Erlton</td>
<td>outbound</td>
<td>24.34***</td>
<td>0.98***</td>
<td>0.69***</td>
<td>0.13*</td>
<td>0.45</td>
<td>12.33%</td>
</tr>
<tr>
<td>Chinook</td>
<td></td>
<td>27.16***</td>
<td>0.36***</td>
<td>0.30***</td>
<td>0.82***</td>
<td>0.24</td>
<td>12.60%</td>
</tr>
<tr>
<td>University</td>
<td></td>
<td>23.88***</td>
<td>0.52***</td>
<td>0.42***</td>
<td>0.81***</td>
<td>0.22</td>
<td>11.08%</td>
</tr>
<tr>
<td>Station</td>
<td>25.17***</td>
<td>0.48***</td>
<td>0.38***</td>
<td>0.80***</td>
<td>0.14</td>
<td>11.04%</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>39th Avenue</td>
<td>24.58***</td>
<td>0.67***</td>
<td>0.43***</td>
<td>1.12***</td>
<td>0.20</td>
<td>12.10%</td>
<td></td>
</tr>
<tr>
<td>Lions Park</td>
<td>25.22***</td>
<td>0.53***</td>
<td>0.38***</td>
<td>0.78***</td>
<td>0.10</td>
<td>11.12%</td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

**Table 4.2** presents the results for each selected station, categorized by direction as inbound and outbound. Notably, all estimated parameters were found to be statistically significant, except for the additional time added due to passenger load for boarding passengers at Erlton station. Erlton station, characterized by a narrower platform width compared to the other selected stations, exhibited the highest coefficients for average boarding and alighting time per passenger in both directions. Conversely, Chinook station, which shares a similar geometric layout but possess a wider platform, displayed the lowest coefficient values for average boarding and alighting time per passenger in both directions. It is also noteworthy to highlight the differences in station layout between Erlton and Chinook stations. Both stations feature a middle platform, but Erlton station has entrances positioned on both sides of the platform, while Chinook station has only one entrance situated on one side of the platform.

In the selected data, dwell time refers to the duration between the departure and arrival times of a train, excluding acceleration and deceleration times but including door opening and closing times. In this study, the intercept represents the combined duration of door opening and closing times, along with the duration a train waits for a green signal or scheduled departure time. Chinook station recorded the highest value for the intercept in both directions. This trend may be attributed to passengers having to walk longer distances to access the train doors due to the single entrance located at one end of the platform. Conversely, University station, with two entrances situated in the middle of the platform, demonstrated the lowest intercept values in both directions. This arrangement likely facilitates more efficient passenger movement, resulting in lower intercept
values in both directions. In the analysis, it was observed that the intercept time for both directions at all selected stations averages around 25 seconds. However, when applying Grubbs' outlier test, it was found that the intercept for Chinook station deviates significantly from this average. This finding suggests that proposing a uniform intercept value for all stations may not be appropriate. Moreover, this test underscores the argument that the intercept time is heavily influenced by the specific geometry of each station.

In summary, the analysis of station-specific results sheds light on the impact of platform geometry and entrance configurations on dwell time dynamics. Erlton's narrower platform and dual entrances contribute to longer boarding and alighting times, while Chinook's wider platform and single entrance lead to lower values for average boarding and alighting time per passenger. University station's centrally located entrances further optimize passenger flow, resulting in the least intercept among the selected stations. The findings underscore the importance of considering the uniqueness of station design in dwell time modeling. We observed that the same dwell time model should not be applied to all stations in the system. Despite the uniform characteristics of the trains, the average boarding and alighting time per passenger is influenced by the specific design of each station. This results in variations in total dwell time, even at the same level of passenger demand.

4.4 Identifying Outliers in the Data

In the previous section, lower R² values were observed for some models, which prompted an investigation into the specific reasons behind these lower values. However, it is important to mention that previous literature on dwell time has frequently reported lower R² values for regression models (Fritz, 1983; Gysin, 2018; Li et al., 2018; Palmqvist et al., 2020; Wirasinghe &
Szplett, 1984). In this section, we tested for nonlinear and polynomial fits for the data. Subsequently, the residual distribution was examined for uniformity, and outliers were investigated to determine if they were associated with specific events.

### 4.4.1 Test for Nonlinearity

In this subsection, we aim to investigate how the number of boarding and alighting passengers at the critical door relates to the dwell time. To analyze this relationship, we have employed various types of curve fits, including linear, nonlinear, and polynomial fits. We then compared the MSE for each direction for all selected stations. **Figure 4.5(a)** and **Figure 4.5(b)** illustrate the results for linear, nonlinear, and polynomial fits for Erlton station in the inbound and outbound directions, respectively.

**Figure 4.5**: MSE comparison for linear, and polynomial data fits (a) inbound and (b) outbound directions for Erlton station.

**Figure 4.5** presents graphs with the x-axis representing the total number of passengers boarding and alighting at the critical door, referred to as passenger demand, and the y-axis
represents the dwell time in seconds. In this graph, we have three different types of data fits: linear (in red), nonlinear (in green), and polynomial (in blue). The polynomial curve is of degree 3, allowing for greater flexibility in fitting the data but with the risk of overfitting. Interestingly, the nonlinear and polynomial trends were identical in both cases. We compared the MSE for linear, nonlinear, and polynomial curve fits to determine the most suitable curve fit without overfitting. It was observed that MSE decreases by a maximum of 0.66 seconds for polynomial curve fit compared to linear curve fit. This analysis shows that linear curve fit is the best trendline for Erlton station without overfitting. To compare the same trend for other selected stations, Table 4.3 is prepared, which compares MSE for Erlton, Chinook, University, Sunnyside, 39th Avenue, and Lions Park stations in each direction.

Table 4.3: MSE for Linear, and Polynomial Data Fits for Inbound and Outbound Directions.

<table>
<thead>
<tr>
<th>Station</th>
<th>Inbound</th>
<th></th>
<th>Outbound</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Linear</td>
<td>Nonlinear</td>
<td>Polynomial</td>
<td>Linear</td>
</tr>
<tr>
<td>Erlton</td>
<td>22.09</td>
<td>21.44</td>
<td>21.43</td>
<td>20.50</td>
</tr>
<tr>
<td>Chinook</td>
<td>30.93</td>
<td>30.75</td>
<td>30.75</td>
<td>26.74</td>
</tr>
<tr>
<td>University</td>
<td>14.92</td>
<td>14.92</td>
<td>14.89</td>
<td>13.88</td>
</tr>
<tr>
<td>Lions Park</td>
<td>17.40</td>
<td>17.40</td>
<td>17.38</td>
<td>18.92</td>
</tr>
</tbody>
</table>

It can be observed that the MSE values for all selected stations only change by fractions of seconds for linear and nonlinear data fits, except for the case of Erlton station in the inbound direction. A similar observation is found when comparing MSE values for nonlinear and
polynomial data fits in all cases. Although the MSE value varies by station, it does not change much across various data fitting methods, indicating that a linear relation between passenger demand and dwell time holds true. The variation in MSE values across stations can be attributed to outliers in the data, specifically observations with very low passenger demand (less than 5 boarding and alighting passengers) but higher dwell times.

One important factor to consider is Calgary Transit's policy of adhering to scheduled departure times. Moreover, trains need to be held at a station for a longer time than the scheduled dwell time due to train traffic congestion, especially for stations close to downtown where both the red and blue lines share the same track.

In conclusion, the analysis suggests that linear curve fitting is generally the most suitable choice for modeling dwell time at the mentioned stations, as it provides reasonably consistent results across different scenarios without overfitting. However, high MSE values in some cases can be attributed to outliers and the impact of operational policies, such as holding trains for scheduled departures and potential train congestion in certain downtown stations. Further investigation to identify outliers in the data is necessary to enhance the accuracy of the dwell time model.

4.4.2 Residuals’ Distribution

The linear regression model relies on several key assumptions. First, it assumes that the residuals in the model predictions are normally distributed with a mean of zero, meaning that the model should not systematically overpredict or underpredict. Second, there is the assumption of homoscedasticity, which means that the residuals should have equal variance. Third, the independence of the residual is assumed, indicating that the residual in one observation should not
be related to the residual in another observation and that there is no autocorrelation among the residual terms. To validate the normality assumption, we explored the distribution of residuals by employing various techniques like box plots, normal probability plots, and variable plots for predicted dwell time and independent variables. **Figure 4.6(a) and Figure 4.6(b)** present box plots for the residuals related to the predicted dwell time for each selected station. Our initial observation revealed that the residuals associated with predicted dwell time exhibited an approximately normal distribution centered around zero. However, it is crucial to note that across all scenarios, a significant number of outliers were identified.

![Figure 4.6: Residuals plots for Erlton, Chinook, University, Sunnyside, 39th Avenue, and Lions Park stations for (a) inbound and (b) outbound directions.](image)

As previously mentioned, Calgary Transit has a policy of adhering to scheduled departure time, leading to an increase in observed dwell time, even when passenger demand remains the same at the critical door. In our previous exercise, we found that residuals are normally distributed with long tail. Furthermore, we investigated whether there is a systematic pattern among the
outliers, which could potentially allow us to exclude them from our analysis. To achieve this, we explored the data distribution concerning the date of the month, the day of the week, the hour of the day, and the onboard passenger demand. This exploration aimed to uncover any relationships between passenger demand at the critical door and dwell time, assisting us in identifying outliers in our analysis.

4.4.3 Date of the Month

As mentioned before, in this thesis, we have obtained LRV data recorded by Calgary Transit. The data covers the entire months of July 2022 and July 2023, as well as 15 days each from January, April, and September 2023. In this comparison, we wanted to find a specific week or a day that has more outliers compared to other days in the month. As we know, summer is the most preferred time for regular maintenance of the rail network as well as road and signal repairs. Rail maintenance or road and signal repair near the LRT station may lead to holding trains at a station for a longer time than scheduled for the safety of workers.
Figure 4.7: Dwell time comparison with passenger demand for the day of July 2022 and 2023 for inbound direction of Erlton station.

Figure 4.7 represents passenger demand at the critical door on the x-axis and dwell time in seconds on the y-axis for inbound direction for Erlton station for the month of July 2022 and 2023. The color bar in Figure 4.7 shows the date of the month of July for the recorded data. In these plots, we focused on the data points located in the top-left side. These points represent instances where dwell time is higher despite lower passenger demand, which can be identified as outliers in the dataset. Our observations reveal that, during both years, there were instances, where passenger demand exceeded the usual levels. Notably, this occurred during the period of the Calgary Stampede, which typically takes place from July 7 to July 16 and is renowned as one of the busiest events in the city. However, aside from this period, there were not any distinct patterns or specific dates in a given week that significantly contributed to the presence of outliers in the data. It should be noted that similar exercise was performed for the month of January, April, and September for the Erlton station as well as other selected stations. However, it is not possible to present all plots in this thesis.

4.4.4 Day of the Week

In our analysis, we initially examined the data for outliers related to specific dates. However, we found that there was not any particular date or week that consistently stood out as a source of outliers in our recorded data. Given this, we extended our investigation to consider the influence of the day of the week. We were interested in understanding whether specific days, such as weekends or particular weekdays, might have had a significant impact on the presence of outliers in our data. For example, we wanted to determine if, say, Saturdays or Mondays were associated with unusually long dwell times with the same level of demand. This involved
examining the data closely to check if there were any patterns or trends that were specific to certain days of the week. By doing this, we aimed to gain a more comprehensive understanding of the factors contributing to the outliers in our recorded information.

Figure 4.8: Dwell time comparison with passenger demand for day of the week for inbound direction of Erlton station.

Figure 4.8 depicts passenger demand on the x-axis and dwell time in seconds on the y-axis for the inbound direction for Erlton station. Each subplot is divided to distinguish passenger flows (as mentioned in Table 3.1), and the colors indicate different days of the week. The data for each day of the week appears to be randomly distributed, and there is not a specific day that is significantly responsible for the outliers. Upon careful examination of these plots, we can conclude that the data for weekends and weekdays are remarkably similar, and at this point, there is insufficient evidence to exclude the weekend data from our analysis. Similar to the previous section, this analysis is conducted for all selected stations for each direction, but the results are not presented here.

4.4.5 Time of the Day

When we could not gather sufficient evidence from examining outliers that might be associated with particular days of the week or a specific date, we turned our attention to analyzing
the data distribution throughout different times of the day. Our goal was to understand whether the presence of peak or off-peak periods had any influence on these outlier occurrences. To explore this, we generated plots that display passenger demand, representing the total number of passengers boarding and alighting at the crucial door, as well as the dwell time. We utilized a unique color-coding scheme to distinguish various times of the day, helping us identify potential patterns or irregularities associated with different periods.

![Figure 4.9: Dwell time comparison with passenger demand for Time of the day for inbound direction of Erlton station.](image)

Figure 4.9, presents passenger demand on the x-axis, and dwell time in seconds on the y-axis, pertaining to the inbound direction at Erlton station. Each of these subplots is divided by passenger flow, and the use of colors serves to indicate the various times throughout the day. It is worth noting that during the Stampede period, the LRT system in Calgary operates 24 hours a day, a unique occurrence compared to other days, when there is a halt in service between 1:00 AM and 4:00 AM. This temporal distinction becomes evident in the recorded data, where we observe a reduced number of data points during this specific timeframe compared to other periods of the day. The Stampede event, lasting for a total of 10 days, contributes 20 days' worth of data, while the remaining days account for the standard 85-day dataset. An interesting observation emerges as we
scrutinize this data: there is no visible pattern that distinguishes peak and off-peak time periods throughout the day. As a result, we can reasonably conclude that excluding data from off-peak periods may not be a wise choice for our analysis. It is crucial to understand that most studies on dwell time collect data during peak hours, particularly when they employ manual data collection methods. This is due to the higher frequency during peak period. However, in our case, the utilization of the APC and AVL systems allows us to effortlessly collect data throughout the entire operational duration without the need for additional, labor-intensive efforts, setting this dataset apart as a valuable resource for analysis.

4.4.6 Passenger Load

Since we were unable to detect substantial evidence indicating the presence of outliers belonging to specific time periods, such as particular weeks, days, or times of the day, we undertook a comprehensive examination of passenger load. It is conceivable that under specific passenger load conditions, dwell time exhibits higher values, particularly when compared to instances of lower passenger load despite the same passenger demand levels. Our investigation sought to shed light on the possibility that, given specific passenger load conditions, there could be a higher value of dwell time, potentially deviating from the trend observed in cases with lower passenger loads but with equivalent passenger demand levels at the critical door.
Figure 4.10: Dwell time comparison with passenger demand for passenger load inbound direction of Erlton station.

Figure 4.10 represents the relationship between passenger demand on the x-axis, and dwell time in seconds on the y-axis. These plots belong to the inbound direction of the Erlton station. Each subplot is related to passenger flow, employing a unique color code to denote passenger load categories. In our comprehensive descriptive analysis, we ascertained that the median passenger load for the Erlton station was approximately 23 passengers per car for inbound direction. Notably, it is imperative to underscore that the passenger load referred to in Figure 4.10 pertains specifically to the car having the critical door. In investigating the association between passenger load and outliers, our focus is specifically on the data points located in the top left corner of the plots. These data points exhibit lower passenger demand but higher dwell time. The objective of this investigation is to identify the specific reasons behind these points. Despite the variation in passenger load, a noticeable pattern indicative of high passenger load's direct association with outliers within the data is not evident from this analysis. Consequently, we may reasonably conclude that the outliers are not primarily linked to instances of exceptionally high passenger load.
In this detailed analysis, we systematically explored the distribution of residuals in our linear regression model, scrutinizing key assumptions regarding normality, homoscedasticity, and independence. We employed various visualization techniques, including box plots and normal probability plots, to assess the normality of residuals. While our findings revealed an approximate normal distribution of residuals with some outliers. Furthermore, our investigation extended to the identification of causes for outliers in the data, where we carefully examined the date of the month, day of the week, hour of the day, and passenger load. Although we observed increased passenger demand during the Calgary Stampede, no specific day or week stood out as a consistent source of outliers. Additionally, we explored time-of-day patterns and passenger load, concluding that neither peak nor off-peak times nor exceptionally high passenger loads appeared to be primary contributors to the outliers.

In summary, this comprehensive analysis suggests that outliers in dwell time data are not significantly associated with specific time-related or passenger load factors; hence, there is no solid reason to exclude some data points from the analysis. This analysis also underscores the importance of considering other potential sources of variation that might be related to the operation of LRT, such as headway and specific incidents like door blocking due to technical failure. It should be noted that all trains operating in Calgary are not equipped with APC and AVL system, which limit this study to investigate the impact of headway on outliers. Additionally, Calgary Transit does not follow a specific pattern when dispatching trains equipped with and without APC and AVL systems. Another potential reason for outliers in the data could be passengers with bicycles or loading trolleys, as well as those using accessible services such as wheelchairs. These passengers typically take longer time to board and alight, thereby contributing to longer dwell times. One limitation of the data used in this study is the lack of information regarding passengers'
characteristics. Therefore, we cannot conclusively determine whether outliers are potentially associated with specific passenger characteristics.

In this chapter, the results for station-specific geometric characteristics were compared without considering passenger flow. Additionally, an investigation was conducted into possible reasons for outliers in the selected data. However, no specific reason for the outliers could be concluded. Chapter 5 explicitly compares the impact of station-specific geometric characteristics and passenger flow on dwell time. For this comparison, the selected data is divided into three distinct passenger groups as mentioned in Table 3.1.
CHAPTER 5. IMPACT OF STATION DESIGN AND PASSENGER FLOW

This chapter focuses on investigating the impact of station design and passenger flow on dwell time. To analyze the impact of station design, the study examined six stations with distinct geometric layouts, which were previously identified in Chapter 4. These stations were chosen specifically for their unique designs, allowing for a comprehensive exploration of how differently layouts affect dwell time. The data for each station and direction was categorized into three groups based on the fraction of boarding passengers \( \psi_{ni} \) mentioned in Table 3.1. The regression analysis was conducted for each direction and each group as mentioned in Table 3.1 using the OLS method in the Statsmodels library in Python. Notably, this process involved carrying out 36 unique regression analyses, as there were multiple stations, directions, and groups to be considered. This approach allowed for a comprehensive evaluation of the factors influencing dwell time, enabling us to make insightful comparisons between various scenarios.

5.1 Alighting-Dominant Passenger Flow

To enhance the comparability of our findings, we segregated the results for each passenger flow and presented them in separate tables. Table 5.1 presents the regression results for alighting-dominant passenger flow \( (0 \leq \psi_{ni} \leq 0.32) \), comparing the inbound and outbound directions at several selected stations: Erlton, Chinook, University, Sunnyside, 39th Avenue, and Lions Park. Chinook station stands out with the highest intercept, 27.4 seconds for the inbound direction and 26.6 seconds for the outbound direction. Conversely, University station demonstrates the lowest intercept values among the selected stations, with 23.5 seconds for inbound and 23.6 seconds for
outbound directions. Erlton station, characterized by a narrower platform compared to others featuring a middle platform, exhibits the highest coefficients for boarding and alighting passengers in both directions for boarding and alighting dominant cases. Notably, we observe that boarding passengers take more time to board in alighting-dominant passenger flows. In the majority of cases, the coefficient for boarding passengers is not statistically significant. This trend persists when considering the extra time added due to passenger load.

Table 5.1 Results for Alighting-Dominant Passenger Flow.

<table>
<thead>
<tr>
<th>Station Name</th>
<th>Direction</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3 \times 100$</th>
<th>$R^2$</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erlton</td>
<td>inbound</td>
<td>24.31***</td>
<td>1.26***</td>
<td>0.86***</td>
<td>-0.38</td>
<td>0.47</td>
<td>11.92%</td>
</tr>
<tr>
<td>Chinook</td>
<td></td>
<td>27.41***</td>
<td>1.63</td>
<td>0.20</td>
<td>-0.99</td>
<td>0.05</td>
<td>15.79%</td>
</tr>
<tr>
<td>University</td>
<td></td>
<td>23.47***</td>
<td>0.17</td>
<td>0.45***</td>
<td>1.37</td>
<td>0.15</td>
<td>9.60%</td>
</tr>
<tr>
<td>Sunnyside</td>
<td></td>
<td>24.57***</td>
<td>0.41</td>
<td>0.87***</td>
<td>0.73</td>
<td>0.10</td>
<td>13.18%</td>
</tr>
<tr>
<td>39th Avenue</td>
<td></td>
<td>25.08***</td>
<td>-0.88</td>
<td>0.84***</td>
<td>1.32</td>
<td>0.04</td>
<td>13.81%</td>
</tr>
<tr>
<td>Lions Park</td>
<td></td>
<td>24.04***</td>
<td>1.17</td>
<td>0.54***</td>
<td>-0.40</td>
<td>0.09</td>
<td>12.20%</td>
</tr>
<tr>
<td>Erlton</td>
<td>outbound</td>
<td>23.80***</td>
<td>1.16**</td>
<td>0.70***</td>
<td>-0.04</td>
<td>0.23</td>
<td>11.34%</td>
</tr>
<tr>
<td>Chinook</td>
<td></td>
<td>26.64***</td>
<td>-0.22</td>
<td>0.36***</td>
<td>2.04***</td>
<td>0.23</td>
<td>12.75%</td>
</tr>
<tr>
<td>University</td>
<td></td>
<td>23.61***</td>
<td>0.19</td>
<td>0.46***</td>
<td>1.59*</td>
<td>0.17</td>
<td>10.88%</td>
</tr>
<tr>
<td>Sunnyside</td>
<td></td>
<td>24.49***</td>
<td>0.05</td>
<td>0.53***</td>
<td>1.48**</td>
<td>0.17</td>
<td>11.47%</td>
</tr>
<tr>
<td>39th Avenue</td>
<td></td>
<td>24.50***</td>
<td>0.53</td>
<td>0.38***</td>
<td>2.14***</td>
<td>0.17</td>
<td>12.22%</td>
</tr>
<tr>
<td>Lions Park</td>
<td></td>
<td>24.72***</td>
<td>0.75*</td>
<td>0.45***</td>
<td>0.78</td>
<td>0.10</td>
<td>11.22%</td>
</tr>
</tbody>
</table>

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
5.2 Mixed Passenger Flow

Table 5.2 shows the results for the mixed passenger flow \((0.33 \leq \psi \leq 0.66)\) in inbound and outbound directions. In this analysis, we also found that Chinook station has the highest intercept for both directions with higher statistical significance level. Boarding and alighting time per passenger are almost the same for all stations in the mixed passenger flow, except for Erlton station. This disparity may be due to the platform width. In the mixed passenger flow case, the extra time added for boarding passengers due to the passenger load is statistically significant for many cases. University station stands out with the highest value for extra time due to passenger load for both directions. However, Chinook station exhibited the lowest value for extra time due to passenger load in both directions.

Table 5.2 Results for Mixed Passenger Flow.

<table>
<thead>
<tr>
<th>Station Name</th>
<th>Direction</th>
<th>(\beta_0)</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
<th>(\beta_3 \times 100)</th>
<th>(R^2)</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erlton</td>
<td>inbound</td>
<td>24.83***</td>
<td>1.31***</td>
<td>0.60***</td>
<td>-0.08</td>
<td>0.58</td>
<td>12.50%</td>
</tr>
<tr>
<td>Chinook</td>
<td></td>
<td>28.02***</td>
<td>0.23</td>
<td>0.25*</td>
<td>0.54***</td>
<td>0.16</td>
<td>12.67%</td>
</tr>
<tr>
<td>University</td>
<td></td>
<td>24.16***</td>
<td>0.14</td>
<td>0.49*</td>
<td>1.39***</td>
<td>0.20</td>
<td>10.63%</td>
</tr>
<tr>
<td>Sunnyside</td>
<td></td>
<td>25.63***</td>
<td>0.24</td>
<td>0.31</td>
<td>1.13***</td>
<td>0.15</td>
<td>11.11%</td>
</tr>
<tr>
<td>39th Avenue</td>
<td></td>
<td>24.51***</td>
<td>0.88*</td>
<td>0.83*</td>
<td>0.16</td>
<td>0.14</td>
<td>12.77%</td>
</tr>
<tr>
<td>Lions Park</td>
<td></td>
<td>25.34***</td>
<td>0.01</td>
<td>0.88**</td>
<td>0.90</td>
<td>0.11</td>
<td>11.03%</td>
</tr>
<tr>
<td>Erlton</td>
<td>outbound</td>
<td>24.39***</td>
<td>0.86***</td>
<td>1.02***</td>
<td>-0.04</td>
<td>0.49</td>
<td>12.61%</td>
</tr>
<tr>
<td>Chinook</td>
<td></td>
<td>27.19***</td>
<td>0.36***</td>
<td>0.31***</td>
<td>0.75***</td>
<td>0.32</td>
<td>11.80%</td>
</tr>
<tr>
<td>University</td>
<td></td>
<td>24.22***</td>
<td>-0.01</td>
<td>0.56***</td>
<td>1.63***</td>
<td>0.19</td>
<td>11.24%</td>
</tr>
</tbody>
</table>
5.3 Boarding-Dominant Passenger Flow

Table 5.3 presents the regression results for the case of boarding-dominant passenger flow \((0.67 \leq \psi_{n1} \leq 1)\) for inbound and outbound directions. In this scenario, the observation regarding the intercept remains consistent: Chinook station exhibits the highest intercept, while University stands out with the lowest intercept for both directions, at a statistically significant level. Similar to the findings in the alighting-dominant passenger flow, the coefficients for alighting passengers are notably high in the boarding-dominant scenario. The significance level for boarding passengers is consistently higher across all cases, including the extra time added due to passenger load. Conversely, the significance level for alighting passengers' coefficients is not notably high in the context of the boarding-dominant passenger flow as expected.

Table 5.3 Results for Boarding-Dominant Passenger Flow.

<table>
<thead>
<tr>
<th>Station Name</th>
<th>Direction</th>
<th>(\beta_0)</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
<th>(\beta_3 \times 100)</th>
<th>(R^2)</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erlton</td>
<td>inbound</td>
<td>24.96***</td>
<td>0.85***</td>
<td>1.24***</td>
<td>0.07</td>
<td>0.37</td>
<td>13.09%</td>
</tr>
<tr>
<td>Chinook</td>
<td>inbound</td>
<td>28.66***</td>
<td>0.33***</td>
<td>0.17</td>
<td>0.34***</td>
<td>0.21</td>
<td>13.67%</td>
</tr>
<tr>
<td>University</td>
<td></td>
<td>24.15***</td>
<td>0.57***</td>
<td>0.34</td>
<td>0.39***</td>
<td>0.24</td>
<td>10.68%</td>
</tr>
<tr>
<td>Sunnyside</td>
<td></td>
<td>25.93***</td>
<td>0.36***</td>
<td>0.26</td>
<td>0.59***</td>
<td>0.14</td>
<td>11.58%</td>
</tr>
</tbody>
</table>
Presented analysis reveals that boarding takes more time in alighting-dominant passenger flow, whereas alighting takes longer time per passenger in boarding-dominant passenger flow. The additional time due to passenger load is significant in most mixed passenger flow cases and consistently significant in boarding-dominant passenger flow, with the exception of Erlton station in the inbound direction. However, in the case of alighting-dominant passenger flow, the term representing extra time due to passenger load lacks statistical significance. This is no surprise as we expect alighting passengers to get closer to the train doors prior to door opening times. As such, when alighting in dominant, higher train load does not affect dwell time as much as the case when boarding is dominate or with mixed flow. The term for extra time added due to passenger load can be omitted from the model for stations with a low significance level. Nevertheless, we have included it to facilitate uniform comparisons across all stations using the same model. We divided the data into multiple groups, such as passenger flows and directions, leading to instances where certain terms were statistically insignificant. It is crucial to acknowledge that our primary objective in this chapter was not to construct a concise dwell time model but rather to compare various
scenarios, extracting insights from each unique case. Analysis conducted in Chapter 4, without considering passenger flow distinctions, revealed that all independent variables were statistically significant using the same data and dwell time model. This implies that the decision to segregate data into multiple groups hinges on the specific objectives of study. In this chapter, we aimed to explore variations in the average passenger boarding and alighting times per passenger, particularly in the context of dominant boarding and alighting passenger flows.

After conducting a regression analysis and excluding coefficients with negative values from the model, we found that there was no notable difference in the results. Table 5.4, Table 5.5, and Table 5.6 present the results after removing negative coefficients for the stations where such values were present. Upon close examination of these results and comparing them with models that reported negative coefficients, we noticed that the $R^2$ value and MAPE are changing by less than 0.01 and 0.45%, respectively. Although there were slight alterations in the values for intercept and other coefficients, these differences were not statistically significant.

Table 5.4 Results for Alighting-Dominant Passenger Flow.

<table>
<thead>
<tr>
<th>Station Name</th>
<th>Direction</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3 \times 100$</th>
<th>$R^2$</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erlton</td>
<td>inbound</td>
<td>24.31***</td>
<td>0.96***</td>
<td>0.86***</td>
<td>NA</td>
<td>0.47</td>
<td>11.93%</td>
</tr>
<tr>
<td>Chinook</td>
<td>inbound</td>
<td>27.43***</td>
<td>1.13</td>
<td>0.21</td>
<td>NA</td>
<td>0.04</td>
<td>15.79%</td>
</tr>
<tr>
<td>39th Avenue</td>
<td>inbound</td>
<td>25.09***</td>
<td>NA</td>
<td>0.81***</td>
<td>0.08</td>
<td>0.04</td>
<td>13.82%</td>
</tr>
<tr>
<td>Lions Park</td>
<td>inbound</td>
<td>24.04***</td>
<td>1.17</td>
<td>1.20***</td>
<td>NA</td>
<td>0.09</td>
<td>12.20%</td>
</tr>
<tr>
<td>Erlton</td>
<td>outbound</td>
<td>23.80***</td>
<td>1.14***</td>
<td>0.70***</td>
<td>NA</td>
<td>0.23</td>
<td>11.34%</td>
</tr>
<tr>
<td>Chinook</td>
<td>outbound</td>
<td>26.63***</td>
<td>NA</td>
<td>0.34***</td>
<td>1.83***</td>
<td>0.22</td>
<td>12.77%</td>
</tr>
</tbody>
</table>

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Table 5.5 Results for Mixed Passenger Flow.

<table>
<thead>
<tr>
<th>Station Name</th>
<th>Direction</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3 \times 100$</th>
<th>$R^2$</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erlton</td>
<td>inbound</td>
<td>24.89***</td>
<td>1.24***</td>
<td>0.60***</td>
<td>NA</td>
<td>0.58</td>
<td>12.49%</td>
</tr>
<tr>
<td>Erlton</td>
<td>outbound</td>
<td>24.40***</td>
<td>0.83***</td>
<td>1.03***</td>
<td>NA</td>
<td>0.49</td>
<td>12.16%</td>
</tr>
<tr>
<td>University</td>
<td>outbound</td>
<td>24.22***</td>
<td>NA</td>
<td>0.56***</td>
<td>1.62***</td>
<td>0.19</td>
<td>11.24%</td>
</tr>
<tr>
<td>Lions Park</td>
<td></td>
<td>25.73***</td>
<td>NA</td>
<td>0.58***</td>
<td>0.96***</td>
<td>0.15</td>
<td>9.99%</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

Table 5.6 Results for Boarding-Dominant Passenger Flow.

<table>
<thead>
<tr>
<th>Station Name</th>
<th>Direction</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3 \times 100$</th>
<th>$R^2$</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sunnyside</td>
<td>outbound</td>
<td>25.38***</td>
<td>0.45***</td>
<td>NA</td>
<td>0.82***</td>
<td>0.11</td>
<td>11.25%</td>
</tr>
<tr>
<td>39th Avenue</td>
<td></td>
<td>24.56***</td>
<td>0.68***</td>
<td>NA</td>
<td>1.17***</td>
<td>0.19</td>
<td>11.88%</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

The goodness of fit values vary by station and direction, and fraction of boarding passengers. It should be noted that we included both peak and off-peak periods for the model estimation. The goodness of fit could be improved by separating flows by peak/off-peak periods. However, this may result in insufficient data size for the comparisons of different passenger flows and station designs. It was observed that this data had many outliers, which resulted in lower $R^2$ values for some models. A comprehensive study conducted in Chapter 4 found that outliers in dwell time data are not associated with specific time-related or passenger load factors; hence, there is no solid reason to exclude outliers from the analysis. Also, literature on dwell time has reported
lower R² values for regression models (Fritz, 1983; Gysin, 2018; Li et al., 2018; Palmqvist et al., 2020; Wirasinghe & Szplett, 1984).

5.4 Discussion

In order to ensure an unbiased and meaningful comparison of station-specific geometric characteristics, we have compared the estimated dwell time with the same level of demand for all stations and are presented in Figure 5.1(a) and Figure 5.1(b) for inbound and outbound direction, respectively. In both figures, we assumed a passenger load of 50 passengers for the car with the critical door. We then increased the level of alighting demand along the x-axis and the level of boarding demand along the y-axis, ranging from 5 passengers to 15 for the critical door. The dwell time in seconds is represented on the z-axis, with different colors (blue, green, red, cyan, magenta, and yellow) corresponding to Erlton, Chinook, University, Sunnyside, 39th Avenue, and Lions Park, respectively. For each station, we have developed three distinct dwell time models Equation 3-3 based on the passenger groups outlined in Table 3.1. In the displayed plot, we have employed one of these models to estimate the dwell time based on the fraction of boarding passengers (ψₙ).
Figure 5.1: Dwell time comparison for boarding, and alighting passengers for (a) inbound and (b) outbound direction.

Upon careful observation, we have noted some interesting findings regarding the dwell times at different stations. University station (Figure 4.2(c)) stands out with the lowest dwell times in various scenarios. This station has a middle platform with two entrances located at the middle of the platform, as well as a heated waiting room positioned at the middle. These observations suggest that a station design with a middle platform and multiple entrances positioned at the middle of the platform outperformed other stations in terms of dwell time in our case study for cases of alighting-dominant and mixed passenger flows. Another notable observation of the boarding-dominant group is that Sunnyside station exhibits shorter dwell times for both directions, particularly under 30 seconds for the outbound direction. However, in most cases, Sunnyside station stands out with a second-lowest dwell time than the other considered stations. Sunnyside station, presented in Figure 4.2(d), has side platforms with multiple entrances at the end and middle of the platform.

Erlton station, presented in Figure 4.2(a), has a middle platform but with a narrower width as compared to other selected stations (around 7.5 m vs around 9-10 m). This station has maximum dwell time in all cases except the case of mixed passenger flow with a lower level of passenger demand in the inbound direction. Our case study revealed that a station with a middle narrow platform and two entrances positioned at the end underperformed other station layouts. The 39th Avenue station, depicted in Figure 4.2(e), is designed with side platforms and features entrances at the end and middle. This station consistently exhibits the second-highest dwell time in most cases. Both 39th Avenue and Lions Park stations have identical platform widths, which are narrower by 0.5 meters compared to the Sunnyside station. It is important to acknowledge that the
difference in dwell time for the same level of passenger demand at 39th Avenue and Lions Park stations is influenced by the location of entrances and the platform area covered by the heated shelter. Lions Park Station has multiple entrances at both ends and in the middle of the platform for the inbound direction and two entrances at both ends along with a sidewalk at a lower grade for the outbound direction.

In the comparison between alighting and boarding dominant passenger flows, it was found that the average alighting time per passenger is higher in the boarding-dominant flow, while the average boarding time per passenger is higher in the alighting-dominant flow. This observation is consistent with findings reported by Wirasinghe & Szplett, (1984) regarding the comparative analysis of average alighting and boarding time per passenger within alighting and boarding dominant passenger flows. This study revealed significant insights into dwell times at different stations. The university station consistently stood out with the lowest dwell times, attributed to its middle platform, and strategically-placed entrances. Conversely, Sunnyside station exhibited notably short dwell times for the case of boarding-dominant passenger flow, particularly under 30 seconds for the outbound direction, yet consistently ranked second-lowest overall. Erlton station, characterized by a narrower middle platform, recorded maximum dwell times in most scenarios. Lastly, 39th Avenue station consistently maintained the second-highest dwell times, highlighting the impact of platform layout on passenger flow dynamics.

This chapter compared the impact of station-specific geometric characteristics and passenger flow on dwell time. Chapter 6 presents the results for the microsimulation model for dwell time. This includes the results for model calibration and then compares the various dwelling strategies to minimize dwell time.
CHAPTER 6. MICROSIMULATION MODEL

In this chapter, a microsimulation model for dwell time using Vissim is developed for the Erlton station, which is located close to the CBD on the red line of Calgary Transit. The developed microsimulation model is then calibrated to enhance the performance of the simulation model by considering passenger flow modeling. Later, the microsimulation model is used to test alternative dwelling strategies.

6.1 Simulation Model for Erlton Station

To create an accurate simulation model, it is crucial to identify the geometric features of the station. Figure 4.2(a) illustrates the platform design for Erlton station, featuring a middle platform with two entrances at both ends. Additionally, there are two shelters and two information boards on the platform, which may obstruct the free movement of passengers. Relevant information was collected through field visits and data collected by the APC and AVL systems.

Figure 6.1 depicts the layout of the developed simulation model in Vissim. The model was developed by first creating two links and Public Transit (PT) lines with a headway of 150 seconds. PT stops were then added to the model. The train's occupancy was adjusted to assess the impact of passenger load on dwell time. Subsequently, the boarding demand per hour was incorporated. Additionally, the alighting demand per vehicle was calculated as a percentage based on occupancy and assigned to the alighting passenger’s percentage parameter in Vissim. These adjustments were based on input variables derived from collected data, including boarding passengers, alighting passengers, and passenger load. The final step in modeling was to assign pedestrian route choices.
Since the station has two entrances at each end of the platform, alighting and boarding passengers were divided between both entrances.

![Platform layout of the simulation model in Vissim](image)

**Figure 6.1: Platform layout of the simulation model in Vissim.**

The parameter values utilized within the Vissim simulation software are established through field observations. The parameter values used in Vissim, including a door lock duration before departure of 8 seconds, a door closure delay of 2 seconds, a clearance time of 5 seconds, and a door lock duration of 2 seconds, which are determined based on field observations. After developing dwell time models and setting the required parameters, passenger flow is modeled to enhance the simulation results and calibrate the model. Average boarding and alighting time per passenger are obtained from the **Table 4.2**. In this research, to calculate the actual speed \(S_1\), jam density and free flow speed described by Fruin, (1971) are considered which are 4 pedestrians per square meter and 1.425 meters per second, respectively. Then, **Equation 3-6** is used to calculate the desired speed factor for walking speed. Subsequently, adjustments are made to the React to \(n\) pedestrians and Tau values to calibrate the model. The optimized values are determined to be 6 for React to \(n\) pedestrians and 0.2 for Tau.
The simulation model in Vissim runs for an hour, but results are recorded for only 30 minutes. The initial 15 minutes and final 15 minutes are designated as warm-up and cool-down periods, respectively, for the simulation. Each case in Vissim is simulated ten times, and the average dwell time is calculated for comparison. Figure 6.2 represents the results of the model calibration, with observed dwell time in seconds on the x-axis and simulated dwell time in seconds on the y-axis.

Figure 6.2: Results for simulation model calibration.

In the context of the simulation results presented in Figure 6.2, the red dotted line serves as a reference point representing the ideal scenario where the simulated dwell time perfectly matches the observed dwell time. The clustering of data points around this line suggests that, in
many cases, the simulated dwell time closely approximates the observed values. However, it is essential to consider the influence of Calgary Transit's operational policy, which prioritizes strict adherence to scheduled departure times. This policy means that the observed dwell time may increase due to the need to maintain the scheduled departure schedule. Typically, Calgary Transit schedules dwell times of around 30 seconds for each station. As evident from Figure 6.2, there are no data points representing observed dwell times less than 30 seconds, indicating the impact of this scheduling policy. However, it is important to note that this policy is not explicitly incorporated into the simulation environment. As a result, the simulation produced dwell times of less than 30 seconds in scenarios with lower passenger demand, where departures could potentially occur sooner than scheduled.

6.2 Alternatively Dwelling Strategies

In this section, we conducted a detailed analysis to explore three different dwelling strategies using a microsimulation approach. The first strategy, Policy 1, assumes that boarding passengers are uniformly distributed across the platform. This means that passengers waiting to board the train are spread out evenly along the length of platform. The second strategy, Policy 2, involves passengers using specific doors for boarding and alighting. This prevents passengers from using the same door for both purposes, ensuring smoother flow and reduced congestion. Policy 3 combines elements of both Policy 1 and Policy 2. It features dedicated doors for boarding and alighting, similar to Policy 2, while also maintaining a uniform distribution of boarding passengers on the platform.

To assess the effectiveness of these strategies, we examined a scenario with 50 onboard passengers for the car which has critical door. The results, depicted in Figure 6.3, provide insights
into how varying levels of alighting and boarding passengers impact dwell time. The figure's x-axe and y-axis represent alighting and boarding passengers, respectively, while the dwell time in seconds is shown on the z-axis. This analysis helps us understand the implications of different dwelling strategies for passenger flows and level of passenger demand.

Figure 6.3: Comparison of alternative dwelling strategies for 50 passenger load.

From Figure 6.3, we can observe that Policy 3 outperformed in scenarios with mixed passenger flow. On the other hand, Policy 1 demonstrates better outcomes in situations characterized by either a boarding-dominant or alighting-dominant passenger flow. This suggests that having dedicated doors for boarding and alighting is effective when coupled with a uniform distribution of boarding passengers on the platform for the mixed passenger flow. However, it may
not be optimal to separate boarding and alighting passengers in cases where one group significantly outweighs the other. Doing so could result in the overutilization of the dedicated boarding or alighting door.
CHAPTER 7. CONCLUSIONS AND RECOMMENDATIONS

7.1 Overview

Dwell time plays a vital role in the overall performance and capacity of rail systems, and making accurate estimation of dwell time is essential for operational planning, system simulation, and optimizing headways. Traditional data collection methods have limitations in terms of data volume and accuracy, prompting the need for employing automated data sources such as APC and ALV systems. The primary objective of this study was to demonstrate the use of APC and AVL data for predicting dwell time, accurately estimating passenger load per train car, comparing dwell time models based on station-specific characteristics, enhancing simulation performance by considering passenger flow modeling, and testing alternative dwelling strategies. By incorporating data from critical doors and considering the fraction of boarding passengers, this study aimed to provide valuable insights into passenger behavior and optimize rail operations. The literature on the distribution of waiting passengers highlighted that passenger distribution on the platform is influenced by the location and number of entrances. Assuming a uniform passenger distribution for dwell time estimation can lead to an underestimation of dwell time. Another literature gap in dwell time studies involves accurately estimating the passenger load for each train car. Moreover, literature is lacking in supporting methodologies to accurately developing simulation models for dwell time optimization. This study aims to fill this gap by using APC data to identify the critical door for dwell time and to estimate the precise passenger load for each train car. Later, we compare dwell time based on the specific geometric characteristics of the station and the flow of passengers. Additionally, this study proposes a methodology to improve simulation performance by
incorporating passenger flow and evaluates alternative dwelling strategies aimed at minimizing total dwell time across different passenger flow groups.

### 7.2 Summary of the Work and Results

The study analyzed APC and AVL data from Calgary Transit for the entire months of July 2022 and July 2023, as well as 15 days each from January, April, and September 2023, with a focus on six selected stations with different platform designs. Regression analyses revealed variations in dwell time across stations, directions, and passenger flows. In regression analyses, a notable finding was that the average alighting time per passenger was higher in the boarding-dominant group, while the average boarding time per passenger was higher in the alighting-dominant group. This means, when the majority of passengers are boarding, the average alighting time per passenger was higher. On the other hand, when the majority of passengers are alighting, the average boarding time per passenger was higher. This result indicates that boarding time and alighting time per passenger heavily depends on the passenger flow dominance. It is important to note that when the passenger flow is alighting-dominant, boarding passengers cannot share time with the alighting passengers and need their own time for boarding. Similar results are reported in existing literature. According to Wirasinghe & Szplett, (1984), the average alighting time per passenger was observed to be greater within the boarding-dominant group, whereas the average boarding time per passenger was found to be higher within the alighting-dominant group. However, their study did not consider the passenger load for the comparison of passenger flows. In contrast, our study considered the exact number passenger load and compared the passenger flows based on the station-specific geometric characteristics.
Based on our observations, we found that for some stations the coefficients for boarding passengers are not statistically significant in the case of an alighting-dominant passenger flow, and similarly, the coefficients for alighting passengers are not statistically significant in the case of a boarding-dominant passenger flow. Moreover, the significance level for the extra time added due to passenger load is low for alighting-dominant passenger flow for the case of Calgary’s Light Rail Transit system. For the sake of comparison, we have selected the same regression model for all stations but variables with low insignificance level should be removed from the model. Our study, without passenger flow distinctions, found all independent variables statistically significant, indicating the importance of data segregation based on study objectives, such as exploring variations in average passenger boarding and alighting times in the context of dominant flows.

The observations in our study indicated that a station with a middle platform and two entrances located at the middle (such as the University station) was more efficient in terms of minimizing dwell times for many instances except the case of a boarding-dominant passenger flow. Moreover, stations with side platforms and multiple entrances at ends and the middle (such as the Sunnyside station) outperformed others in handling boarding-dominant passenger flows in our case study. Stations with a narrower middle platform and entrances at both ends of the platform (such as the Erlton station) underperformed other designs. This station exhibited the highest intercept and coefficients among the stations considered in this study. In our analysis, we discovered that a wider platform with multiple entrances was more effective in minimizing the dwell time. Specifically, a side platform with three entrances performed better for boarding-dominant passenger flows, while a middle platform with two entrances in the middle excelled in case of alighting-dominant passenger flows.
This study also developed and calibrated a microsimulation model for dwell time using Vissim, specifically for the Erlton station. Geometric features of the station were identified through field visits, and the simulation model was developed accordingly. Results from the calibration indicated that the simulated dwell time closely approximated observed dwell time. The study also explored three alternative dwelling strategies: Policy 1, which considered uniform distribution of boarding passengers on platform; Policy 2 which involves dedicated door for boarding and alighting passengers; and Policy 3 which is a combination of Policy 1 and Policy 2. Policy 3, outperformed in case of mixed passenger flow, while Policy 1 demonstrated better outcomes in boarding-dominant or alighting-dominant situations.

While some of the observations and conclusions may not be generic for all systems and different stations, the proposed method and station categorization can help to answer the efficiency of station design in terms of dwell time minimization. This method provides valuable insights into the dynamics of pedestrian movement by utilizing existing data, without the need for additional simulations. Different mixes of passenger flows have different impacts on dwell time and a particular station design may be more appropriate given certain passenger flows, as such similar studies can be performed in different systems in order to design new or modify existing stations. The insights derived from this study hold a significant value for transit authorities, as they can aid in implementing improved crowd management strategies and optimizing resource allocations.

7.3 Contributions

The study contributes significantly to the field of urban rail transit. Firstly, it showcases the effective utilization of APC and AVL systems for detailed dwell time estimation and precise passenger load estimation per train car. Leveraging these automated data sources enables the study
to provide more accurate insights into the performance of urban rail transit systems. Secondly, the proposed method and station categorization presented in this study offer a significant contribution to the field by addressing the efficiency of station design with a focus on minimizing dwell time. By categorizing stations based on their specific characteristics and analyzing the impact of these characteristics on dwell time, the study provides valuable insights into the dynamics of pedestrian movement within rail transit systems. Through this approach, transit authorities and urban planners can make informed decisions to enhance station efficiency and improve the overall performance of rail transit systems.

Finally, the study investigates how to improve the performance of simulation models by integrating realistic passenger walking behavior, particularly in crowded platform environments. This innovative approach offers a more precise depiction of real-world situations, enhancing the accuracy of simulations. Moreover, the study investigates alternative dwelling strategies, including uniform passenger distribution on platforms and the implementation of dedicated boarding and alighting doors. By exploring these strategies, policymakers gain valuable insights into effective approaches for minimizing dwell time.

7.4 Recommendations

Stations tend to experience boarding-dominant passenger flows during morning peak hours and alighting-dominant flows during evening peak hours, and vice versa. In such scenarios, a side platform with more than two entrances and a wider platform is recommended for the boarding-dominant direction, while more than one entrance with a wider platform is advised for the alighting-dominant direction. For stations experiencing mixed passenger flows, it is recommended
to have a middle platform with more than one entrances and a width of more than 9.0 meters, as both morning and evening peak hours will involve mixed passenger flows.

Based on the microsimulation analysis in various passenger flow scenarios, it is recommended for transit authorities to consider a flexible approach to door allocation at station platforms. This approach involves implementing dedicated doors for both boarding and alighting passengers with uniform distribution of boarding passengers when faced with mixed passenger flows with heavy passenger load. However, in situations of boarding-dominant and alighting-dominant passengers flows, it is advisable to avoid separating boarding and alighting passengers to prevent the overutilization of dedicated doors. By adopting this flexible approach, transit systems can optimize dwell time and enhance passenger experience.

7.5 Future Research

Based on the limitations identified in this study, several promising directions for future research emerge. Future research could investigate into the development of more comprehensive data collection methodologies that capture passengers' entry and exit choices and their preferences for waiting areas on the platform. By employing advanced tracking technologies or performing image processing through CCTV cameras at stations, researchers can gain deeper insights into passenger behavior and distribution patterns. This enhanced understanding would enable more accurate modeling of passenger flows and leading to more effective strategies for optimizing rail operations and enhancing the passenger experience.

Additionally, there is a need for studies exploring the relationship between headways and dwell time, particularly in the context of longer headways. Investigating how variations in headways impact passenger boarding and alighting behavior, as well as overall dwell time, would
provide valuable insights for transit planners and operators. One limitation of this study is the lack of information about passengers' characteristics, such as age, gender, and accessibility requirements. This limitation can be addressed using the developed microsimulation model. One can identify the gender of passengers and those with accessibility requirements and estimate the average time required for them to board and alight.

While we utilized data from the LRV system, this approach can be extended and applied to any rail transit systems equipped with APC and AVL systems for collecting similar data. For future research, it is recommended to study the role of station design on passenger distribution along the platform and the magnitude of boarding and alighting at the critical door. Moreover, extending this approach to other urban rail systems could lead to enhanced dwell time modeling and contribute to more efficient rail operations.
REFERENCES

https://doi.org/10.1080/19439962.2010.487636


https://doi.org/10.1109/iwobi.2014.6913934


http://www.strc.ethz.ch/2015/Bosina_EtAl.pdf


Calgary Transit. (2023). *Calgary Transit’s Light Rail Transit (LRT) Map* [dataset].

https://www.calgarytransit.com/content/dam/transit/rider-information/CTrain-Map-June2023.pdf


https://doi.org/103747


https://doi.org/10.1243/09544097JRRT115

https://doi.org/10.1007/s40864-021-00161-8


Hayes, S., Charlton, J., Fletcher, D., & Richmond, P. (2022). Validation of Agent-Based Passenger Movement Modeling for Railway Stations Subject to Social Distancing During the COVID-19 Pandemic. Transportation Research Record.
https://doi.org/10.1177/03611981221093634


