The influence of the gap on platform train interface occurrences

Myoya, Rozina L.

A dissertation submitted in partial fulfilment of the requirements for the degree of

MASTER OF ENGINEERING (TRANSPORTATION ENGINEERING)

In the

FACULTY OF ENGINEERING, BUILT ENVIRONMENT AND INFORMATION TECHNOLOGY

UNIVERSITY OF PRETORIA

February 2020

Dissertation Summary

The influence of the gap on platform train interface occurrences

RL MYOYA

Supervisor:	Gräbe, Hannes P.J.
Co-Supervisor:	Professor C.J. Venter
Department:	Civil Engineering
University:	University of Pretoria
Degree:	Master of Engineering (Transportation engineering)

Investigation into the causes of Platform Train Interface (PTI) occurrences has led to the conclusions that PTI occurrences can be attributed to human factors and platform design. In South Africa, PTI occurrences account for the majority of operational related incidents at train stations, where it was predicted by the Railway Safety Regulator that for 95% of all occurrences at least one serious injury is recorded. The purpose of investigating the causes of PTI occurrences would be to determine the most effective interventions that would lead to the most significant decrease in occurrences, especially considering stations where the reduction of the gap size to a standard of 75 mm horizontal and ±50 mm vertical (as stipulated by the Railway Safety and Standards Board, UK) would not be feasible due to operational constraints. The research aimed to identify and quantify all the factors that influenced PTI occurrences through the application of statistical techniques such as Multilinear Regression and Multinomial Logistic regression analysis. Thereafter, the factors identified were ranked according to the level of significance. From the research it was concluded that the main factors influencing PTI occurrences at stations were overcrowding, train design and platform design. It was further concluded that the train design was the most significant factor and that the change to a safer train design would have resulted in the highest reduction in PTI occurrences. The research also measured the level of accessibility into the train by special needs passengers to comment on the level of service provision by the Metrorail to the most vulnerable in society. Findings from the research showed that special needs passengers were, indeed, excluded from the service because of the large gap at the PTI.

ACKNOWLEDGEMENTS

I wish to express my appreciation to the following organisations and persons who made this project dissertation possible:

a) The Railway Safety Regulator for sponsoring this research carried out by the Chair in Railway Safety at the University of Pretoria, and for the provision of historical data that was used in the research.

b) Professor P.J. Gräbe my supervisor, and Professor C.J. Venter, my co-supervisor for their continuous guidance and support.

c) The Marketing and Communication department at PRASA, Pretoria for providing access to station platforms along the Pretoria-Pienaarspoort corridor where the research was conducted.

d) The following persons are gratefully acknowledged for their assistance during the course of the study:

    i) Sadiq Wanjala for the insight into statistical analysis.

    ii) Andre Broekman for his assistance in the development of the Arduino prototype developed in the study.

    iii) Dr. C. Malan for her continuous support and willingness to assist in providing the necessary resources to conduct the research.

e) My family and friends for their unwavering encouragement and support during the study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

DoT	Department of Transport
FWI	Fatalities and Weighted Injuries
HDL	Human Dynamics Laboratory
LCG	Linear Congruent Method
ML	Multilinear
MNL	Multinomial Logistic
MSA	Moving South Africa
MT	Moving train
NJDOT	New Jersey Department of Transport
OC	Overcrowding

PAMELA	Pedestrian Accessibility and Movement Environment Laboratory
PRASA	Passenger Rail Agency of South Africa
PRM	Persons with Reduced Mobility
PTI	Platform Train Interface
RRL	Road Research Laboratory
RSR	Railway Safety Regulator
RSSB	Railway Safety and Standards Board
SR	Staff riding (train surfing)
TU	Train units
WESMLE	Weighted Exogenous Sample Maximum Likelihood Estimator

1 INTRODUCTION

1.1 BACKGROUND

International research has indicated that urbanisation will continue to increase at an alarming rate in the future. The current pattern seen in South Africa indicates an increase of 1 million people in Gauteng since the 2011 census (Statistics South Africa, 2013). This increase was due to the migration of people from the Eastern Cape, Northern Cape, Free State, KwaZulu Natal, Limpopo as well as immigrants from Southern Africa. This drastic increase in the urban population could have a crippling effect on the transport system if it is not upgraded to accommodate this growing volume of commuters.

Metrorail, the commuter rail service arm of the Passenger Rail Agency of South Africa (PRASA), is an example of a commuter metro rail system that has the potential to serve as the backbone of the Gauteng transport system to address the growing volume of commuters. Metrorail is responsible for transporting approximately 2 million commuters per day nationwide (Metrorail, 2018). However, the Metrorail service is plagued with accounts of maladministration and a substandard mode of operation that impacts passenger safety and train ridership (Mathebula and Sopazi, 2016). This current state of operation can be traced back to the trend that became prevalent in the railway industry in South Africa after the publication of the De Villiers Report (1986). The trend observed after the publication of that report was the minimal resource and political support in promoting rail as the preferred mode of transport.

This lack of political support has led to the extreme degradation of the Metrorail railway network, where the current mode of operation leads to unsafe conditions for train commuters and personnel. During the 2017/2018 reporting period, the Railway Safety Regulator (RSR) reported 4 478 operational occurrences. Occurrences at the platform train interface (PTI) accounted for approximately 16.6 % (744 occurrences) of the total recorded operational occurrences. The precursors of these incidents were attributed to human behaviour and platform design.

One of the main aspects of platform design was the gap size between the train and the platform. According to the Platform Train Interface Strategy published by the Railway Safety and Standards Board (RSSB) in the United Kingdom (UK) in 2015, a gap size of 75 mm (horizontal) and ± 50 mm (vertical) is considered safe for commuters and enables unassisted accessibility for Persons with Reduced Mobility (PRM). Figure 1-1 presents a schematic of the gap between the train and the platform.

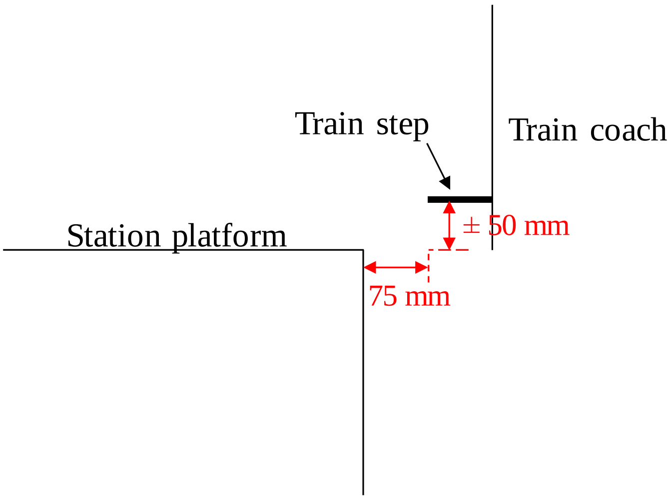

Figure 1-1: Schematic of PTI gap

In South Africa, the Metrorail network shares the same track as freight trains thus making it unsafe to design stations that have a gap clearance (horizontal or vertical) less than 100 mm wide. Due to this limitation, PTI occurrence investigations also incorporates human factors research to determine the optimum solution to eliminate these occurrences - even in the presence of the 100 mm gap clearance. Consequently, this research seeks to quantify the effect of platform and train design on passenger behaviour to determine methods of increasing passenger safety at the PTI.

This research also seeks to determine methods of continuous measurement of the gap clearance as a means of implementing proactive maintenance. The successful implementation of continuous gap measurements would result in the determination of the rate of track deterioration based on the change in the gap size. Access to this data would contribute to the development of an optimum maintenance schedule.

Lastly, the research seeks to determine the optimum combination of interventions that could be implemented to address the shortfalls in platform and train design, as well as to determine measures to mitigate human behaviour that would result in effectively eliminating PTI occurrences.

1.2 PROBLEM STATEMENT

Platform train interchange (PTI) occurrences accounted for the highest number of total annual occurrences according to the 2016/2017 State of Safety Report. These occurrences have the following effects on train operations:

- Disruption of normal train operations resulting in excessive delays, thus resulting in losses incurred by operators. These losses include both financial and reputational losses;
- Loss of commuter confidence (due to unsafe conditions and delays), thus resulting in reduced train ridership
- Possible loss of life

Figure 1-2 presents the statistics on the total annual operational occurrences from 2013 to 2017.

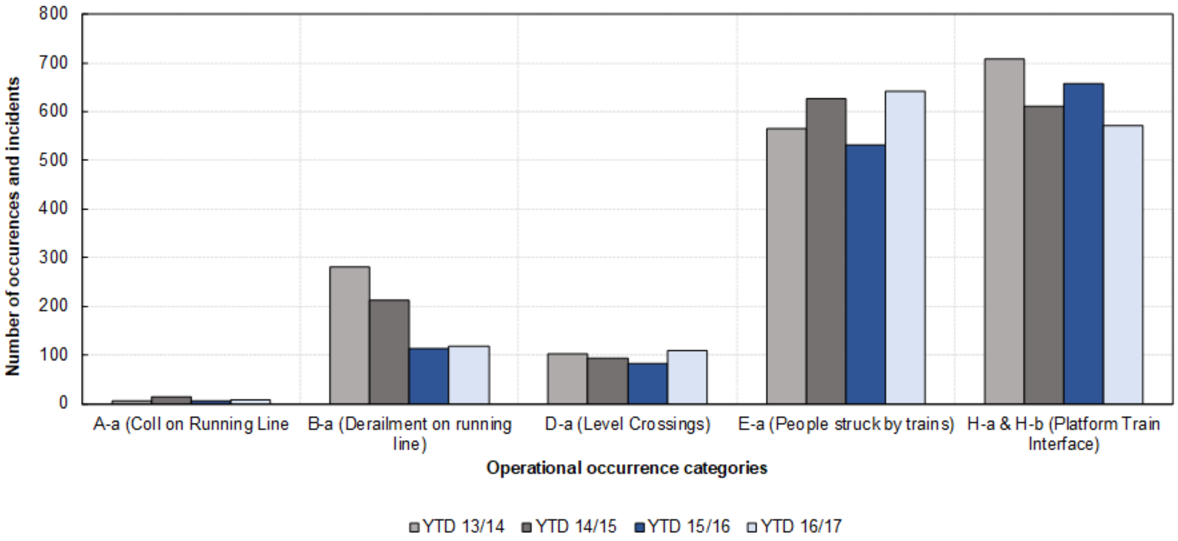

Figure 1-2: Total annual operational occurrences recorded from 2013 – 2017 by the RSR (RSR, 2017).

The numbers presented indicate that the state of safety of South African railways requires further analysis to determine the real underlying causes of these statistics and the reason these numbers remain high despite investigations taking place. For every person killed, injured or disabled due to a PTI occurrence, many others are affected. Affected families are driven into poverty as a result of the loss of a breadwinner, the high cost of prolonged medical care, or the extra financing needed to care for people with disabilities (Ala, 2008). For this reason, research is required particularly in the South African context to determine the possible precursors of occurrences, incidents and accidents and identify means to eliminate and/or mitigate these precursors.

1.3 OBJECTIVES OF THE STUDY

The main objective of this study was to quantify the risk of PTI occurrences taking into account technical, individual and human factors. These factors are further elaborated as:

- Technical factors:
    - The gap size (vertical and horizontal) between the train and the platform;
    - The train design, specifically the functioning of the train doors (manual vs automatic); and
    - Train scheduling and the effect on corridor capacity because of the delays caused by PTI occurrences.
- Individual factors:
    - The susceptibility to experiencing PTI occurrences of the different gender groups;
    - The level of accessibility into the train by special needs passengers; and
    - The possible causes of service inaccessibility of different passenger groups.
- Human factors:
    - Travel patterns of commuters observed depending on the time of day or day of the week which could lead to rushing and over crowding
    - The effect of overcrowding at stations and the likelihood for a PTI occurrence;
    - The effect of overcrowding on passenger behaviour while crossing over the PTI; and
    - The effect of train delays on human behaviour at station platforms e.g. rushing, crowding, frustration etc.

By quantifying the risk of PTI occurrences, the effect of the abovementioned factors can also be quantified and ranked to determine the most critical ones to consider in PTI occurrence mitigation. Furthermore, the quantification of the effects of the abovementioned factors was carried out with the objective of:

- Determining the risk of a PTI occurrence based on station parameters such as the size of the gap at the PTI, the level of crowding at the station and the time/day; and.
- Also, the risk profile of the different gender groups was carried out to determine the individual passenger profile that would be most susceptible to a PTI occurrence.

1.4 SCOPE OF THE STUDY

The proposed study focused on PTI incidents that occurred at Metrorail train stations along the Pretoria-Pienaarspoort corridor, along which freight and passenger trains travel. This corridor was selected for the research because PTI occurrences reported along this corridor accounted for 11% of the total occurrences recorded on the entire Gauteng network during the reporting period of 2014 to 2017, thus classifying it as a high-risk corridor. Also, the new Class 10M4 ("blue") trains were first implemented along this route, therefore creating the opportunity of comparing the service provision with the older Class 5M2A ("yellow") trains

The data used in the research was provided by the RSR and included only the 2014 to 2017 reported occurrences nationwide. All data analysis was focused on the PTI occurrences recorded along the Pretoria-Pienaarspoort corridor. The field measurements were conducted at Metrorail stations and focused only on what occurred as passengers crossed over the PTI.

The assumptions made in the calculations carried out in the research were based on informal interviews carried out with PRASA officials and observations made during the surveys carried out throughout the research. These assumptions being:

1. The percentage distribution of the different gender groups for the Pretoria-Pienaarspoort passenger population;
2. The passenger saturation flow rates for each station along the Pretoria-Pienaarspoort corridor;
3. The train coach capacities for the different train types;
4. The loading factor used to determine how full each train coach gets depending on the time of day; and

5. The train schedule and the train frequency per hour.

1.5 METHODOLOGY

The following methodology was employed:

1. Measurement of the gap size between the train and station platform along the Pretoria to Pienaarspoort corridor:
    - A distance measuring prototype was developed using Arduino technology to test the feasibility of implementing a device at the PTI with the ability to continuously measure the gap size.
    - The vertical and horizontal gap size between the train and platform at each station along the Pretoria to Pienaarspoort corridor was measured using a tape measure to determine the stations with the largest gap sizes.
    - The results obtained from the measurement of the gap sizes along this corridor were used to determine the appropriate stations to carry out the observation surveys that sought to determine the factors that influence passenger behaviour at train stations.
2. Observational surveys were conducted to record passenger behaviour in relation to the gap size and to identify the stress factors that may result in PTI incidents occurring. The data collected included:
    - The number and gender of passengers boarding and alighting the train.
    - The number of passengers looking down as they were boarding or alighting the train (signifying their awareness of the gap).
    - The number of passengers carrying luggage or using a cell phone while boarding or alighting the train.
    - Identification of any other stress factors or behavioural characteristics of commuters that were believed to increase the risk of a PTI incident occurring.
3. The Road Note 34 method (Road Research Laboratory, 1963) was conducted to measure passenger saturation flow at the Metrorail station platforms. The RN34 method is traditionally used to measure vehicle saturation flow at traffic signals, however with slight

modifications the method can be applied in measuring passenger through train doors. The measure of the passenger saturation flow was carried out as a means of quantifying the effects of platform design (specifically the gap size) and train design on passengers' behaviour.

4. A Multilinear regression analysis and Multinomial Logistic regression analysis was carried out on the data provided by the RSR as a means of developing a predictive function to determine the risk of PTI occurrences along a corridor based on inputs such as the vertical gap size at a station, station risk classification, passenger gender and the time and day of the week.

1.6 CONTRIBUTION TO THE STATE OF KNOWLEDGE

The research carried out contributes to the railway industry in South Africa as follows:

1. Conducting a survey dedicated to understanding passenger behaviour at train stations provides a better understanding of the factors that influence passenger behaviour, thus resulting in better planning and station design taking into consideration passenger satisfaction in the South African context.

2. Testing the feasibility of applying a method for continuous gap size measurement at stations allows for the possibility of the adoption of a proactive maintenance approach by PRASA. This approach in maintenance will lead to lower maintenance costs in the long term with regard to track maintenance at train stations.

3. Quantifying the risk of PTI occurrences would contribute to the development of a safety benchmark for PTI occurrences and other safety incidents in South Africa. Developing such a standard would provide operators with a clear benchmark by which to measure the level of safety of operations and service provision. Such a benchmark would be similar to the Common Safety Targets (CSTs) set out by the European Union's Railway Safety Directive (Directive 2004/49/EC).

1.7 ORGANISATION OF THE REPORT

The layout of the dissertation is as follows:

- Chapter 1 serves as the introduction to the dissertation summarising the motivation, objectives, methodology, and contribution of the research to the state of knowledge
- Chapter 2 outlines the literature review on the research topic and other related topics that contributed to the research
- Chapter 3 provides the field tests and results obtained from the tests carried out during the research
- Chapter 4 outlines the data analysis carried out and the conclusions drawn from the results obtained.
- Chapter 5 concludes the dissertation with findings of the research and recommendations.
- Chapter 6 lists the references used in the research
- Appendix A contains results obtained from the gap size measurement, observation surveys carried out and the results from the Multinomial Logistic regression analysis carried out.

2 LITERATURE REVIEW

2.1 INTRODUCTION

The PTI is the gap between the train and the platform. Based on a study conducted by Cheng (2010), the gap size at the PTI is one of the major causes of passenger anxiety associated with train travel. The gap size affects the level of accessibility for train users and it also affects the level of safety at train stations. For example, special needs passengers would not easily have access to the train, or in the worst case, have completely no access to the train if the horizontal and vertical gap sizes were larger than 75 mm and ±50 mm respectively. This specification is based on the Platform Train Interface Strategy published by the RSSB in 2015 (RSSB, 2015). Special needs passengers are classified as people living with disabilities, children/scholars and elderly customers. This classification was based on the Moving South Africa report (Department of Transport, 1999) published by the South African Department of Transport.

Because of the limitation caused by the 100 mm gap clearance requirement at South African stations (Regulator standard on railway stations, 2015), it was necessary to determine interventions that could be applied to limit the effects of reduced accessibility into the train, as well as the effects of reduced safety on the platform for all passengers. According to a study conducted by Wahl (2014), the two main factors that affected passenger safety on station platforms were human behaviour and platform design (gap size).

The following section explores literature on how human factors and platform design influence passenger safety at train stations by exploring methods of measuring human behaviour at train stations and the optimisation of the gap size to ensure accessibility into the train for all passengers (especially special needs passengers). This chapter also explores novel methods of monitoring the gap size in train stations as a means of adopting a proactive maintenance approach to ensure the gap size is always within the required standard. Methods of quantifying the direct and indirect costs of PTI incidents will be investigated, as well as statistical methods that can be employed to quantify the influence of human behaviour and platform design on the risk of PTI occurrences at train stations.

2.2 PLATFORM DESIGN

The platform design, specifically the gap size between the train and the platform, is considered to be one of the main factors that contributes to PTI occurrences (Wahl, 2014). According to the Regulator standard on railway stations (RSR, 2015), a 100 mm clearance (horizontal or vertical) is required to ensure the safe travel of all trains through train stations. This gap clearance is however too large to ensure universal accessibility into the train for all passenger types, whether able-bodied or people living with a disability.

According to the RSSB (2015) a horizontal gap size of 75 mm and a vertical gap size of ±50 mm ensures accessibility and safety for all passengers as they cross over the PTI. This standard was further scrutinised by Moug (2016) in Melbourne, where it was concluded that a more conservative standard was required to ensure universal access and safety for all passengers. The conclusions derived from this study stated that a vertical gap size of 20 mm and a horizontal gap size of 40 mm would ensure the safety of passengers and guarantee unassisted accessibility for people living with a disability.

Both standards outlined in the previous paragraph cannot be implemented at South African train stations due to the clearance requirement stipulated in the Regulator’s standard (RSR, 2014). This clearance requirement may result in large gap sizes at the PTI in train stations throughout the country. For example, it was observed in the study conducted by Wahl (2014) that gap sizes along the Mabopane-Naledi corridor ranged from 200 mm to 400 mm or larger. These large gap sizes were as a result of inadequate maintenance carried out in some of the stations that resulted in the track structure shifting away from the platform over time. This lack of maintenance was as a result of minimal resources dedicated in support of rail as a preferred mode of transport motivated by the De Villiers Report (1986). This report recommended the restriction of further capital investment in the railways.

This lack of resource dedication coupled with maladministration resulted in the extreme degradation of Metrorail train stations and the railway network, where the service is currently operating in a substandard mode (Mathebula and Sopazi, 2016). The result of operating in this mode leads to unsafe conditions for train commuters as well as employees. In addition, the accessibility of special needs passengers is compromised, resulting in the exclusion of this passenger group from the use of the affordable Metrorail service.

In order to reverse the effects of this extreme degradation, interventions to reduce the effects of the gap size are required. These interventions could be applied to either the track structure in stations or the station platforms. Both of these interventions would involve the disruption of normal train operations as well as significant financial investment in the implementation process.

Another approach to consider would be the modification of the passenger train design by developing a mechanical step that adjusts its position according to the size of the gap at each station. This approach would require the development of a mechanical and electronic system operating simultaneously. The mechanical step would be required to adjust its position based on the information received from electronic sensors with the capability of measuring the gap between the train and the platform. This approach would have the added advantage of not disrupting normal operations excessively and it was assumed that the financial investment would be less significant due to the availability of affordable and robust electronic components for the system.

All the interventions mentioned would involve the determination of the gap size as an initial requirement. When considering the intervention on the station infrastructure, the first step taken would be to determine the gap sizes at the different stations to categorise the high risk and low risk stations (based on gap size). When considering the mechanical step, a sensor would be required to accurately measure the gap size between the train and the platform in order to provide feedback for the appropriate position of the mechanical step. For these reasons it was therefore necessary to explore the available methods of gap size measurement.

2.2.1 Methods of measuring the gap between the train and the platform

In Japan, the gap clearance is measured manually on the track using a system composed of a 2D range scanner, a “note-type computer”, an odometer, a battery, and a hand cart (Shimizu, 2016). Figure 2-1 illustrates this technology

Figure 2-1: The current measurement system used in Japanese train stations (Shimizu, 2016)

A laser beam is emitted from the scanner and reflected off the platform surface. The distance between the scanner and the surface of the platform is then measured by applying the time-of-flight principle. This system can also be used to measure the distance and height from the centre of the track to the edge of the platform. In addition, this “current system” could measure the clearance from the track to any point on the platform. This was possible because of the presence of the odometer, which measured the longitudinal distance from the first measuring point to each subsequent measuring point. Measurements were carried out at intervals of 0.5 m along the track at a walking pace (Shimizu, 2016).

One main disadvantage identified by Shimizu (2016) relating to this clearance measurement method was that track closure was always required for measurements to be carried out, thereby disrupting normal train operations. For this reason, a novel system of gap measurement was proposed by using a 3D laser scanner coupled with a stop-and-go operation, where measurements of the gap clearance were carried out from the platform. By carrying out measurements from the platform, the requirement for track closure was bypassed, thereby making it possible to carry out measurements during normal operating hours. The proposed method is depicted in Figure 2-2.

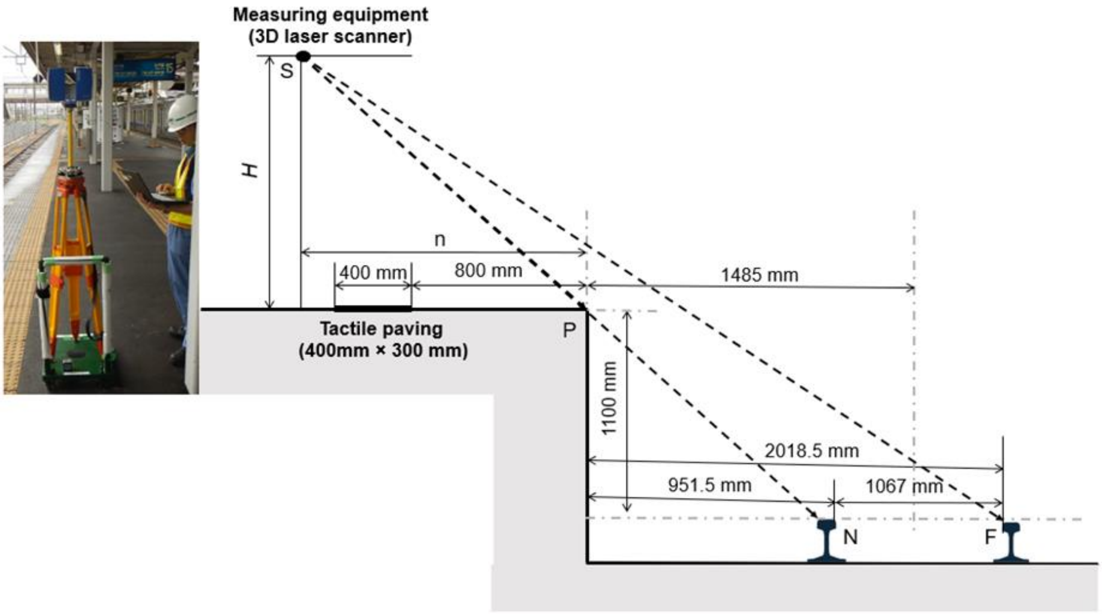

Figure 2-2: Depiction of the system of gap measurement proposed using a 3D laser scanner adopted from Shimizu (2016) (NTS)

Figure 2-2 presents a schematic of the cross section of a typical platform of a Japanese commuter station. According to Shimizu (2016), while testing the proposed method, it was required that a space of 1 200 mm from the edge of the platform be maintained to ensure train traffic safety. Measurements were also carried out at a 10 m interval along the platform to maximize the 3D laser scanner coverage as depicted in Figure 2-3.

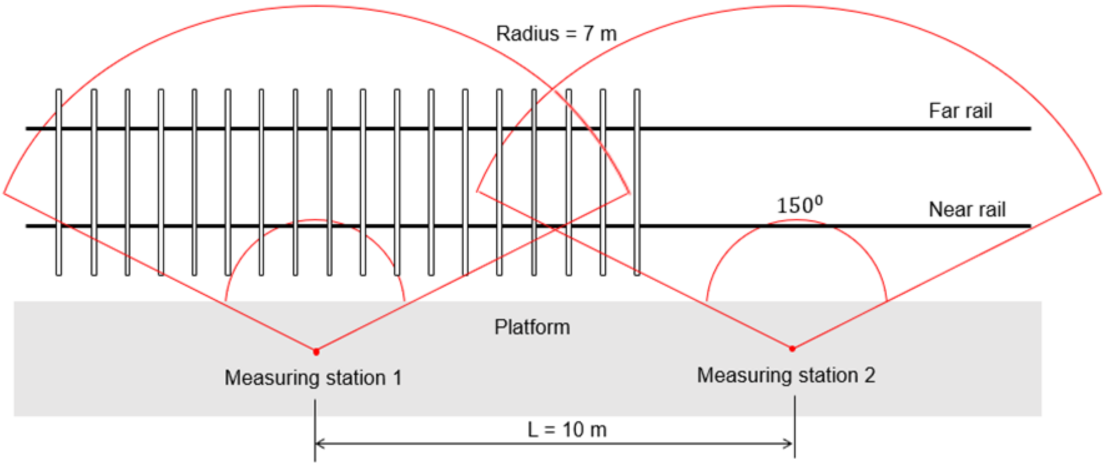

Figure 2-3: Experimental set up of the proposed system adopted from Shimizu (2016) (NTS).

The results obtained from testing the proposed method in Shimizu’s research proved that the method was feasible and extensive in its applications. Using this method, it was possible to get a clear indication of the rail profile and position at an accuracy of 3 mm. Examples of the readings obtained are presented in Figure 2-4 and 2-5.

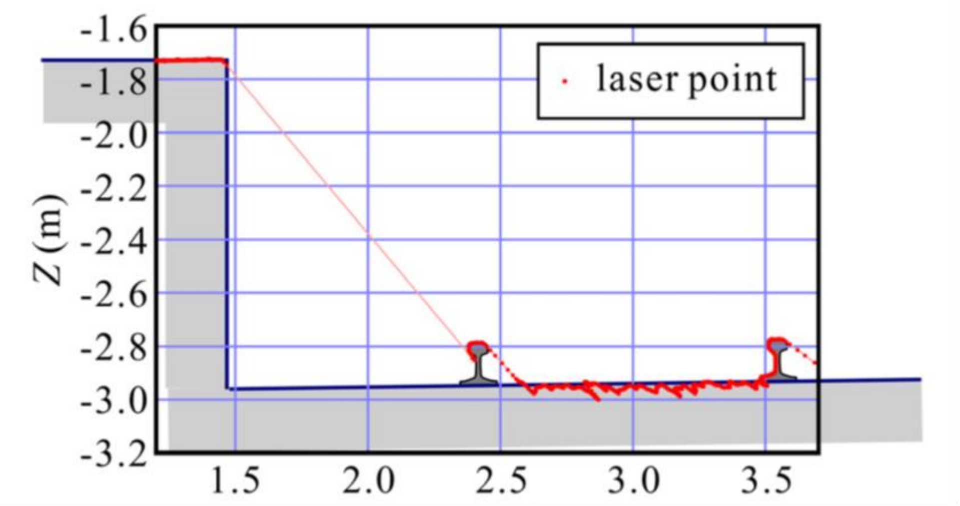

Figure 2-4: Example of readings obtained from the proposed method (Shimizu, 2016)

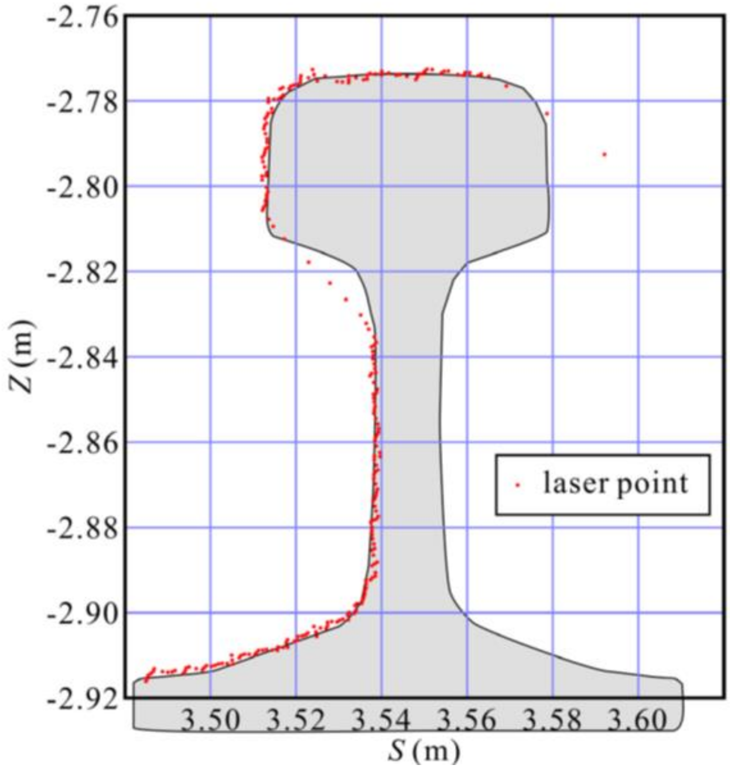

Figure 2-5: Detailed example of readings obtained from the proposed method (Shimizu, 2016)

Figure 2-4 and 2-5 present the data points obtained from the experiment with the x and y axes representing the position of the points relative to the 3D laser scanner. The cross sections of the platform and rail were superimposed onto the graphs to give a clear presentation of the results obtained. These measurements allow for accurate determination of the position of the platform and track structure, thereby making it possible to monitor track movement relative to the platform to an accuracy of 3 mm.

Based on the results presented in Figures 2-4 and 2-5, the proposed method was considered feasible and perceived to have the potential of extensive application in track condition measurement in train stations. However, implementation of the method would involve a significant financial cost in the purchase of the laser scanner (FARO Focus3D X 330) used in the study. For this reason, it was concluded that the method was not cost effective for this research.

Another disadvantage identified during the testing was the interference caused due to passenger intrusion into the measurement path. There were also instances where it was not possible to secure a space of 1200 mm away from the edge of the platform. In such instances setting up the hardware nearer to the edge of the platform was not a solution due to the dead angle of the 3D laser scanner that would have resulted in the platform edge not being measured.

Considering the outlined limitations from the study conducted by Shimizu (2016), a different approach can be considered of applying simple distance measuring sensors that would be used to measure only the gap between the train and the platform. The main advantage of this suggested approach would be the availability and affordability of the components. Furthermore, the successful implementation of the prototype would create the possibility of continuous measurement of the clearance size at the platform during normal operational hours.

Considering the study conducted by Wahl (2014), the two main factors identified that influenced the risk of passengers experiencing a PTI incident were platform design and human factors. For this reason, the development of the optimum solution to mitigating the risk of PTI incidents occurring was not isolated to the modification of the platform design or train design; human factors were also considered in the research.

2.3 HUMAN FACTORS

Research conducted by the New Jersey Department of Transport (2009) showed that the majority of the accidents documented were as a result of human factors and not rolling stock or train operations. According to the study, it was suggested that focusing on passenger characteristics may be more efficient in reducing PTI accidents than changing the platform design. This conclusion was motivated by the results obtained from the study that showed there were no reductions in injuries or risk of injury with the proposed solution of covering the gap between the platform and the train (New Jersey Department of Transport, 2009).

From the study carried out by the NJDOT (2009), no correlation was found between the average gap size and the gap injury frequency or rate, thus implying that the gap size had no observable influence on PTI occurrences. The study found that passenger characteristics, such as age and gender, could influence the risk of a PTI occurrence. For example, the analysis of the gap injuries according to gender showed a significant difference between gap injuries associated with male and female passengers.

It was also found that distractions such as carrying luggage, pushing a stroller, holding a child's hand or using a cell phone influenced the risk of a PTI occurrence. The most significant distraction observed was that of passengers carrying luggage, while cell phone usage was the least common distraction, and this was due to the high noise levels on the platform that made cell phone usage impractical.

Behavioural characteristics, such as the passenger's awareness of the gap as they crossed over the PTI showed to also have an influence on the risk of a PTI occurrence. It was found that 69% of the gap injuries recorded involved female passengers. The study attributed this higher level of PTI occurrences with female passengers to the lack of awareness of the gap. It was observed that 86% of females looked down at the gap, compared to 90% of male passengers. Thus, explaining the reason more women than men experienced PTI incidents.

This same study, however, still concluded that by removing large gaps, opportunities for gap injuries would also be removed, thus the risk of PTI occurrences would decrease. This observation was supported by results of a study conducted by the RSR in South Africa (RSR, 2014) that showed a slight correlation between gap size and PTI occurrences.

The study conducted by the RSR revealed that in the presence of a vertical gap that was 400 mm or larger, a correlation could be observed between PTI occurrence frequency and the gap size, however when the vertical gap size was 200 mm or smaller the effect of the gap size on PTI occurrence frequency was insignificant.

Studies such as the one conducted by the NJDOT (2009) emphasised the importance of considering human factors that could influence passenger safety on station platforms. Such a study proved that in order to mitigate the risk of PTI occurrences, a solution that combined both platform design (e.g. gap size, signage and infrastructure condition) and human factors (e.g. human behaviour, overcrowding and individual factors) was required. To determine the most significant human factors to consider, it was a requirement that the investigation process of PTI incidents be accurate enough to determine the underlying causes of PTI occurrences.

2.4 PTI OCCURRENCE INVESTIGATION IN SOUTH AFRICA

The current trend in incident investigation is that of focusing on the root causes of PTI incidents. According to the State of Safety Report published by the RSR (2017), approximately 60% of all occurrences investigated were as a result of human factor-related root causes. Examples of the root causes identified were negligence and a lack of supervision from the personnel of the train service providers. Other human factors issues identified were passenger behaviour that was largely influenced by the lack of communication and information dissemination in train stations. One of the highlighted human factor-related root causes was that of a lack of training of employees, which can be linked to the issue of skills shortage that has been identified as one of the main shortcomings to be addressed in the Corporate plan published by PRASA (2016/2019).

Despite the shift to focusing on the root causes of PTI incidents in the investigative process, further improvement of the process is still required. According to a study conducted by Hutchings (2017), systemic factors exist in the occurrence investigation process that has resulted in flawed investigations and often not leading to the real reasons for why occurrences happen. An example of the systemic factors identified was the lack of a National Rail Policy in South Africa.

The National Rail Policy is currently a Draft White Paper that was published in June 2017 (Department of Transport, 2017). The goal is for the full implementation of the draft as a Bill that the Railway industry in South Africa would adhere to. The National Rail Policy seeks to address the systemic shortcomings outlined in the study conducted by Hutchings (2017). These shortcomings included financial constraints experienced in the railway industry, insufficient resource allocation and inadequate training of personnel that if addressed would improve the incident investigation process.

Discrepancies were also identified in the number of railway occurrences reported in the media with the number of railway occurrences that happened annually reported by the Regulator. According to Hutchings (2017), if the process is flawed, only the symptoms of accident causation would be identified and therefore the suitability of the recommendations made to address the actual cause would be ineffective. The shortcomings outlined affect the quality, reliability, validity, accuracy and objectivity of the findings and consequently, the relevance of the recommendations derived from the occurrence investigations (Hutchings, 2017).

Internationally, there has been a visible trend in prioritising human factors research in many industries. This trend is as a result of the shift in viewing human interaction with infrastructure as systems of components interacting with each other. The focus of human factors studies is primarily in the interconnections of the system components (Wilson, 2014). A good example where this approach is applied is in maintenance and replacement regimes, where the performance of the whole system is influenced by the interactions between the different system components.

Research conducted by Liming, Jianwei, Guoqiang and Limin (2010) depicts this method of systems using a Monte-Carlo simulation based on fault tree analysis to analyse the reliability of a door system of metro vehicles in an urban railway system. The fault tree analysis depicted the different door components as a system of elements influencing each other (Figure 2-6). Considering a similar approach, the application of a Monte-Carlo simulation to determine the likelihood of a PTI occurrence and the effects thereof would provide a means of quantifying the effects (or cost) of an incident to either the rail service provider or the victim of the occurrence and the rail service users.

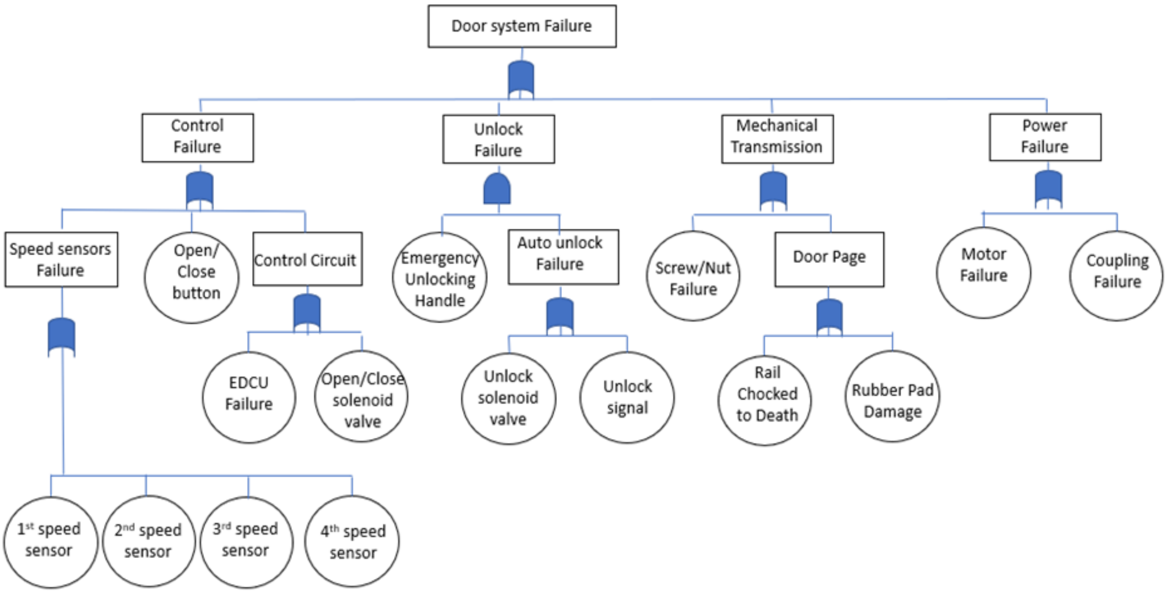

Figure 2-6: Fault tree of a metro vehicle door system adapted from Liming et al. (2010)

Using the Fault tree analysis method, the probability of failure of the door system was calculated based on the failure rate recorded for each minor component. This calculation was carried out by means of a simulation that incorporated Monte-Carlo principles. However, instead of applying the method of generating uniformly distributed random numbers to apply to the Monte-Carlo simulation, pseudo-random numbers were generated. This was carried out using the Linear Congruent Method (LCG) (Liming, 2010). It was assumed that the LCG approach was applied because of its replicability characteristic, where it would have been possible to repeatedly create the same sequence and thus compare alternative systems/models to be used to determine the probability of failure of the system.

In addition, the importance of each component was also ranked based on the effect the component had on the system functioning. The importance of each component was calculated using Equation 2-1.

$$\boldsymbol{W(X_i) = \frac{N_{\Phi(X_i)=0}}{N_{X_i=0}}} \qquad \textbf{Eq 2-1}$$

Where $N_{X_i=0}$ indicated the total number of failures of the basic components, and $N_{\Phi(X_i)=0}$ indicated the total number of system failures caused by component X_i and $0 \leq W(X_i) \leq 1$. If $W(X_i) = 1$, it indicated that the component X_i was a critical component to the functioning of the system. Additionally, the “Model importance” was calculated and defined by Equation 2-2.

$$\boldsymbol{W_M(X_i) = \frac{N_{\Phi(X_i)=0}}{N_{\Phi=0}}} \qquad \textbf{Eq 2-2}$$

Where $N_{\Phi=0}$ indicated the total number of system failures. $W_M(X_i)$ indicated the percentage system failures caused by X_i in all the system failures. By considering both $W(X_i)$ and $W_M(X_i)$, the most critical components were identified and thus given higher priority during inspections and maintenance.

A similar approach to that applied in the study carried out by Liming et al. (2010) can be used in human factors research, especially in the investigation of PTI incidents. The different factors that resulted in a PTI incident can be identified and represent the different components of the system. By applying Liming's approach, the importance of each factor can be ranked to determine the most critical factors that could result in a PTI occurrence.

2.5 FACTORS INFLUENCING PASSENGER BEHAVIOUR

Based on the literature covered, factors that influence the risk of PTI occurrences can be categorised as either platform design, train design, human factors, or a combination of these categories. Further analysis of historic data to determine the recorded causes of PTI occurrences could refine and prove the conclusions drawn from the literature reviewed. In addition, the results from conducting a similar study to that of the NJDOT (2009) would result in the identification of factors at train stations that influence passenger behaviour. These factors would then be categorised as outlined earlier. Thereafter, it would be possible to identify the most critical causes and passenger behavioural factors in both the historical data and the collected data, thereby enabling the prioritisation of the most critical interventions to mitigate the risk of PTI occurrences.

Key indicators can be applied to quantify the effects of platform design and train design on passenger behaviour. One such indicator being the passenger flow rate through the train doors. Evaluating the variation in passenger flow across the PTI would give an indication of the most critical factors that affect the flow rate.

One technique that can be applied in the evaluation of passenger flow is the Road Note 34 (RN34) method (Road Research Laboratory, 1963). This method is used in measuring vehicle saturation flow at traffic signals. The method involves counting the number of vehicles crossing the “stop-line” at a signalised intersection/junction in successive intervals during a saturated green period. According to RN34, the saturated green period (g) was defined as the the saturated time intervals which are free from “lost time” effects $(L_1$ and $L_2)$ as presented in Figure 2-7.

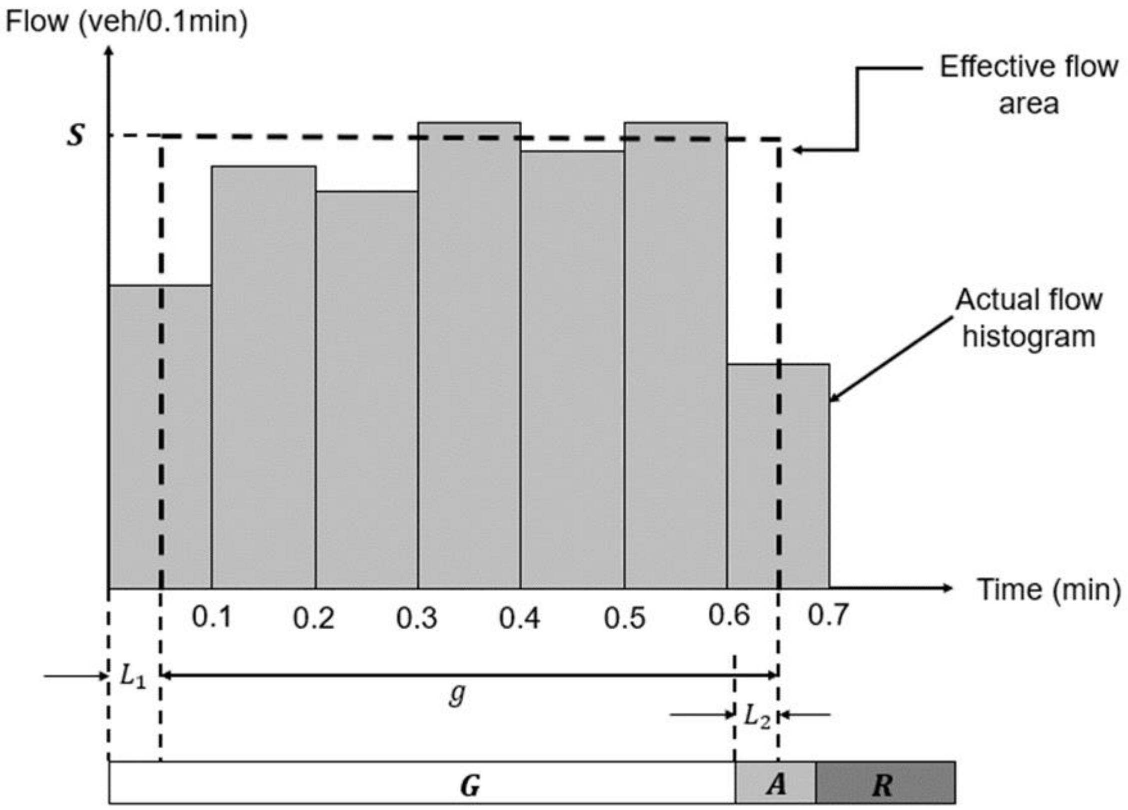

Figure 2-7: Typical flow discharge at a traffic signal adapted from Fernández, Valencia and Seriani (2015).

At the beginning of the green period (G) there is a transient period (L_1) defined as the "start loss", when vehicles begin to flow through the intersection/junction before the discharge rate reaches its maximum rate (S). S is also defined as the saturation flow for that intersection/junction. If a queue remains until the start of the amber cycle (A), another transient period (L_2) is experienced, where vehicles begin to slow down in anticipation of the red cycle and this period is defined as the "end gain" until the start of the red cycle. The effective green time is thus defined as $g = G - L_1 + L_2$.

The RN34 method was used in measuring passenger saturation at the Pedestrian Accessibility and Movement Environment Laboratory (PAMELA) and the Human Dynamics Laboratory (HDL) by Fernández, Valencia and Seriani (2015). The research was carried out through real-scale laboratory experimentation under controlled conditions. The results from the study revealed that similar discharge curves can be derived from measuring vehicle flow at signalised intersections/junctions and from measuring passenger flow through public transport doors, thus validating the feasibility of using the RN34 method to measure passenger flow through metro train doors.

The factors identified that affected passenger flow indicated that the flow rate was dependent on external factors such as vehicle design (door width), platform design (gap size) and the level of overcrowding (Fernández et al., 2015). Because the experiments were conducted in real-scale experimentation conditions, other external factors that would be prevalent at real stations were not considered. These factors were train reliability and effective information dissemination. These two factors have been shown to affect passenger behaviour on the platform. Train delays increase the urgency of boarding the train when the train does arrive, and this results in excessive pushing and shoving at the PTI. Inadequate information dissemination has been observed to only further aggravate this urgency (Wahl, 2014). Therefore, it can be assumed that increased pushing and shoving at the PTI would affect passenger flow rates and increase the risk of PTI occurrences.

A study conducted by Cheng (2010) explored passenger anxiety associated with train travel to further investigate factors that affect passenger behaviour in train stations. This type of research was classified in the qualitative data sphere and this created the issue of quantifying the data collected into a form where statistical techniques can be applied to it in order to derive quantitative results.

In Cheng's (2010) study, 412 train passengers were surveyed, where passengers were required to respond to questions associated with factors that possibly cause passenger anxiety. The answers were presented on a Likert-type scale that provides ordinal data results. The ordinal results were later used to transform the qualitative data to quantitative data on which statistical techniques could be applied. The scale consisted of a five-point score with each point defined as: 5 for “strongly disagree”, 4 for “disagree”, 3 for “neutral”, 2 for “agree” and 1 for “strongly agree”. The possible sources of anxiety were identified as:

1. Accessibility to the train station from original destination
2. The reliability of the timetable
3. Crowding on the platform and in the train
4. Possible delays of the train
5. The gap between the train and the platform
6. Finding a seat at the platform while waiting for the next train

7. Passenger guiding information during transfers
8. Noise on the platform and in the train
9. Taking the train alone
10. Sitting next to a stranger in the train

The results from the survey were then analysed by means of the Rasch model (Rasch, 1960). This model is used to assess an individual's ability and item difficulty through a formulation of the relationship between ability and difficulty and the probability of success (Bezruczko and Lincre, 2005). The Racsh model is an example of an objective measurement method whereby ability and difficulty are measured in logits (log-odd units). This model enables the classification of different passenger groups according to ability and enables the ranking of task difficulty based on the responses from the different passenger groups.

Figure 2-8 presents the conceptual results that could be obtained from the Rasch model when measuring passenger ability and ranking the item difficulty of the sources of travel anxiety. This model was adopted from Cheng (2010).

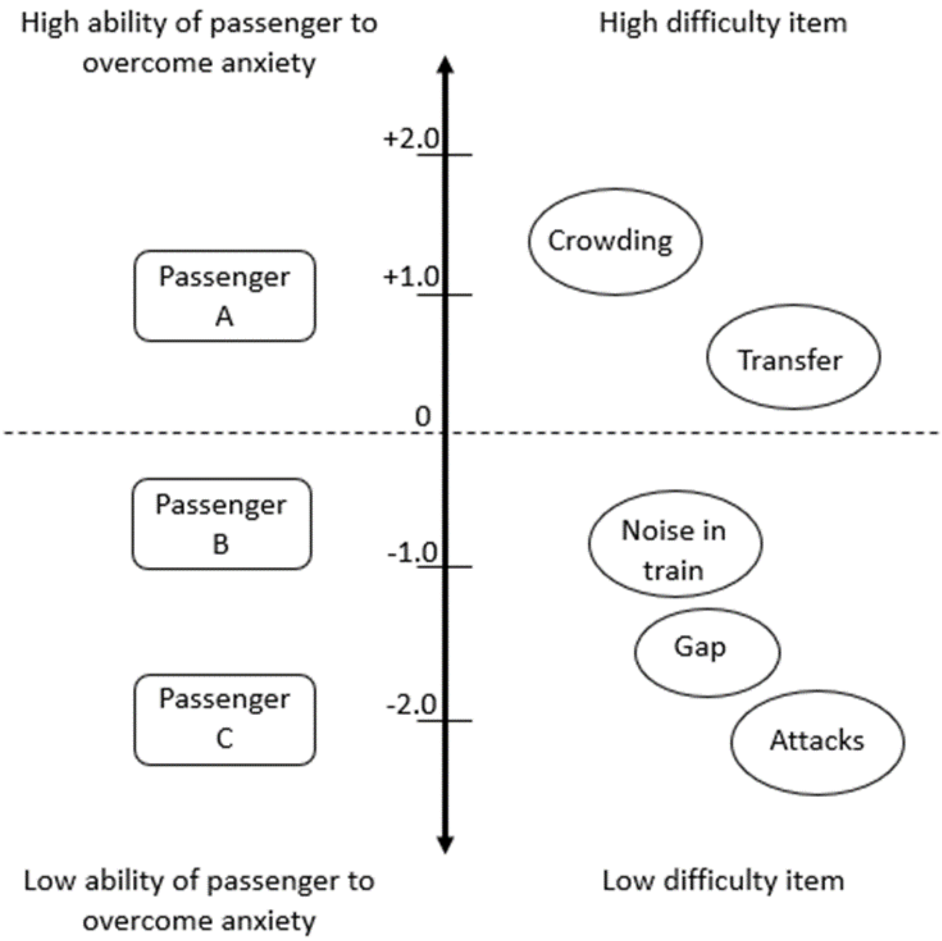

Figure 2-8: Conceptual model for measuring passengers' anxieties by applying the Rasch model (Cheng, 2010)

Referring to Figure 2-8 the y-axis would represent the logit score obtained from the Rasch model, the left side of the y axis represents the different passengers based on the their ability score and the right side of the y axis represents the items ranked according to their difficulty level. By comparing the passenger’s position and item’s position on the logit scale, the relationship between the item difficulty and passenger ability can be identified. For example, considering the position of Passenger B on the logit scale, the diagram shows that Passenger B has the ability to overcome the anxiety caused by noise in the train, the gap between the train and the platform, and possible attacks, however Passenger B does not have the ability to overcome the anxiety caused by train transfers and overcrowding.

The actual results obtained from the study revealed that items such as overcrowding, train delays, noise in the train and the gap between the train and the platform were some of the major causes of passenger anxiety. The most difficult items being overcrowding and train delays. Based on the study carried out by Wahl (2014), train delays caused increased agitation and thus resulted in more pushing and shoving as passengers were crossing over the PTI. This increased aggression affected passenger flow and also, increased the risk of a PTI incidents occurring. The study conducted by Cheng (2010) further supported this observation, where it was found that reducing train delays resulted in a higher percentage of passengers feeling less anxious as compared to reducing the level of overcrowding, thus implying that train delays was found to be a more critical issue than overcrowding.

Research is required to determine whether a similar observation would be found in the South African context. The Metrorail service in South Africa is infamous for having an unpredictable timetable and lengthy delays (Mathebula and Sopazi, 2016). This type of service delivery is as a result of many factors such as: vandalism that results in less trains being available for operation, old rolling stock, and an inadequate signalling system, which, according to Mathebula and Sopazi (2016), results in operations being carried out in a degraded mode.

2.6 SPECIAL NEEDS PASSENGERS

The study conducted by Cheng (2010) highlighted the importance of considering passenger ability. The study highlighted that different passenger groups had different ability levels. The ability levels of a passenger group would determine their likelihood of crossing over the PTI successfully, which in turn would dictate their level of service exclusion. By collecting information on passenger ability levels, train service providers can use this information to determine aspects of the service that need improvement to cater to the abilities of the excluded passenger groups.

According to the action agenda issued by the Minister of Transport and the Department of Transport (DoT) in 1999 (Moving South Africa, 1999), the goal was to develop a transport system that would meet the requirements of all passengers, especially passengers with special needs by 2020. While compiling the Moving South Africa report (1999), it was observed that the most marginalised passengers were those with special needs who were defined as those whose needs differed from the average, able-bodied adult. The Moving South Africa (MSA) team segmented special needs customers into the following three groups:

1. **Life Cycle passengers:** These were defined as customers who had special transport needs by virtue of the fact that they happened to be in a normal stage of the human life cycle. Examples would include children 5 to 14 years old to whom transport is particularly unsafe. Women during pregnancy who may need special assistance or who should be protected from exposure to particular health risks associated with pregnancy. The elderly who, as a result of age-related impairments, require assistance, security and access.

2. **Impairment passengers**: These were defined as customers with physical or cognitive impairments and disabilities from whom special assistance, adapted technologies and special safety requirements are necessary.

3. **Signage passengers**: These were defined as customers who for reasons of literacy, age or lack of familiarity with language of signage were unable to access enough information to use the transport system effectively.

The Action Agenda articulated that its vision was that of a transport system with the capability of differentiating and meeting the needs of all its customers. Achieving this type of transport system would entail knowing the characteristics of the commuters using the service. A transport operator would be required to have information on the types of passengers they were catering for and consequentially the passengers that are marginalised by the service.

According to Statistics South Africa’s 2011 census, 7.5% of the South African population consists of people living with disabilities. In Gauteng, 5.3% of the population are living with disabilities. Without access to public transport this percentage of people find themselves isolated and unable to contribute, or benefit from the services and commercial activities available to most of their fellow citizens. This results in the systemic socioeconomic exclusion of these individuals (Lionjanga and Venter, 2017).

Providing universally accessible public transport would require extra investment into the refurbishment of many of the current operating train stations in South Africa. This increased investment may not be viewed as a favourable financial strategy, however when considering the overall socioeconomic benefits to society, such a strategy is required and is imperative. In the UK public transport operators are becoming more responsive to the needs of PRMs as they are recognising the benefits of an accessible service not just to PRMs and other passengers, but also to their businesses (Tennøy and Leiren, 2008).

In a report published by Fearnley, Aarhaug and Leiren (2015) investigating the benefits of implementing universal accessibility measures, participants in the survey mentioned how reducing the gap between the vehicle and the platform made it easier to board and alight the bus with children and prams. Thus, making it easier for children and parents with babies to travel.

Another prominent feature mentioned in the surveys was having access to real-time information on the vehicle status. Having access to this information reduced passenger travel anxiety and also increased the passengers’ perceived impression of the quality of the service provided. In addition, one of the focus groups participating in the surveys that consisted of mothers expressed a willingness to pay a higher ticket price in order to achieve better accessibility.

The focus group comprising of participants older than 65 years introduced a topic that was hardly mentioned in the other groups, which was safety. This group valued good illumination and clearness both at the stop and on board the public transport vehicle as important for giving them a feeling of security. Being seen felt reassuring for this group and cameras on board the public transport vehicle was requested. Also, in this focus group, the gap between the vehicle and the platform was a prominent theme. It was also mentioned that crowding during rush hour reduced their perception of accessibility into the service. However, this group was not receptive to paying a higher ticket price to have more accessibility measures carried out. They perceived that commuting with public transport was already expensive.

From the surveys carried out in the study by Tennøy and Leiren (2008), it was observed that (in general) the accessibility measures implemented in the three cities were perceived as a quality enhancement rather than being measures specifically targeted towards a specific group, namely PRMs (Figure 2-9). This observation supports the conclusion drawn by Tennøy and Leiren (2008) that accessibility measures increase the quality of facilities for all - not only for people living with disabilities.

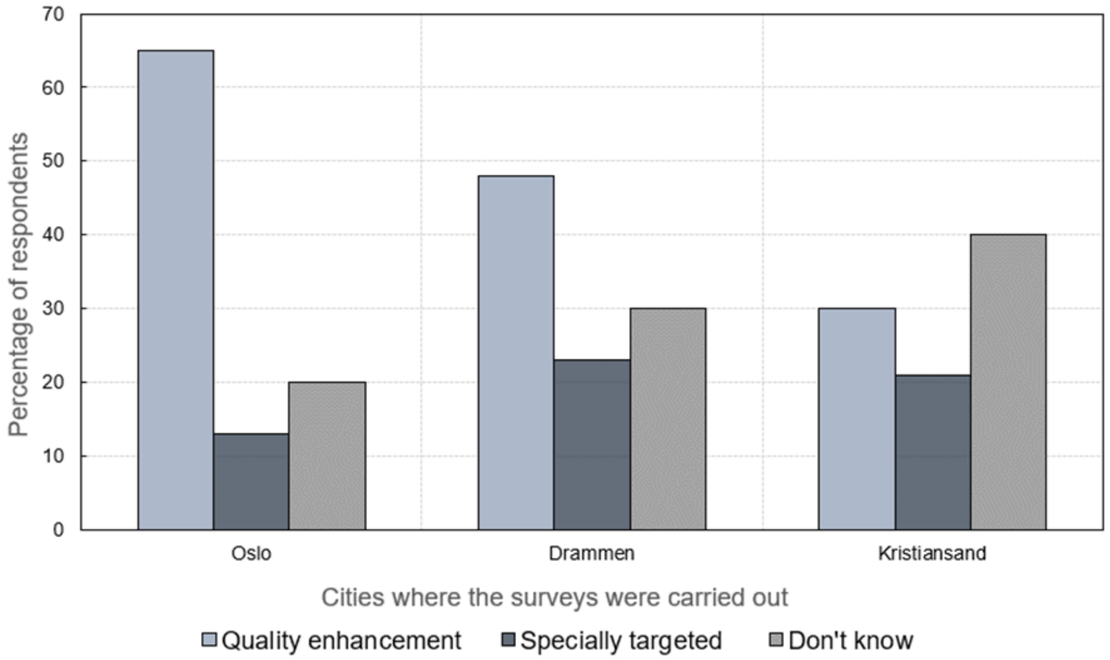

Figure 2-9: Results from the surveys carried out by Fearnley et al. (2015) to measure how the implementation of accessibility measures were perceived by commuters

It was also recorded in the study that a large portion of passengers reported that they travelled more frequently due to the universal design measures. The results are presented in Figure 2-10 and indicate that implementing measures for universal accessibility would not only result in the inclusion of special needs passengers, but would also increase ridership for the general population, thus resulting in a higher revenue stream for the service provider.

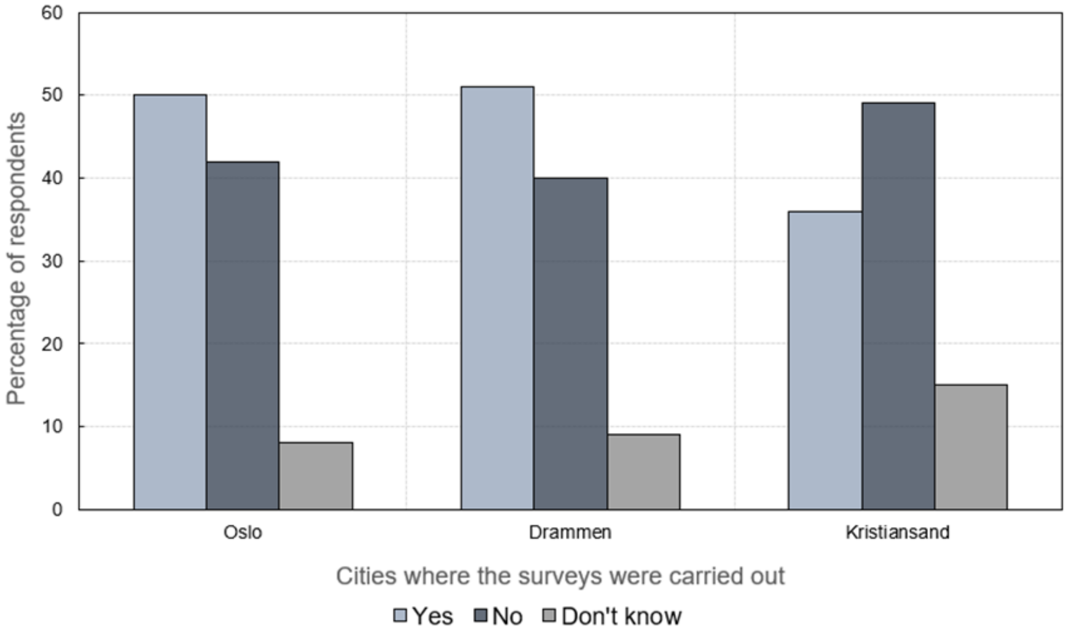

Figure 2-10: Results from the surveys carried out by Fearnley et al. (2015) to measure if the implementation of universal accessibility increased ridership

The results obtained from the study conducted by Fearnley et al. (2015) highlighted that factors such as the gap between the train and the platform, access to real-time information and security all affected the participant's perception of accessibility of the service. The study also highlighted the importance of working towards a compromise in the implementation of accessibility measures while also considering how the extra financial investment would affect the price of travel. It should be noted that by virtue of resorting to increasing the ticket prices of a service as a means to cover the costs of implementing accessibility measures, a group of passengers would find themselves excluded from the service due to their inability to afford the ticket price.

For this reason, universal design measures should be included in a cost-benefit framework, where different interventions may be prioritised, ranked and compared with other investments in the transport sector. In this way, the optimal solution in terms of direct and indirect costs and benefits can be motivated and implemented.

2.7 SOCIOECONOMIC ASPECTS

The ongoing PRASA Station modernisation program was implemented with the goal of making railway the backbone of the public transport system in the respective metropolitan areas (RSR, 2017). As part of the RSR PRASA Station study (2016), it was clearly outlined that PRASA's motivation for the station upgrade program was primarily commercially driven. This approach in itself cannot be criticised because the aim of its implementation was to encourage commuters to move from road to rail, thus increasing the potential revenue to be gained in the railway industry. However, if the issue of safety and accessibility is not addressed at the same time, the ultimate objective of increased train ridership will not be achieved. As highlighted in the previous sections, accessibility and safety have a great influence on user satisfaction, which consequently would affect train ridership and therefore productivity for PRASA.

Based on the Guidance Note published by the World Bank on improving accessibility to transport (2013), to address the factor of passenger accessibility into the train, the following aspects of the station design would require intervention:

1. **The gap between the train and the platform**: Trains should ideally have level access from the station platform to the coach.
2. **Where level access is not possible, the following measures should be taken**:
   - There should be raised areas of the platform that align with the train doors that can accommodate wheelchair users. The use of portable ramps is suggested.
   - Any portable ramps should have the capability of being securely fixed to the train.
3. **Wheelchair users' access**: Portable hand operated lifts can also be used for wheelchair users to enable them to board the train while seated.
4. **Boarding and alighting sections for PRM**: Consideration should be given to the locations where PRMs board. These locations should be away from curved sections of the track where the gap between the train and the platform could be large.
5. **Train door design**: Doorway widths should be a minimum of 800 mm and unobstructed. In addition, train doors should open automatically, or when a passenger pushes a button which should be located no higher than 1300 mm above the floor of the train.

Table 2-1 presents the likely costs and benefits of the accessibility features highlighted earlier.

Table 2-1: The likely costs and benefits of implementing accessibility features at train stations (The World Bank, 2013)

Feature	**Costs**	**Likely benefits**
Incorporating fully accessible rolling stock	≈ 5% additional cost of new accessible rail vehicles to make accessible	Benefits all PRM, but particularly wheel-chair users, to board, alight and move about the train with ease
Portable manual train lifts	≈ \$10 – 14 000 per lift	Enables wheelchair users only to board and alight with ease.
Wheelchair ramps	≈ \$275 per meter	Enable wheelchair users to board, alight and move around the train with ease. Also benefits those with children or luggage.
Automatic doors on train	≈ \$20-25 000 per double width opening door	Makes boarding and alighting easier for all passengers, but particularly those unable to use door handles.
Handrails on stairs	≈ \$140 per meter	Enable PRM (and all other passengers) to use stairs with greater support.

The cost figures presented in Table 2-1 only serve as an indication of the possible costs involved in implementing accessibility features. It should be noted that the figures presented were based on 2013 prices, therefore the inflation rate since then would be considered if these prices were to be applied in the research.

The implementation of these features would not only improve the accessibility into the train for PRMs but would also improve the safety level for all passengers in the station, especially at the station platform. As highlighted in previous studies, safety is one of the main factors influencing the level of ridership in any public transport service (Cheng, 2010; Fearnley et al., 2015). It is a common notion in the transport industry that finding a method of defining the “real" cost of safety incidents on passengers and service operators would effectively motivate the importance of adequately addressing passenger safety issues by the train operator.

An example of a method used to define the "real" cost of occurrences and accidents was applied in finding the systemic cost of risk for heavy haul operations in South Africa. The systemic cost of risk was defined as an artificial risk threshold programmed into the railway over many years using quasi-static conventions and imperfect policies, procedures, standards and guidelines (PPSGs). According to Van Der Merwe, Malan and Havenga (2017), the systemic cost of risk for heavy haul operations can be determined using the method presented in the flow diagram in Figure 2-11. The methodology illustrated in Figure 2-11 only considered the direct costs incurred by the rail service provider.

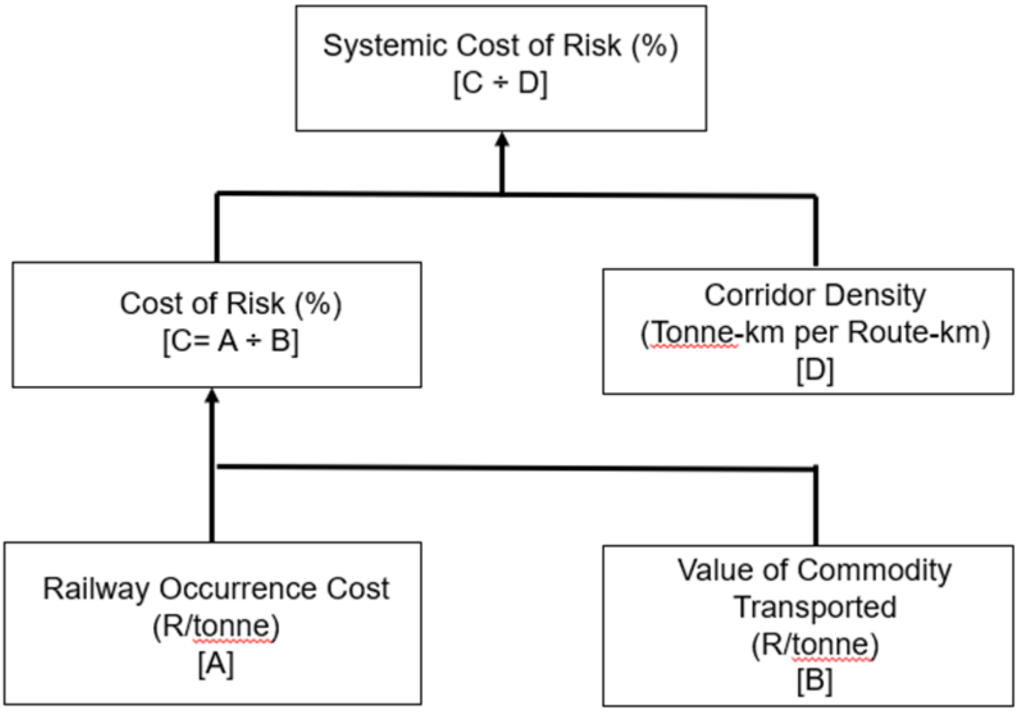

Figure 2-11: Cost of Risk model adopted from Van der Merwe et al. (2017)

The flow diagram presented in Figure 2-11 can also be depicted as an equation for ease of readability:

$$\boldsymbol{Systemic\ Cost\ of\ Risk}\ (\%) = \frac{\boldsymbol{(A \div B)}}{\boldsymbol{D}} \qquad \textbf{Eq 2-3}$$

It was concluded that the main factors contributing to the systemic cost of risk were socioeconomic factors, topography and railway operational considerations. This conclusion was based on the cost of risk determined in the freight rail/heavy haul sector. The same approach could be adopted to determine the cost of risk of PTI occurrences in the passenger rail sector. Figure 2-12 presents a conceptual flow chart similar to the flow chart presented in Figure 2-11 but adapted for application in the passenger rail sector.

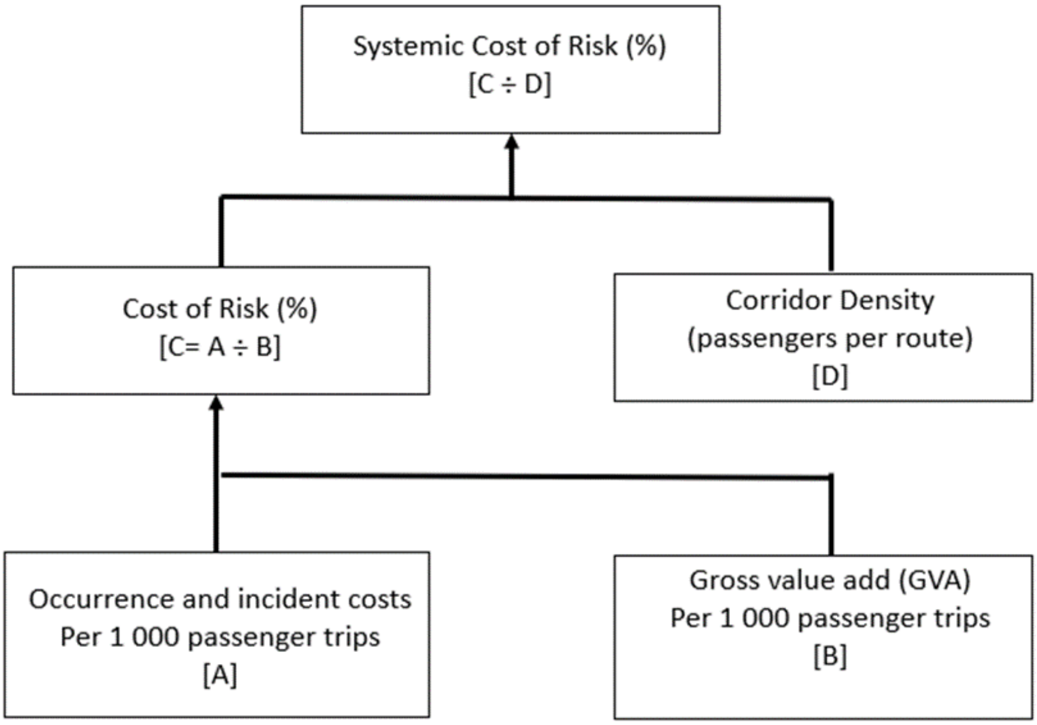

Figure 2-12: Conceptual Cost of Risk model adapted for the passenger rail sector derived from Van der Merwe et al. (2017)

A further beneficial approach to quantifying the cost of risk (specifically from a user perspective) would be to include indirect costs such as opportunity costs, costs incurred due to delays and costs incurred by the individual in the event of an injury. Obtaining this type of data is challenging because of the ambiguity in identifying and quantifying the variables to consider (Cheng, 2010; Martens and Di Ciommo, 2017). In most instances, the Value of Time (VoT) is used to represent the costs to the public transport user.

When evaluating the cost to users in the motivation of transport projects or interventions, the assumption made is that in most instances individuals make a trade-off between all elements that affect their utility of the service, these elements being travel time, travel costs, safety risks, exposure to air pollution etc. (Jara-Díaz, 2007). Research conducted by Martens and Di

Ciommo (2017) revealed that the trade-offs of the elements outlined earlier are systematically unaccounted for when conducting cost-benefit analysis in practice. Martens and Di Ciommo (2017) outlined that the cost-benefit analysis as traditionally employed in practice is primarily a systematic method of measuring benefits and costs using a single denominator of money. These benefits and cost are traditionally expressed using travel time savings. What was proposed in the study conducted was the expansion of the analysis to include accessibility gains in the cost-benefit analysis.

Accessibility gains can be defined as the improvement in the number or range of destinations a person can reach. In the context of researching the PTI in particular, accessibility gains can be roughly defined as the inclusion of special needs passengers into the service, and consequently the improvement of the number and range of destinations that these individuals can access due to accessibility measures implemented on the PTI.

Three distinct accessibility gains were outlined, the first being the identification of benefits of transport investments to persons who would have been at the time excluded from the transport service. The second gain proposed was the assumption that the marginal utility a person benefits from one unit of accessibility gain will decrease if a person's initial level of accessibility to the service was high (Martens, 2006; Metz, 2008). This relationship between service utility and initial level of accessibility is depicted in Figure 2-13.

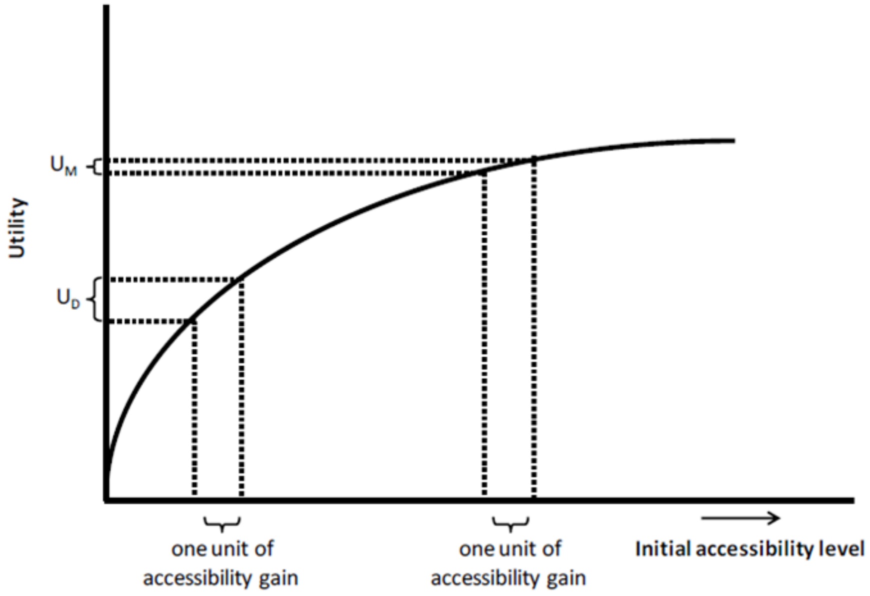

Figure 2-13: Graphical representation of the phenomenon of diminishing marginal returns on utility (Martens and Di Ciommo, 2017)

The relationship depicted in Figure 2-13 shows that by increasing the level of accessibility by one unit, the individuals with the lower initial level of accessibility would generate higher levels of utility than the same unit of accessibility gain for an individual with a higher initial level of accessibility.

The final accessibility gain proposed was that of the effect of the network size. The argument made in the study conducted by Martens and Di Ciommo (2017) was that the impact of transport improvement on overall accessibility will decrease with the size of the transport network, that is, the larger the network, the smaller the increment in accessibility improvement with the addition of a link or improvement of a link. This observation was supported by the research conducted by Levinson and Krizek (2008). For example, considering this observation, an improvement of a developed road link may be expected to generate less accessibility gains than a comparable improvement of an underdeveloped public transport system.

The application of measuring accessibility gains thus expands the costs and benefits that can be applied when motivating the implementation of accessibility interventions (such as the reduction of the gap size between the train and the platform). The gains outlined above propose a method of introducing more factors to consider that would eliminate the tendency of the cost-benefit model steering towards projects that only benefit the able-bodied and high-income population.

2.8 DISCUSSION

Based on the literature reviewed, the primary factors that influence the risk of PTI occurrences can be summarised as platform design, train design and human factors. The platform design is related specifically to the gap size between the train and the platform. According to the World Bank Guidance Note on improving accessibility in transport (2013), trains should ideally have level access from the station platform to the coach. Where level access is not available, provisions should be made to raise areas of the platform that can accommodate wheelchair users (i.e. the use of portable ramps). Such provisions require added financial input from the train service provider, thus increasing the operator's costs. However, Fearnley et al. (2015) argue that the provisions made to improve accessibility for PRMs benefits all passengers (not only PRMs) and increases train ridership, thus resulting in increased revenue for the operator.

The feasibility of any intervention suggested in improving accessibility into the train would require cost effective alternatives. Martens and Di Ciommo (2017) argue that the method of the cost-benefit analysis should be expanded and not only consider the single denominator of money to define the costs and benefits of an intervention. The argument raised is that accessibility gains should be included in the cost-benefit analysis. The inclusion of accessibility gains would consider the population excluded from the transport service, the initial level of accessibility of that population and the effect of the network size. Their argument was that the level of accessibility gains for people who were initially excluded from the service would be higher than for people who had a high initial accessibility level. The introduction of accessibility gains factors would therefore eliminate the tendency of the cost-benefit model steering towards projects that only benefit the able-bodied and high-income population.

Considering the modification of the train design, implementation of technology such as the addition of an adaptive mechanical step at train doors could be a more permanent alternative to the portable ramps suggested in the Guidance Note (World Bank, 2013). The presence of an adaptive step (even if only at specified doors) would have the benefit of reducing train dwell times (compared to having a portable ramp) and provide accessibility into the train for special needs passengers. When considering this alternative, a method of accurately measuring the gap size between the train and the platform would be required. Investigation into the use of open-source hardware and software provided by Arduino is suggested, however, testing of sensor robustness would be required to determine the feasibility of application of the sensors during normal operational hours. The successful implementation of the sensors for continuous gap size measurement could also be applied in collecting data on the track structure deterioration rate at train stations. This data would give insight into the methods of improving platform and track design at stations. In addition, access to this data would also contribute to optimising the maintenance schedule.

Research conducted by the NJDOT (2009) proved that individual factors such as gender, age and distractions at the platform contributed to the risk of PTI occurrences. Key indicators such as passenger flow can also be used in the identification of the most critical factors that influence the risk of a PTI occurrence. The level of variation in passenger flow across the PTI would give insight into the most influential factors to be prioritised for safety interventions.

Cheng (2010) highlights the importance of considering different passenger abilities in overcoming factors in train travel that cause anxiety. It can be argued that a passenger's ability to cross over the PTI successfully determines the level of risk of a PTI occurrence for that passenger. By identifying and ranking the different factors that cause PTI occurrences, it would be possible to prioritise interventions to mitigate the most critical factors. Identification and ranking would also determine the optimum solution to eliminating these occurrences.

3 FIELD TESTS AND RESULTS

3.1 INTRODUCTION

This research focused on PTI incidents that occurred at Metrorail train stations along the Pretoria-Pienaarspoort corridor from 2014 to 2017. The Metrorail service was selected as the focus of the research because of the large number of PTI occurrences reported (by the RSR) at train stations. The Metrorail network is composed of different corridors that are categorised according to the commuter volumes traveling along the corridor. For example, the corridors with the highest passenger volumes were classified as A-corridors, while the medium and low passenger volume corridors were classified as B- and C-corridors respectively. Table 3-1 presents the different corridors and their classification.

Table 3-1: Metrorail corridors, including stations in the Gauteng region

A-Corridor	B-Corridor	C-Corridor
Pretoria	Randfontein	Vereeniging
Leralla	Springs	Oberholzer
Naledi	Pretoria-Saulsville	Nigel
Daveyton	Mabopane-Pienaarspoort via Capital	Booysens
Vereeniging (Midway)	Hercules-Koedoespoort	Pretoria - De wildt
Bell Obmre - Pretoria - Mabopane	Saulville-Pienaarspoort via Capital	
Pretoria-Rissik		
Pretoria-Pienaarspoort		
Kwesine		

The research was conducted along the Pretoria-Pienaarspoort corridor, which was classified as an A-corridor. The PTI occurrences recorded along this corridor accounted for 11% of the total occurrences recorded on the entire Gauteng network during the reporting period of 2014 to 2017, thus classifying it as a high-risk corridor. Lastly, the Pretoria-Pienaarspoort corridor was selected because the new Class 10M4 ("blue") trains that were first implemented along this route, therefore creating the opportunity of comparing the service provision with the older Class 5M2A ("yellow") trains. Figure 3-1 depicts the position of the corridor in relation to the entire Metrorail network in Gauteng.

Figure 3-1: Map of the Metrorail network showing the position of the Pretoria-Pienaarspoortcorridor indicated by the yellow highlighted route Metrorail.co.za, 2018)

This section describes the procedure that was followed in carrying out the field tests to determine the factors that influence the risk of PTI occurrences at train stations. The measure of accessibility of special needs passengers using the train was also carried out to determine the level of exclusion from the Metrorail service by individuals that fell into this category. Special needs passengers consisted of the elderly, children/scholars (aged 5 to 14 years), PRM and visibly pregnant women. This classification was derived from the definition of special needs passengers outlined in the Moving South Africa report published by the Department of Transport (DoT, 1999).

Lastly, the passenger flow rate was measured at specified stations and applied as a method of quantifying the effects of the platform design, train design and human behaviour on safety and operations efficiency. Evaluating the variation in flow rate across the PTI provided an indication of the most critical factors that affected passenger flow, thus allowing for the prioritisation of interventions to optimise the passenger flow rate (especially during peak hours).

3.2 GAP SIZE MEASUREMENT

The gap measurement experiment involved constructing a prototype using Arduino technology and distance measuring sensors. The aim of the experiment was to develop a method of continuous monitoring of the gap size at stations. The prototype consisted of a time-of-flight (laser) sensor and an ultrasonic sensor. The sensors used were the VL53L0X sensor and the HC-SR04 sensor presented in Figure 3-2 and 3-3 respectively, which were chosen based on cost effectiveness and sufficient performance robustness.

Figure 3-2: VL53L0X sensor NTS (SparkFun Electronics, 2017)

Figure 3-3: HC-SR04 sensor NTS (SparkFun Electronics, 2017)

The VL53L0X is a Time-of-Flight (ToF) laser-ranging module housed in a 4.4 × 2.4 × 1.0 mm package and had the capability of measuring absolute distances of a maximum of 2 000 mm. The HC-SR04 is an ultrasonic sensor that measured distances within the range of 20 mm to 5 000 mm. This sensor was accurate to approximately 3 mm. The ultrasonic sensor included an ultrasonic transmitter, a receiver and a control circuit.

The microcontroller used in the prototype development was the Arduino Uno REV3. The microcontroller was manufactured and designed by Arduino, which is an open-source hardware and software company (Arduino.cc, 2018). The Arduino Uno REV3 is presented in Figure 3-4 and the prototypes that were developed are presented in Figure 3-5.

Figure 3-4: Arduino Uno REV3 microcontroller (Arduino, 2018) (NTS)

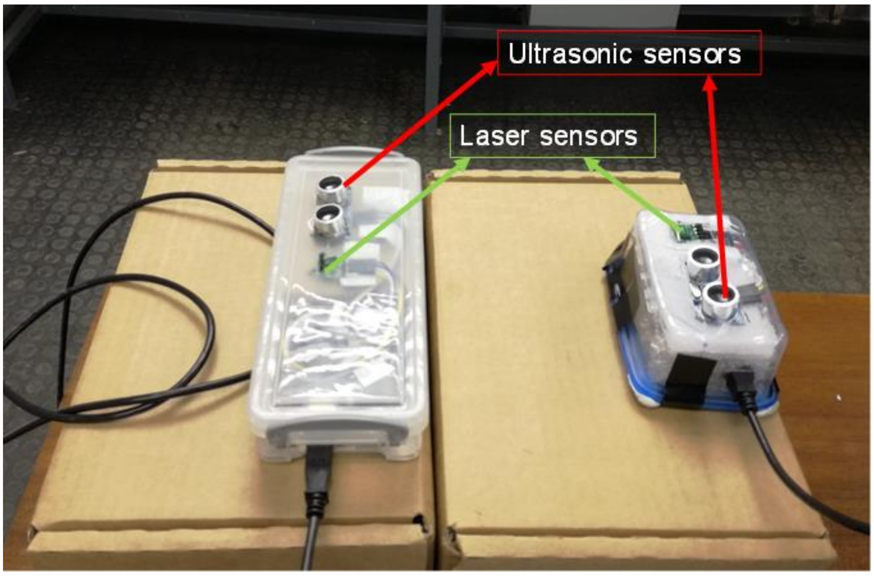

Figure 3-5: Arduino prototypes developed

The continuous measurement of the gap in the long term would provide data on the rate at which the gap size changes, consequently providing an indication of the deterioration rate of the track structure. By obtaining this data, it would be possible to design a proactive maintenance regime that could ensure that the gap size is kept within the required standard. This approach to maintenance would be especially relevant in stations with a ballast track design, where it has been observed that the deterioration rate is higher than that of a ballastless track structure experiencing similar conditions (Čebašek, Esen, Woodward, Laghrouche and Connolly, 2018).

The development of the prototype involved the calibration of the sensors and thereafter, testing the feasibility of field applications. The feasibility testing involved conducting experiments on the prototypes under different conditions, namely stationary conditions and mobile conditions. The approach taken was to compare the gap size measured by the prototype versus the gap size measured manually using a tape measure. The following experimental layouts were used to create the abovementioned conditions:

- **Stationary conditions**: Measurement of the gap size between the train and the platform at Rissik station, which was one of the stations along the corridor that underwent renovations as part of the PRASA Station modernisation program. The prototype was mounted on the platform.

- **Mobile conditions**: Measurement of the gap size between a wall and a moving vehicle with the prototype mounted on the moving vehicle

The experiments were conducted to test the performance of the sensors in different conditions to determine their reliability in long term application and measurement in the field. The successful implementation of the prototype would result in a cost-effective gap measurement method that could create the possibility of the development of an adjustable mechanical step on trains. This development could result in the implementation of a timely and cost-effective solution to the issue of large gap sizes at Metrorail stations throughout the country.

3.2.1 Calibration process

The aim of the calibration process was to determine the accuracy of the sensors' readings, to compare the performance of the laser and ultrasonic sensor, and lastly, to test the performance of the sensors in field application. The following calibration plan was developed:

1. Testing the effects of the target reflectance level.
2. Testing the effects of the shape of the target.
3. Determining the optimum measuring frequency to be applied.

The first calibration test involved carrying out measurements against a black (non-reflective surface) and silver (reflective) surface to determine the effect of the surface's reflectance on the sensors' performance. The measurements were carried out at distances ranging from 100 mm to 500 mm at 100 mm intervals. The experimental setup is presented in Figure 3-6 and 3.7.

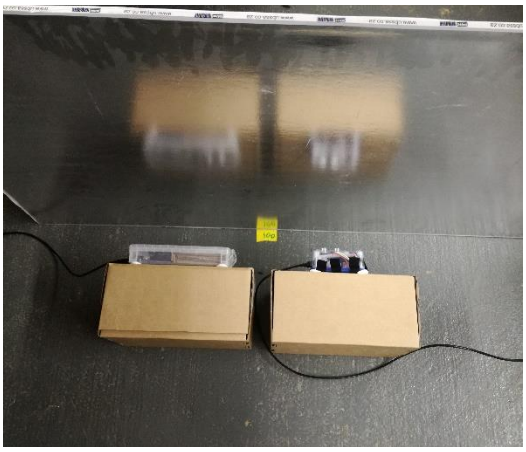

Figure 3-6: Silver reflective surface

Figure 3-7: Black non-reflective surface

The percentage error of the sensor readings obtained at each distance is plotted and presented in Figure 3-8 and 3.9.

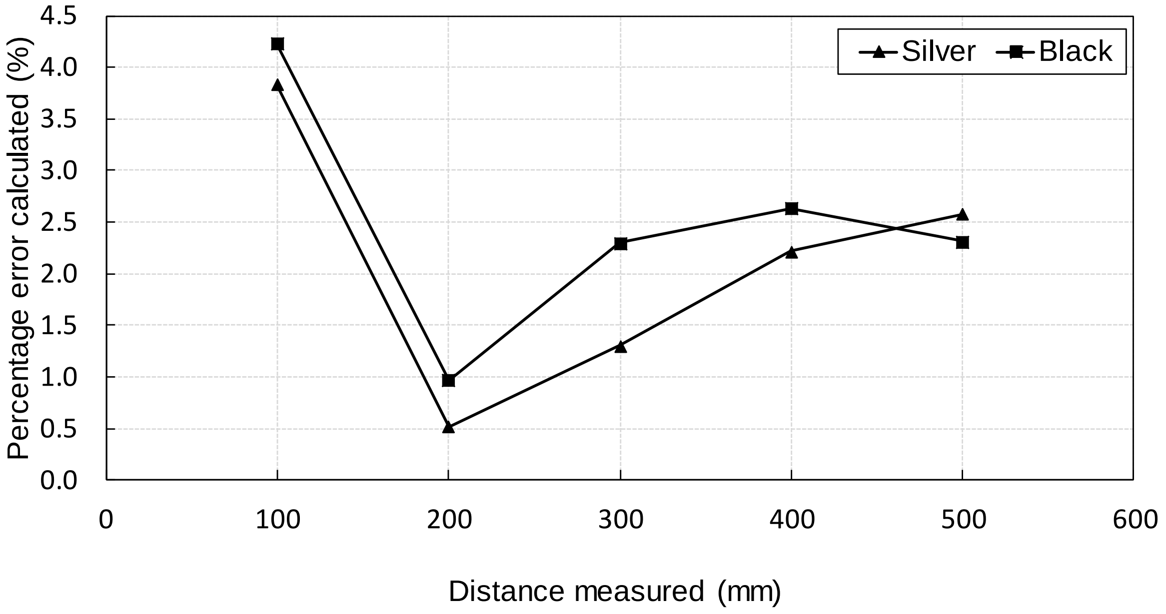

Figure 3-8: Percentage error calculated from the ultrasonic sensor readings

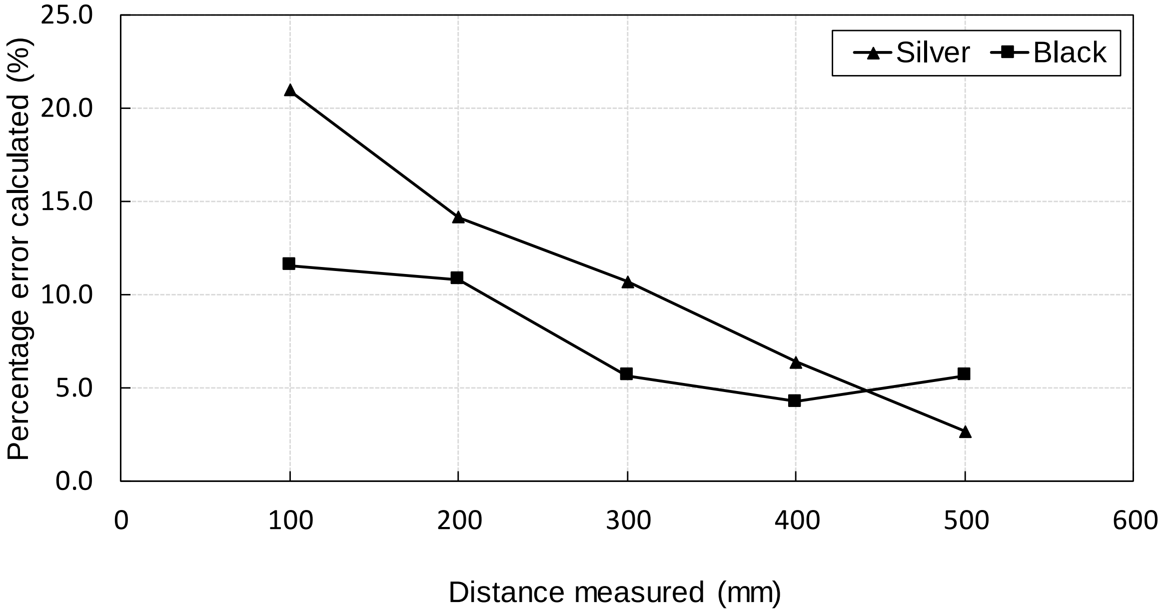

Figure 3-9: Percentage error calculated from the laser sensor readings

Based on the results in Figure 3-8 and 3-9, the lowest percentage error was obtained from the ultrasonic sensor readings and the trend observed in the readings were similar when measuring the distance from the black and silver target, thus indicating that the sensor was not affected by the target reflectance. However, the higher accuracy observed with the silver surface implied that the texture of the surface may have had an influence on the performance of the ultrasonic sensor. The silver surface was smooth, while the black surface was a rough rubber material. It is expected that the ultrasonic waves emitted would reflect more effectively against a smooth surface.

When comparing the readings obtained from the two sensors, the laser sensor measurements were less accurate than those of the ultrasonic sensor. The measurements obtained from the black target were more accurate than those of the silver target. From the results obtained it was concluded that the laser sensor was susceptible to the effects of the target reflectance level, therefore application of measurement correction factors would be required when taking measurements against a reflective surface.

The second calibration test involved carrying out measurements against an irregular shaped target. The measurements were carried out at distances ranging from 100 mm to 500 mm at 100 mm intervals. This experiment was carried out to determine the performance of the sensors in the field in a scenario where the prototype was mounted onto the platform and reflection occurred on the surface of the train. The shape of the train exterior at the level the sensors would be mounted was irregular, thus necessitating the investigation of the effects of the target's shape on the accuracy of the readings. The experimental setup is presented in Figure 3-10. The results obtained from the experiment are presented in Figure 3-11.

Figure 3-10: Experimental setup of an irregularly shaped surface

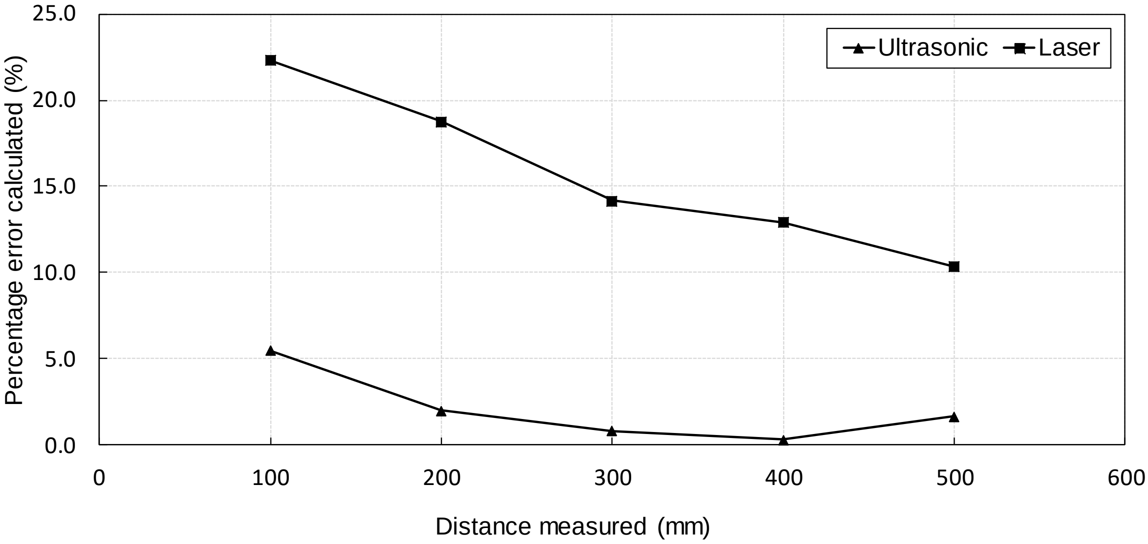

Figure 3-11: Percentage error calculated from the measurements obtained from the ultrasonic and laser sensors

Based on the results presented in Figure 3-11, the lowest percentage error was calculated from the ultrasonic sensor readings thus indicating that the ultrasonic sensor would be more reliable when taking readings against a target with an irregular shape.

The final phase of the calibration process was to determine the optimum measuring frequency. The aim of this phase was to determine the ideal frequency at which the sensors would perform as required. Performance was based on the accuracy of the readings (measured by the percentage error) as well as the stability of the readings (measured by the standard deviation). The approach taken was to measure a distance of 100 mm against a smooth cardboard surface at different frequencies ranging from 4 Hz to 100 Hz. Thereafter, the percentage error and the standard deviation were plotted against the frequency. The results are presented in Figure 3-12 and 3.13.

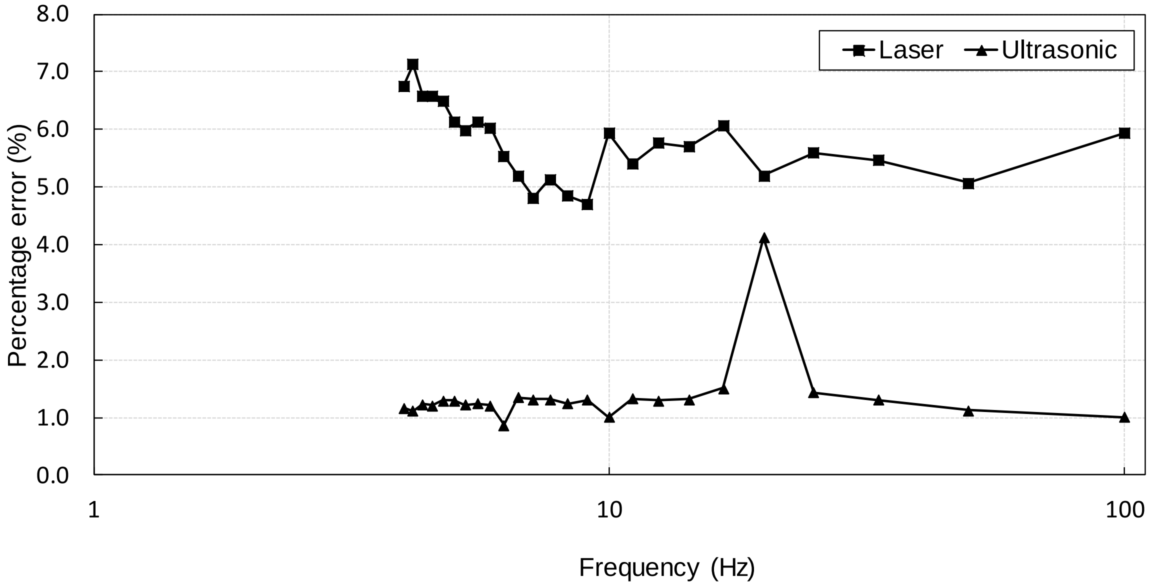

Figure 3-12: Percentage error calculated from the ultrasonic and laser sensor readings against the corresponding measurement frequencies

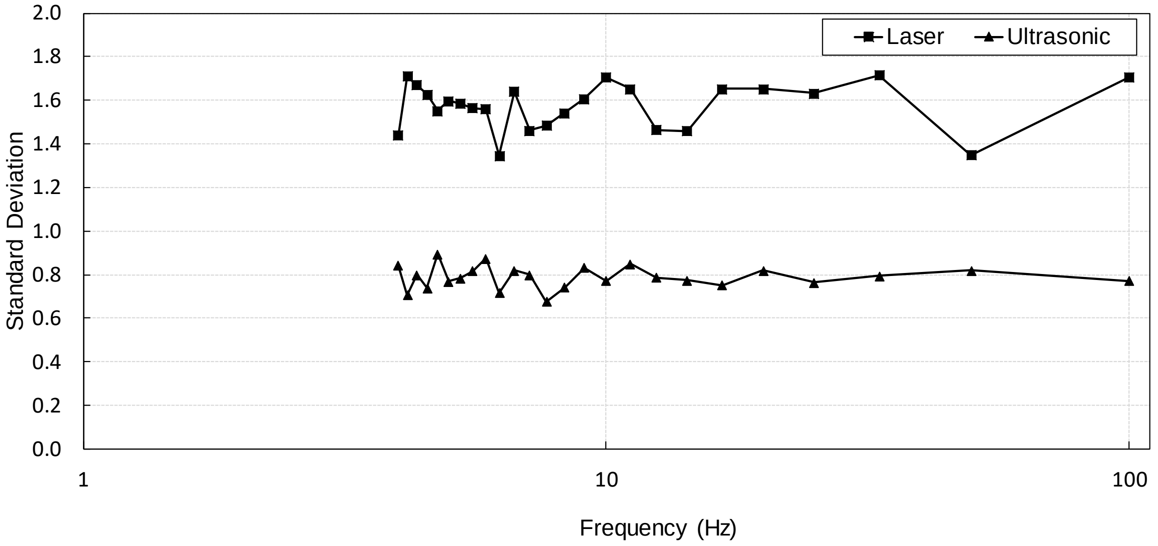

Figure 3-13:Standard deviation calculated from the ultrasonic and laser sensor readings against the corresponding measurement frequencies

Based on the results presented in Figure 3-12 and 3-13, on average, the laser sensor readings displayed a higher percentage error and standard deviation than the readings obtained from the ultrasonic sensor, indicating the better performance of the ultrasonic sensor. However, considering the readings obtained from the ultrasonic sensor, a significant increase in the percentage error and standard deviation was observed at 20 Hz. This observation may have been as a result of a measurement error and was therefore considered an outlier and excluded from the results presented in Figure 3-13.

It was decided that the overruling factor to be considered when determining the optimum frequency would be the stability of the readings because the inaccuracy of the readings could be remedied by applying a correction factor or by applying the moving average method. The findings are summarised in Table 3-2 presenting the measurement frequencies corresponding to the lowest standard deviation obtained from the sensors' readings.

Table 3-2: Summary of findings to determine the optimum frequency

	Laser	Ultrasonic
Minimum Std dev (mm/mm)	1.35	0.68
Corresponding frequency (Hz)	6.25	7.96

3.2.2 Field Testing

The second phase in the prototype testing was to determine the sensors' performance in the field, specifically at a train station during normal operating off-peak hours (14h00 to 15h00). The following conditions were considered:

1. **Stationary conditions**: Measuring the stationary horizontal gap size between the train and the platform at a Metrorail station, where the device was mounted on the platform;
2. **Mobile conditions**: Measurement of the gap size between a wall and a moving vehicle with the device mounted on the moving vehicle

The aim of these tests was to determine the robustness of the prototype during field operation, with the goal of possible implementation of the prototypes for continuous gap measurement. The main characteristics considered in measuring sensor reliability was measurement accuracy (percentage error) and measurement stability (standard deviation)

The first field test was carried out at a Metrorail station to determine the stationary horizontal gap size at the PTI as the train was idling. The experimental configuration involved mounting the prototypes on the platform as presented in Figure 3-14. The horizontal gap size measured at the sensor position on the platform was 100 mm (on average).

Figure 3-14: Prototypes mounted on the platform to measure the horizontal gap size between the train and the platform

The results obtained are presented in Figure 3-15. The following steps were carried out in improving the quality of the raw data obtained from the prototypes:

1. Removal of outliers using the Box and Whisker approach.

2. Shortening of the reading interval to only focus on the stationary readings.

3. Calibration of the readings obtained from the ultrasonic and laser sensors using data translation.

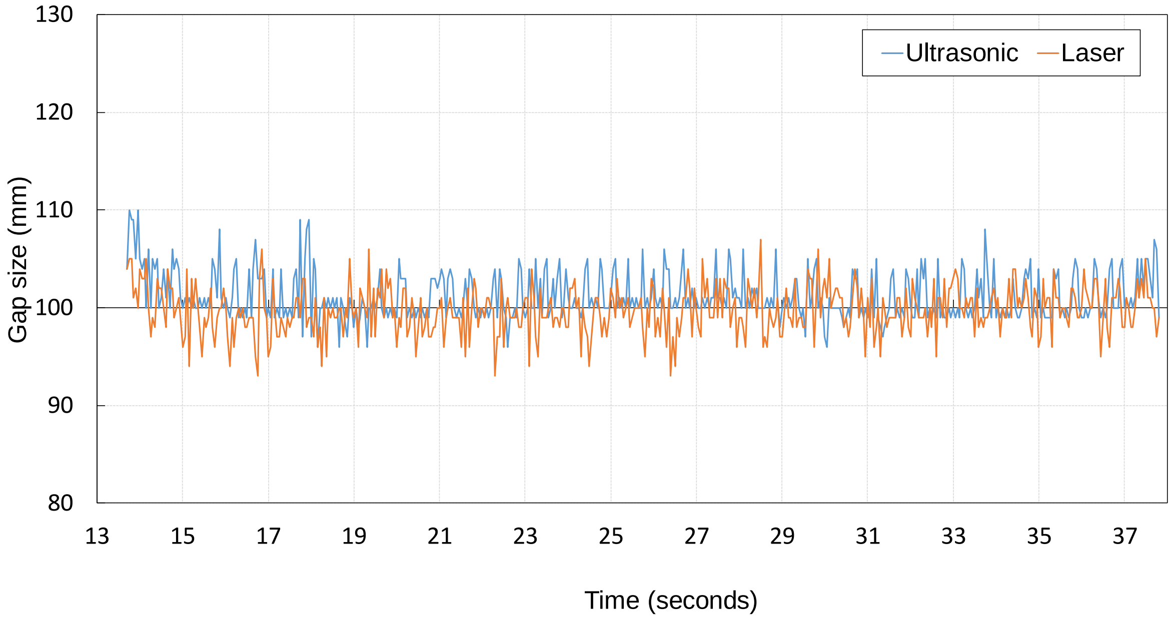

Figure 3-15: Amended horizontal gap size measured of stationary train in the station

The accuracy of the sensors was further described using data characteristics such as the standard deviation and variance. These characteristics are presented in Table 3-3.

Table 3-3: Data characteristic of the horizontal gap measurements

	Raw data		Calibrated data	
	Ultrasonic	Laser	Ultrasonic	Laser
Average gap size measured (mm)	101	125	100	100
Standard deviation	2.44	2.42	2.44	2.42
Variance	5.95	5.84	5.95	5.84

With the removal of outliers and calibration, the sensors proved to be accurate and robust enough for application in stationary conditions. The second field test involved mounting the distance measuring prototype on a moving vehicle to determine the performance of the sensors under mobile conditions. A road vehicle was used in this experiment to simulate the stopping period as the train approaches the station platform. The prototype measured the gap between the vehicle and the wall, similar to how (it would be assumed) the gap between the train and the platform would be measured. This experiment was considered an approximation of the conditions present at a normal operating station bearing in mind that the magnetic interference present while the train is travelling could not be simulated in the experiment. However, there was the advantage of having the ability to control the speed of the vehicle and the distance between the vehicle and the wall during the prototype testing.

The experimental plan involved testing the accuracy of the prototype at different distances. These distances being 750 mm, 1100 mm and 1400 mm. The road vehicle travelled at a speed of 10-15 km/h during the measurement period simulating the speed of a train arriving at a station. The experimental setup is presented in Figure 3-16 and 3-17.

Figure 3-16: Experimental setup depicting measurements recorded at a distance of 750 mm

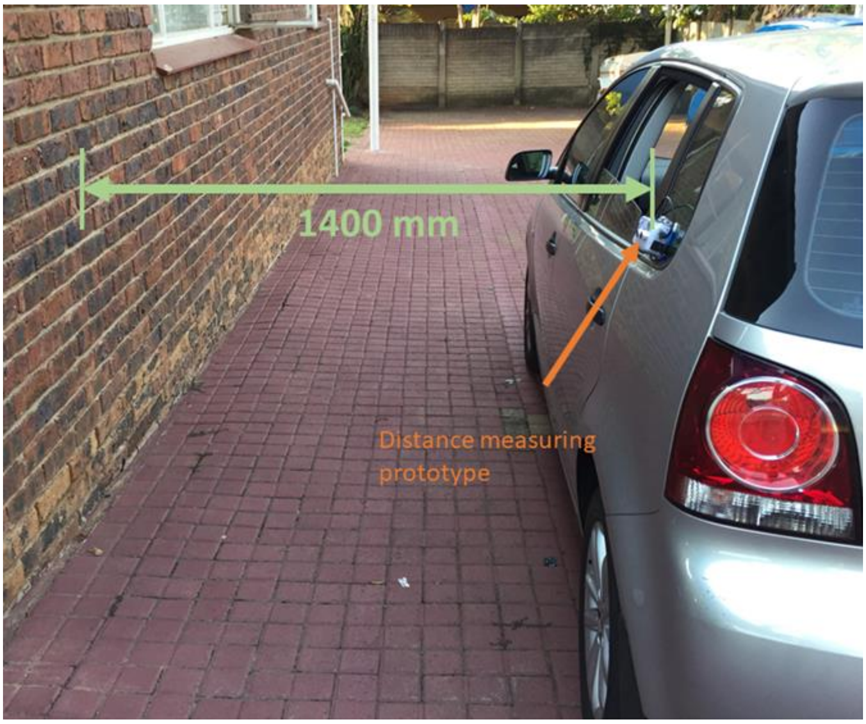

Figure 3-17: Experimental set-up depicting measurements recorded at a distance of 1400 mm

During testing, the readings were stable before the vehicle started moving. However, once the vehicle was in motion, the readings became erratic. This trend is presented in Figure 3-18 and 3-19. Through calibration and application of the moving average method, fairly accurate measurements were obtained from both sensors (Refer to Table 3-4). The results presented indicate that the sensors' reading accuracy improved with increasing distance, however, it should be noted that the readings taken at larger distances were carried out last, thus implying that in fact, the sensor readings improved with time. This observation was attributed to the "warm-up drift" that occurs with sensors when they are initially powered up; the components heat up and this temperature fluctuation from a cold start was assumed to have affected the initial readings taken at shorter distances (Banner Engineering, 2019). Future application of the sensors would require initial readings be taken before full operation of the sensors to compensate for the "warm-up drift" as specified by the sensor manufacturer.

Table 3-4: Summary of percentage error obtained from the recorded measurements

	750 mm		1100 mm		1400 mm	
	Ultrasonic	Laser	Ultrasonic	Laser	Ultrasonic	Laser
Percentage Error	2%	6%	7%	1%	2%	Out-of-range

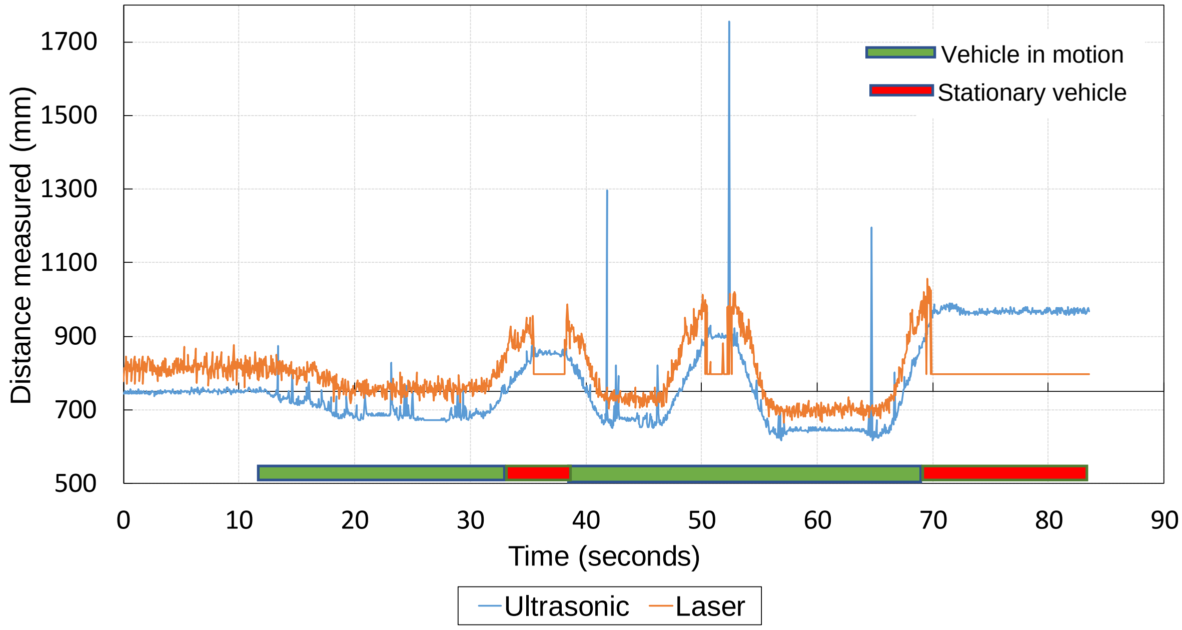

Figure 3-18: Results obtained from the 750 mm measurements

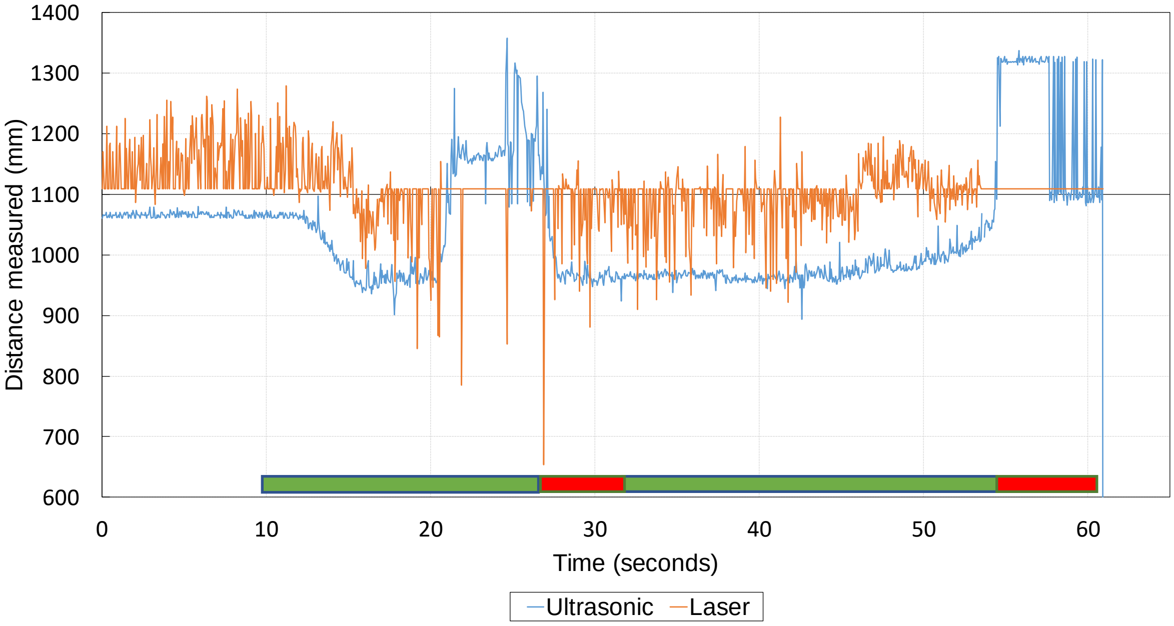

Figure 3-19: Results obtained from the 1100 mm measurements (ultrasonic sensor readings only)

The performance of the sensors differed depending on the distance measured and a slight drift was observed in the ultrasonic sensor readings as the recording period lengthened. This was to be expected according to the manufacturer's guide that stated an allowable drift of approximately 3% would be present with the lengthening of the measurement period. The drift observed was best depicted after the calibration of the readings obtained from the 1400 mm measurement (Refer to Figure 3-20).

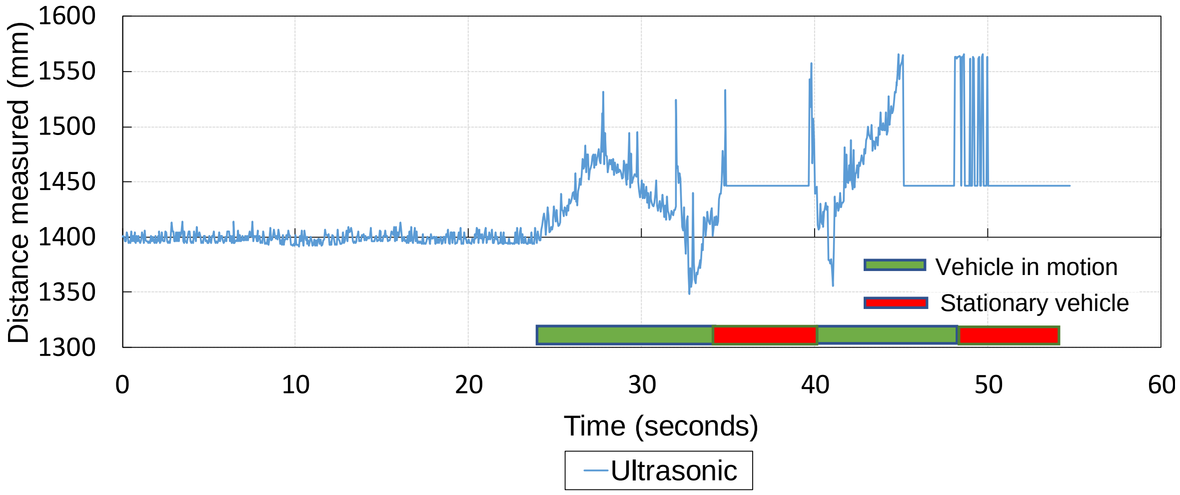

Figure 3-20: Results obtained from the 1 400 mm measurements

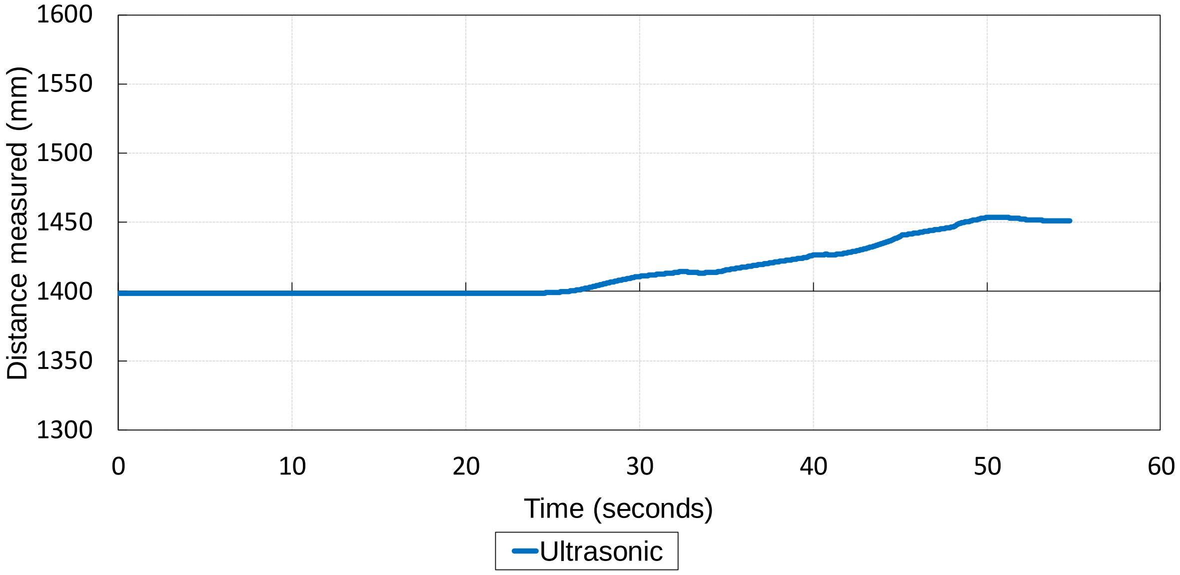

Figure 3-21: Results obtained after applying the moving average method to the data presented in Figure 3-20

Further analysis of the data collected would always be required when measuring distance using the sensors under stationary conditions and conditions of motion. Data cleaning techniques such as the translation of the data points and the application of the moving average method are examples of techniques that can be applied to the raw data collected. Alternatively, these techniques can be written into the coding that dictates the functioning of the sensors.

The application of the sensors in measuring the gap at the PTI at an operating station was not tested because of safety clearance constraints placed by the Metrorail. The successful implementation of the prototypes would require clearance from the service operators to incorporate them into the train design or mount them onto station platforms. Therefore, the testing of the prototypes in mobile condition was not carried further into implementation during normal train operating hours.

3.2.3 Manual Measurement of the gap size along the Pretoria-Pienaarspoort corridor

During testing under stationary conditions at Rissik station, the experimental configuration resulted in only the horizontal gap size at the PTI being recorded. Limitations to the mounting configuration were placed due to safety requirements at the station to ensure the safe travel of the train through the station. The vertical gap size was therefore measured using a tape measure. Furthermore, the manual gap size measurement was expanded to include all stations along the Pretoria-Pienaarspoort corridor. This measurement was conducted by a team of 5 volunteers who measured the gap size at five doors on the train to obtain the range of gap sizes along the platform at each station.

The results obtained were also used to identify the possible high-risk stations along the corridor. A station was considered high risk if the gap size (vertical and/or horizontal) was approximately 300 mm or larger. Conversely, a station was considered low risk if the gap size (vertical and/or horizontal) was 200 mm or smaller (Wahl, 2014). According to a study conducted by Spanjaard, Reeves, Van Dieen, Baltzopoulos and Maganaris (2008) increasing the step height from the stardard 170 mm rise results in increased strain on the ankle and knee joints when ascending the step. This increased strain reduces the likelihood passengers (especially special needs passengers) crossing over the step successfully (Cheng, 2010).

The proposed data collection was carried out over a period of two days, during off-peak hours (7h30 to 11h30) for ease of data collection and safety reasons. The measurements were carried out on the Class 10M4 (“blue”) and Class 5M2A (“yellow”) trains to compare the services in retrospect. The data was collected using the data form presented in Appendix A.

The information collected using the form included:

1. **The station name and the type of track structure** (ballast or non-ballast).

2. **The number of platforms** were counted to determine whether the station was an "end station" or a "through station".

3. **The number of security personnel** on the platform and in the train were counted to give an indication of the level of security presence. It was assumed that visible security presence would assist in crowd control and thus reduce the risk of a PTI occurrence as a result of unsafe human behaviour such as pushing and shoving during peak hours.

4. **The level of crowding** in the train and on the platform was measured. This was carried out using a Likert scale ranging from 1 to 5 with 5 being the highest level of crowding and 1 being the lowest level of crowding (based on the user's perception). This approach was used to gauge the level of comfort perceived by the user based on the level of crowding experienced.

5. **The door position** was recorded to give an indication of the position on the platform where the gap was measured.

6. The **overall perception of the service** was recorded using a Likert scale, with 5 being the poorest service perception and 1 being the best service perception (based on the user's judgment).

7. Lastly, the **availability of accessibility features** for PRMs into the station was also recorded (i.e. functioning lifts, ramps etc.)

Figure 3-22 to 3-24 present examples of the gap measurements obtained at Walker street station, Loftus station and Pretoria station.

Figure 3-22: Walker street station

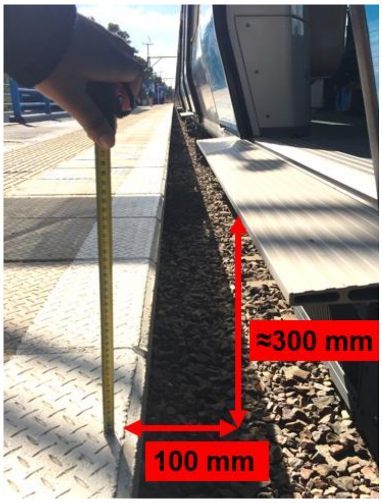

Figure 3-23: Loftus station

Figure 3-24: Pretoria station

A summary of the largest gap sizes at each station along the corridor is presented in Table 3-5. Based on the definition of high risk and low risk stations (Wahl, 2010), it was observed that 73 % of the stations along the Pretoria-Pienaarspoort corridor fell within the medium to high risk category.

Table 3-5: Worst case gap sizes measured at each station along the Pretoria-Pienaarspoort corridor

Station	Risk classification based on vertical gap size	Horizontal gap size (mm)	Vertical gap size (mm)
Pretoria	High	0	340
Loftus	High	100	300
Koedoespoort	High	0	395
Silverton	High	50	320
Watloo	High	25	310
Pienaarspoort	High	0	315
Eerste Fabrieke	High	0	310
Walker street	Medium	300	10
Hartbeespruit	Medium	30	240
Eersterust	Medium	50	260
Denneboom	Medium	0	260
Mears street	Low	130	30
Mamelodi Gardens	Low	140	5
Devenish street	Low	130	28
Rissik	Low	100	50

The complete results from the experiment are presented in Appendix A. The results revealed that none of the stations along the Pretoria-Pienaarspoort corridor are accessible to PRMs (especially wheelchair users). The gap sizes were either too large for a PRM to navigate unassisted or access to the station platform was not possible because the elevators were not operational and there were no wheelchair ramps available. The results also revealed that there was little to no security presence in stations and platforms, therefore implying crowd control and law enforcement was not administered during peak hours. Observation surveys were therefore carried out to observe operational conditions at chosen station during peak hour to reconcile the results obtained from the above informal study.

3.3 PASSENGER BEHAVIOUR OBSERVATION SURVEY

Observation surveys were conducted on passengers at three stations along the corridor, namely Rissik, Walker street and Loftus station. These three stations were classified as low risk, medium risk and high risk, respectively. These risk classifications were based on the gap sizes measured at each of the stations. According to a study conducted by Moug and Coxon (2013), the horizontal gap at a station may not have a higher effect on passenger safety than the vetical gap size, but a large horizontal gap increases the chances of passengers falling through the gap (especially special needs passengers). Table 3-6 presents gap size measurements obtained and the corresponding risk classification of the stations.

Table 3-6: Gap sizes measured at Rissik, Walker street and Loftus station

Station	Risk classification	Horizontal gap size (mm)	Vertical gap size (mm)
Rissik	Low	100	50
Walker street	Medium	300	10
Loftus	High	100	300

The Metrorail service shares the rail network with freight and heavy haul trains, therefore, according to the RSR’s standard on railway stations (2015), a clearance of 100 mm (horizontal or vertical) is required at station platforms to ensure the safe travel of freight and heavy haul trains through the station. Based on this standard, Rissik station was within the limits specified and was therefore classified as low-risk and served as the control station of the experiment.

Based on the study carried out by the Wahl (2014), it was observed that a vertical gap size had a larger influence on passenger safety at the PTI compared to the horizontal gap size. Therefore, Walker street was classified as a medium risk station and consequently, Loftus was classified as a high-risk station.

Using a data collection team of five observers and with the aid of two digital cameras (to compensate for human error), data was collected on the behaviour of passengers boarding and alighting the train. The observation surveys were carried out during the morning (06h01 to 08h00) to and afternoon (16h01 to 20h00) peak hours because it was recorded in the

State of Safety Report (RSR, 2017) that the majority of PTI incidents occurred during peak hours due to the high volume of passengers traveling at that time.

A gap injury analysis carried out by the New Jersey Institute of Technology (2009) determined that 69 % of gap injuries involved women. For this reason, the observation study sought to distinguish the difference between the behaviour of female and male passengers. The proposed methodology in carrying out the observation survey was to gather data on boarding and alighting passengers at three to four doors of the arriving trains at the identified stations. The data collection forms used in the study are attached in Appendix A. The following data was collected:

1. The **number and gender of passengers** boarding and alighting the train;
2. The **number of passengers looking down** as they boarded or alighted the train;
3. The number of **passengers carrying luggage** while crossing the PTI;
4. The number of **passengers using a cell phone** while crossing the PTI;
5. Identification of any other factors or behavioural characteristics of commuters that were observed to influence the commuter’s risk of experiencing a PTI occurrence.

Figure 3-25 to 3.27 present the observation survey experimental set-up at the three different stations.

Figure 3-25: Observation survey at Rissik station

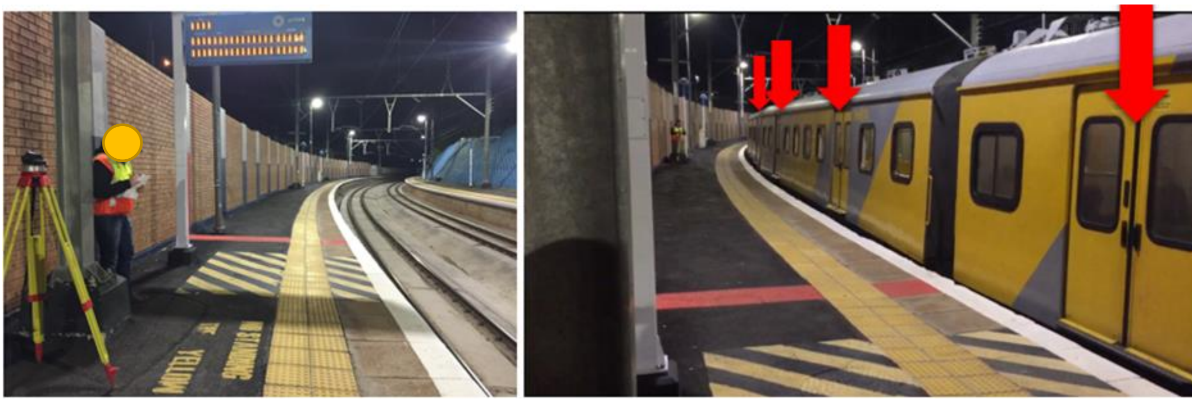

Figure 3-26: Observation survey at Walker street station

Figure 3-27: Observation survey at Loftus station

Passenger movement was recorded at 4 doors (minimum) at each station. Data collection was carried out over a period of 5 working days (Monday to Friday). The main results and conclusions from the observation surveys are presented in subsequent sections.

3.3.1 Main results and conclusions from the passenger observation surveys

The results from the observation survey included the percentage gender distribution of the commuter population and the variation in commuter volumes on each day during the five days of testing. The results also gave an indication of the passenger group that displayed behaviour that could result in a PTI occurrence. The number of special needs passengers observed was also recorded and thus gave an indication of the level of accessibility into the train for this passenger group.

The observation surveys were carried out according to the study that was conducted by the NJDOT (2009). This study sought to distinguish the difference in behaviour between male and female commuters. This distinction was carried out to determine the PTI occurrence risk profile of the different gender groups. The results of the percentage gender distribution at the three stations are presented in Figure 3-28. In addition, the percentage gender distribution of passengers that looked down at the gap while crossing over the PTI are presented in Figure 3-29.

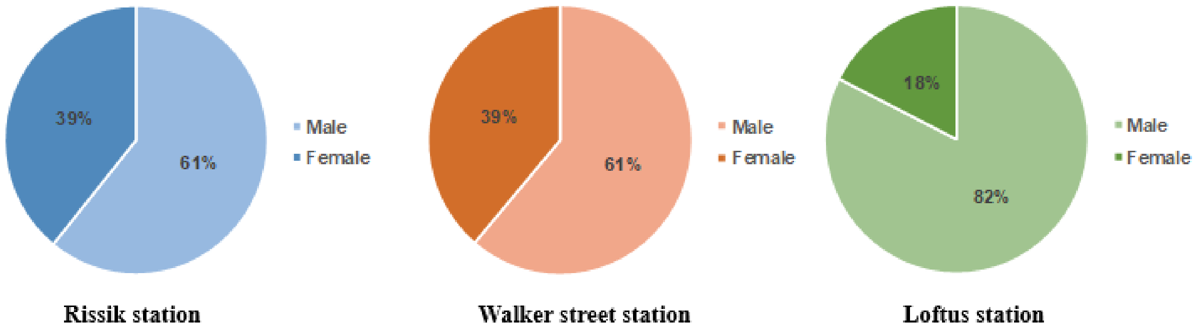

Figure 3-28: Percentage gender distribution at Rissik, Walker street and Loftus station

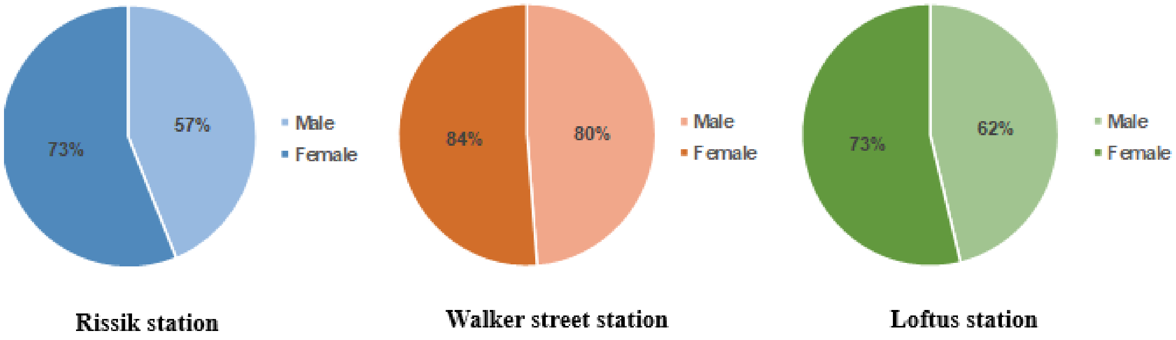

Figure 3-29: Percentage gender distribution of passengers looking down at gap while crossing PTI.

The results presented in Figure 3-28 outline that (on average) the total commuter population consisted of 68 % male commuters and 32 % female commuters. The results in Figure 3-29 further outline that on average a larger percentage of the female commuter population looked down at the gap while crossing over the PTI compared to the male commuter population. This result indicated that the female commuter population had a higher awareness of the gap than the male population, thus implying that the female commuter would (potentially) be less likely to experience a PTI occurrence assuming that being aware of the gap reduced an individual's chances of experiencing a PTI occurrence.

The variation in passenger volumes throughout the five days was observed and the results presented in Figure 3-30. The passenger volume on each day is presented as a percentage of the total passenger volume observed at each station throughout the testing period. In addition, the average of the percentages obtained in Figure 3-30 are presented in Figure 3-31 to represent the average total passenger volume variation on each day. Figure 3-31 outlines that on average the highest volume of passengers was observed on a Friday.

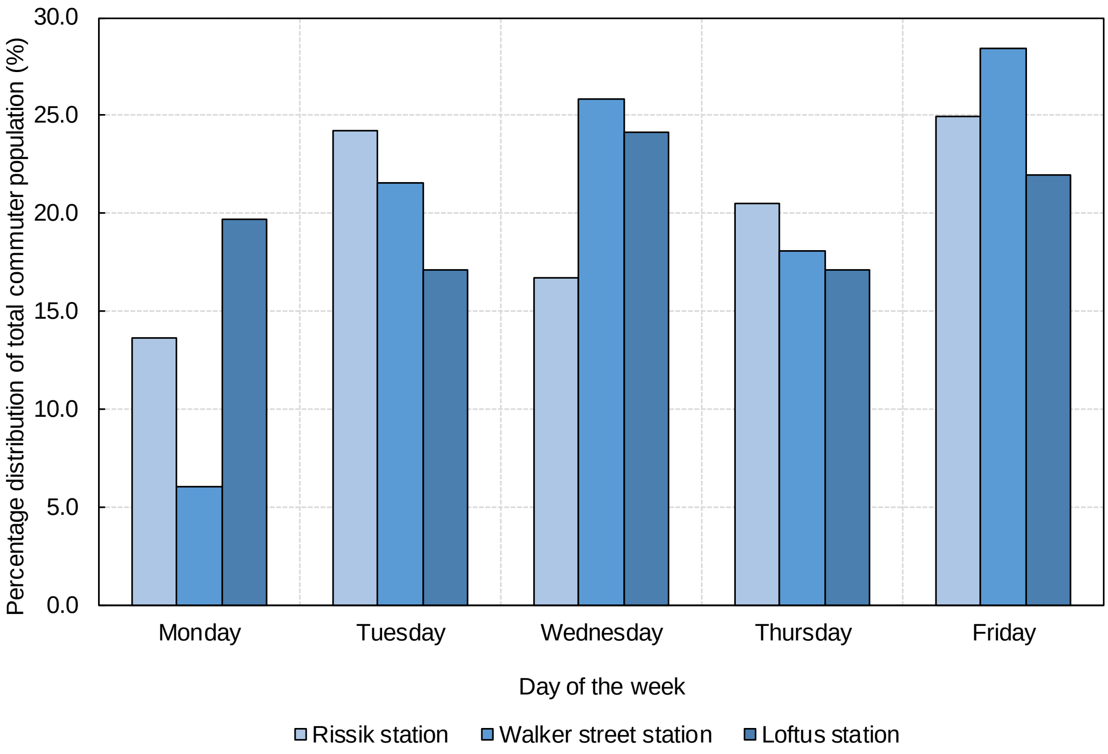

Figure 3-30: Change in passenger volume throughout the week at each station

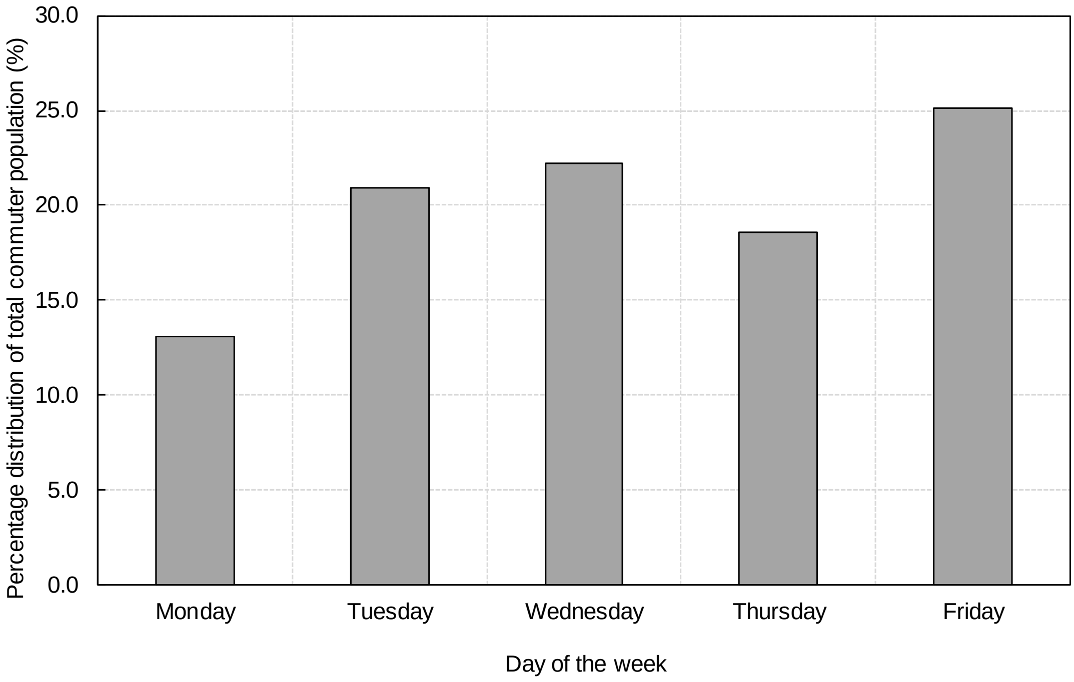

Figure 3-31: Average change in passenger volume throughout the week

According to the State of Safety Report (RSR, 2017), one of the major factors that influenced the risk of a PTI occurrence was the level of overcrowding. Therefore, the results presented in Figure 3-31 imply that Friday may be the "riskiest" day of the week with regard to PTI occurrences because of the high passenger volumes observed on this day compared to the other days of the week.

According to the action agenda issued by the Minister of Transport and the Department of Transport (DoT) in 1999 (Moving South Africa), the goal was to develop a transport system that would meet the requirements of all passengers, especially passengers with special needs, by 2020. Part of the research involved measuring the level of accessibility into the train by special needs passengers as a means of commenting on the progress of the implementation of the action agenda. The total commuter population was considered and categorised according to non-special needs passengers and special needs passengers. Special needs passengers were defined as the elderly (who were distinguished through a process of elimination, meaning, individuals who did not appear to be children, young adults or middle-aged individuals), children/scholars, People with Reduced Mobility (PRMs) and visibly pregnant women. The results are presented in Figure 3-32.

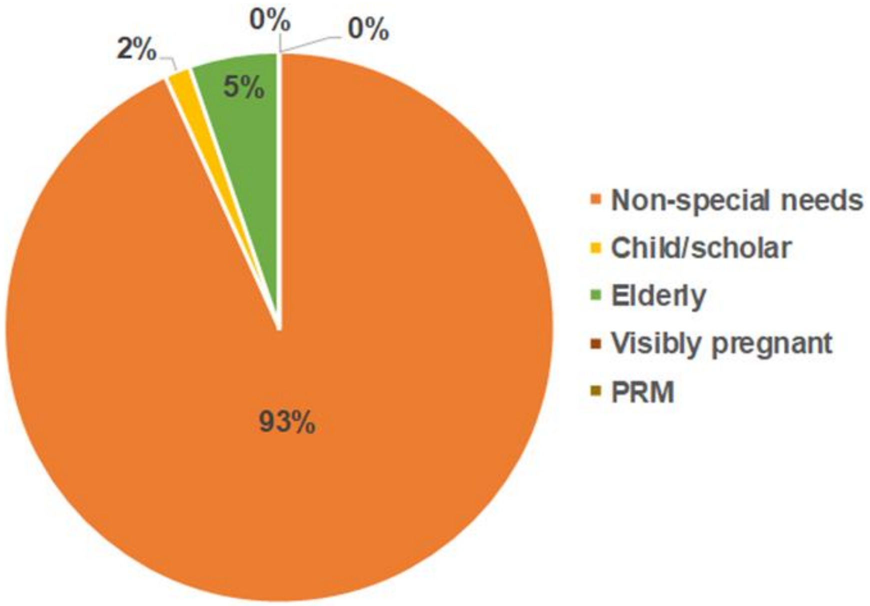

Figure 3-32: Percentage distribution of the different passenger groups observed

Visibly pregnant women and PRMs were not observed throughout any of the observation surveys conducted. Only 7 % of the total commuter population consisted of elderly people and children/scholars. The percentages obtained raised a question about the level of accessibility of the Metrorail service to special needs passengers. It also raised the question of what alternative modes special needs passengers opted to use instead of the Metrorail service. In addition, what were the possible socioeconomic implications of this exclusion of special needs passengers from the affordable Metrorail service?

During the observation surveys, informal unstructured interviews were carried out with passengers at each station. Passengers recounted incidents of theft, vandalism and service unreliability. There was also common mention of the difficulty of navigating the gap at the PTI when boarding and alighting the train. For these reasons, the Metrorail service in most accounts was not the preferred means of transport and most commuters were captive to the service because of its low price. This finding further illustrates that inaccessibility into the service by special needs passengers translated into paying the "penalty" of a higher priced mode of transport. These commuter accounts share (not discounting the unreliability and lack of safety of the service) may also explain the reason passenger rail holds only 15 % of the commuter market in Gauteng even though in most instances the Metrorail service is the most affordable mode of transport (Mondliwa, 2019).

From the observation surveys conducted it was concluded that the major stress factors that influenced the risk of PTI occurrences at train station platforms were:

1. **Level of overcrowding**: The increase in passenger density on the platform and in the train resulted in more pushing and shoving as passengers crossed over the PTI. This behaviour increased the risk of a passenger tripping over the PTI or falling into the gap, especially in the presence of a large gap.

2. **Train design**: The train doors of the older Class 5M2A ("yellow") trains were faulty, resulting in manual operation and in some instances the train door would stick and remain open as the override functionality of the doors was outdated. For this reason, passengers would attempt to board or alight the train while the train was still moving.

3. **Platform design:** The large gap size was observed to be an obstacle especially for the elderly population and passengers carrying luggage or travelling with children.

The three highlighted factors correlated with the difficulty items outlined in the study conducted by Cheng (2010). The results obtained from the study revealed that difficulty items such as overcrowding and the gap between the train and the platform were part of the list of the main causes of passenger anxiety associated with train travel. Performing regular observation surveys on a larger scale would provide invaluable data on passenger behaviour. The results from the surveys can be applied in the determination of the most effective user-centric interventions that could improve train services and result in rail becoming the preferred mode of transport.

3.4 RN34 METHOD

The final test carried out was the measurement of the passenger flow rate through the train doors. The flow rate was determined by analysing the video footage captured during the observation surveys carried out. The variation of the flow rate at each station served as an indicator of the possible factors that influenced passenger movement across the PTI.

The passenger flow rate was determined using the RN34 method (RRL, 1963) and was defined as the number of passengers crossing over the PTI per second. The results presented in this section were based on the method presented in Figure 2-7 with the modification of using a scatter plot instead of a bar graph. Rissik station was considered the control in this experiment because the gap size between the train and the platform adhered to the Regulator’s standard (RSR, 2014). A comparison was conducted between the flow rate at Rissik station and the other two stations to determine the effects of factors such as the gap size, train design and passenger density (in the train and on the platform). The results of the comparison are presented in the subsequent figures. Figure 3-33 presents the passenger flow rates observed at Rissik station and Loftus station during one boarding and alighting cycle in the afternoon peak (between 16h30 and 17h00).

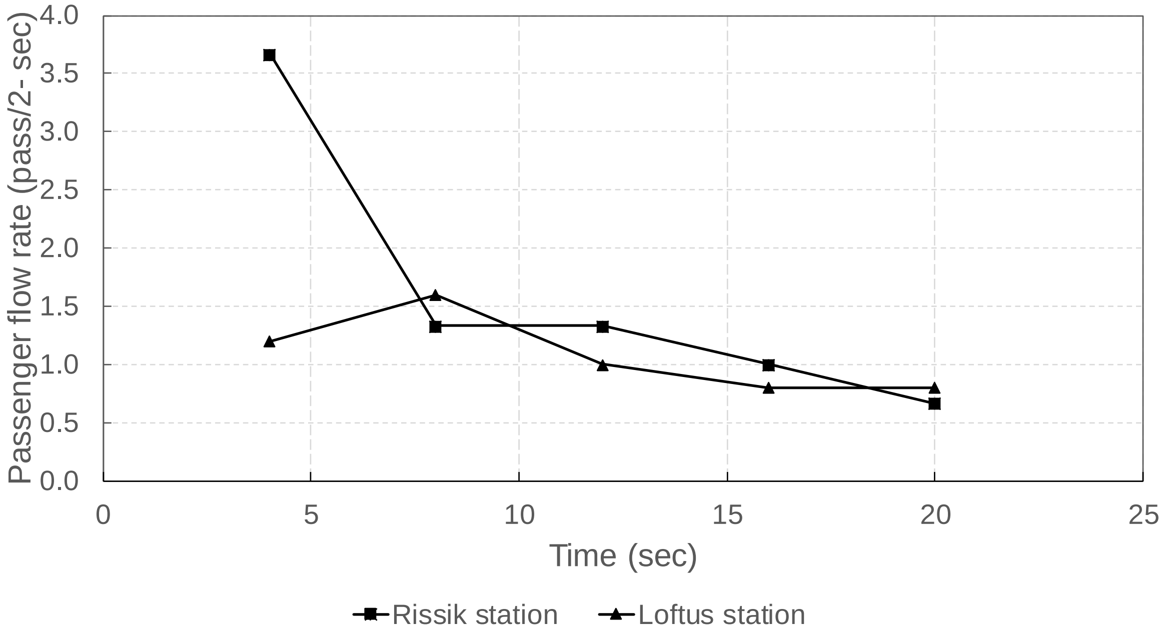

Figure 3-33: Passenger flow rate at Rissik and Loftus station during the afternoon peak

The average flow rate at Rissik station was 8 % higher than the average flow rate at Loftus station. This difference in flow rate was attributed to the vertical gap size, where the vertical gap size at Loftus station was measured to be approximately 6 times larger than the gap size at Rissik station.

The difference in flow rate was also attributed to the different train types that were operating during the time of the observation surveys. At Loftus station the “yellow” trains were in operation and in 40 % of the boarding and alighting cycles recorded, the train doors either operated manually or were stuck causing passengers to rush to the next door. This resulted in bottlenecks forming at the train doors, thus restricting the passenger flow in and out of the train.

Considering the passenger flow observed at Rissik station, the flow rate gradually decreased with time because of the high density of passengers in the train. The passenger flow rate was approximately 63 % higher at the beginning of the boarding cycle because of the automatic door function, but this flow decreased as the train started to fill up.

Figure 3-34 presents the flow rates measured at Rissik station and Walker street station during one boarding and alighting cycle in the morning peak (between 07h00 and 08h30).

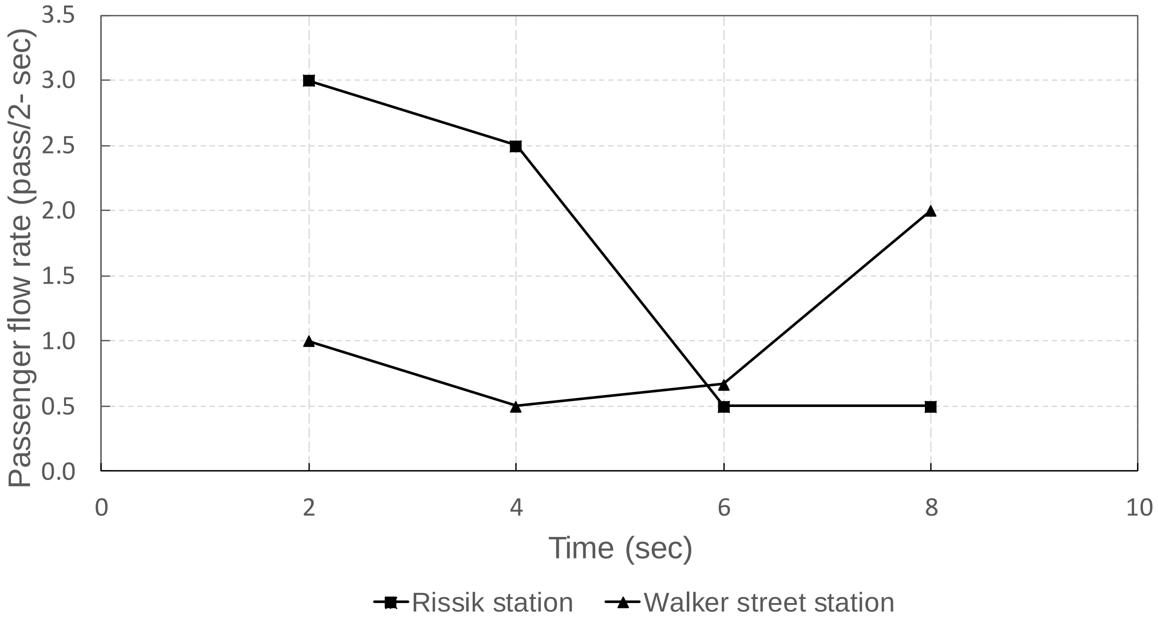

Figure 3-34: Passenger flow rate at Rissik and Walker street station during the morning peak

The average flow rate at Rissik station was 11 % higher than the average flow rate at Walker street station. This difference in flow rate was attributed to the volume of passengers; at Walker street station the passenger volume was lower than the passenger volume at Rissik station during the morning peak hours. For this reason, passengers were not rushing to board the train and the low volume of passengers meant that less passengers crossed over the PTI per second.

The difference in flow rate was also attributed to the horizontal gap size and train design (specifically the train door functionality) at Walker street station. The horizontal gap size at Walker street station was measured to be 3 times larger than the horizonal gap size at Rissik station. The presence of a 300 mm horizontal gap size in addition to manually operating train doors (and sometimes “stuck” doors) resulted in a lower flow rate at Walker street station because of the delayed entrance and/or exit into or out of the train.

When considering the passenger flow observed at Rissik station, a similar trend in the flow rate is observed in Figure 3-34 to the trend observed in Figure 3-33. The gradual decrease in the flow was again attributed to passenger density, however, in this instance it was passenger density on the platform. Passengers moved freely through the doors when the alighting cycle began, but the flow gradually decreased as more passengers filled the platform. Considering Walker street station, the passenger flow gradually increased with time. This observation was attributed to the fact that the train was empty when the boarding cycle began, therefore passenger flow was not obstructed because there was enough space for passengers to move freely into the train.

From these results presented, it was concluded that the most significant factors affecting passenger flow at a Metrorail station were the gap size, train design (manual versus automatic doors) and the passenger density in the train and/or on the platform.

The saturated passenger flow rates at each station were also calculated and are presented in Table 3-7. The saturation flow was defined as the maximum flow rate observed during a saturated boarding and alighting period free from "lost time" effects caused by delayed entry and/or exit into or out of the train (Fernández et al., 2015). The flow rate was further normalised based on the door width (1.65 m) and the number of doors observed (in this case, the door at which maximum flow was observed during the observation surveys).

Table 3-7: Passenger saturation flows (passengers per second)

	Morning Peak		Afternoon Peak	
	Rissik	Walker street	Rissik	Loftus
Pass/sec	1.95	1.92	2.71	2.40
Pass/sec/m	1.18	1.16	1.64	1.45

The passenger saturation flow at Rissik station was observed to be higher than the saturation flows determined at both Walker street station (morning peak) and Loftus station (afternoon peak) by 1.7% and 13% respectively. In a study conducted by Fernández et al. (2015) to determine passenger saturation flow values for design application, the results of the study concluded that the range of 0.9 to 2.0 passengers per second was a sufficient representation of "in-station" saturation flow rates and the results in Table 3-7 further validate this assumption.

3.5 DISCUSSION

The field tests and experiments carried out during the research were aimed at quantifying the effects of platform design, train design and human factors on passenger behaviour and how these factors affect the level of safety of passengers at station platforms. Testing the effects of the platform design was related specifically to testing the effects of the gap between the train and the platform. According to the Regulator's standard on railway stations (2015), a clearance of 100 mm was required to ensure the safe travel of all trains through stations. However, according to the RSSB (2015), a gap size (vertical and/or horizontal) of 100 mm is considered unsafe for passengers and restricts access into the train for PRMs. It is suggested that a gap size of 75 mm (horizontal) and ± 50 mm (vertical) would ensure passenger safety and unassisted access into the train by PRMs. To reduce the effects of the 100 mm gap clearance, strict monitoring of the gap would be required to ensure the gap is constantly within the stipulated standard, thus motivating the development of a distance measuring prototype that could serve this purpose.

The prototype consisted of an Arduino Uno microprocessor coupled with two distance measuring sensors (one ultrasonic sensor and one laser sensor). During the calibration and testing phase, it was observed that the sensor readings were sufficiently accurate when measurements were carried out under stationary conditions, however, when the sensors were tested while mounted on a moving vehicle, the readings became erratic and considerable data cleaning was required. Data cleaning techniques such as translation and application of the moving average method were examples of techniques applied during data analysis. In addition, the ultrasonic sensor readings were sensitive to surface texture and the measurements drifted by approximately 3 % as the testing duration lengthened. The laser sensor was observed to be sensitive to surface reflectance and on average was 5 % less accurate than the ultrasonic sensor. However, no drift was observed with the readings obtained using the laser sensor. Simultaneously implementing the ultrasonic and laser sensors was recommended to create redundancy as a method of cross-checking the accuracy of the readings obtained.

Further research into the successful implementation of continuous gap measurements would provide data on the rate at which the gap size at the PTI changes and in the long term would provide an indication of the deterioration rate of the track structure at stations. Using this data, interventions or regular maintenance can be implemented at stations to ensure the gap clearance is constantly within the stipulated standard, thus reducing the effects of the gap on passenger safety.

Testing the effects of human behaviour on the level of safety of passengers at station platforms involved carrying out observation surveys at Rissik, Loftus and Walker street station. The observation surveys revealed that (on average) the total commuter population observed consisted of 68 % male passengers and 32 % female passengers. This disparity in the percentage gender distribution is assumed to be representative of the gender distribution throughout the Metrorail service (RSR, 2017). In addition, only 7 % of the total passenger population observed were classified as special needs passengers, thus indicating that the Metrorail service was not sufficiently accessible to the special needs commuter population. Informal interviews conducted at the stations revealed that this disparity in the different passenger categories could be attributed to the prevalent incidents of theft, vandalism, service unreliability, and inaccessibility into the train due to the large gap sizes at the PTI.

From the observation surveys it was concluded that the major stress factors that influenced the risk of PTI occurrences at station platforms were overcrowding levels, train design and platform design. Expanding further on this conclusion, the results obtained from measuring passenger flow during the observation surveys revealed that similarly, passenger flow is affected by the gap size at the PTI, train door functioning and passenger density in the train and on the platform. The abovementioned factors therefore affect passenger safety and operation efficiency.

However, the factors highlighted in the study cannot be considered in isolation. Other underlying factors such as train delays and the lack of information dissemination directly influenced the behaviour of passengers while they waited for the train to arrive. Train delays and minimal information dissemination resulted in increased passenger agitation that manifested into more pushing and shoving when passengers were crossing over the PTI when the train arrived. Therefore, these two factors cannot be ignored when deciding the most effective interventions to reduce the risk of PTI occurrences.

The conclusions drawn from the study were not without limitations, for example, limited time access to the stations meant that only peak hour data was collected, therefore limiting the conclusions drawn to only peak hour situations. Recording off peak passenger behaviour would provide an indication of travel patterns at the different stations and thus add to the insight needed in the implementation of the most effective interventions to improve passenger safety and service efficiency. Conducting regular passenger observation surveys on a larger scale and for longer time periods is recommended to effectively determine the optimum user-centric interventions that would lead to the most significant improvements in the service. These improvements could (in the long-term) result in increased train ridership and rail becoming the preferred mode of transport in Gauteng.

4 DATA ANALYSIS

4.1 INTRODUCTION

This section outlines the data analysis carried out in the calculation of the risk of PTI occurrences along the Pretoria-Pienaarspoort corridor. The data analysis process involved investigating the statistical relationships that existed between the factors that were identified to influence PTI occurrences and the frequency of occurrences recorded. The data used in the analysis was provided by the Railway Safety Regulator (RSR) for the 2014 – 2017 reporting periods.

The data analysis process also involved the risk classification of the different stations along the Pretoria-Pienaarspoort corridor. The risk classification was based on the relative risk calculated for each station and indicated the stations that required priority maintenance or interventions to improve passenger safety on the platform.

The main factors considered in the data analysis process were the platform design, human behaviour and train design. The influence of these factors on the risk of PTI occurrences was quantified through regression analyses and the development of estimation functions. The data analysis process is summarised as follows:

1. **Simple Linear regression analyses** of the different factors that were assumed to have an influence on the risk of a PTI occurrence. These factors being: passenger volume in a station, the vertical gap size at the PTI, the gender of the passengers, the day of the week and the time of day. This step was carried out to determine the most significant variables to consider in further data analysis. The abovementioned factors were identified from the observation surveys carried out and the causes identified from historical data provided by the RSR.

2. **Multiple Linear regression analysis** was carried out to approximate a function that could predict the expected number of PTI occurrences per 10 000 passengers at each station along the corridor. The variables chosen in the regression analysis were based on the results obtained from the multiple linear regression analyses carried out. These results further indicated the highest risk stations along the corridor based on the vertical gap size and the population size at each station.

3. **Fault tree analysis** to determine the relative risk of PTI occurrences at each station along the corridor. The relative risk was calculated based on the day of the week, the time of day, gender, and each station served as proxy for the vertical gap size. This calculation was

carried out to quantify the level of influence of other underlying factors on the risk of PTI occurrences that could not be adequately presented through applying linear regression.

4. **Multinomial Logistic regression** was carried out to estimate a function that could predict the probability of the type of occurrence at each station along the corridor. From the data obtained from the RSR it was determined that four types of PTI occurrences were prevalent along the corridor, namely (i) occurrences caused by a moving train, (ii) staff riding (train surfing), (iii) overcrowding and (iv) other (unspecified causes).

In addition to the results obtained from the steps carried out in the data analysis, the current rating of safety with regard to PTI occurrences was established based on the measure of Fatalities and Weighted injuries (FWIs). FWIs was a method applied to quantify the consequences of PTI occurrences introduced in the State of Safety report published by the RSR (2018).

From the results obtained in the data analysis, the effects of PTI occurrences on the capacity and efficiency of the Metrorail service was hypothesised by quantifying the probable loss in capacity along the corridor if a PTI incident were to occur. This finding provided an indication of the probable cost of PTI occurrences to the service provider in terms of corridor capacity losses.

4.2 RSR DATA (2014 to 2017)

The RSR records railway safety incidents every year as part of its mandate since 2008. These safety records are then compiled into an annual State of Safety report that is published and serves as an indicator of the level of safety of freight, heavy haul and passenger rail services in South Africa. The focus of this research was on the PTI occurrences recorded during the 2014 to 2017 reporting period. The data analysis involved categorising the PTI occurrences into the following groups:

1. Gender classification

2. Causes of PTI occurrences

3. Time of PTI occurrences (time of day, day of the week and year of occurrence)

The total number of PTI occurrences recorded nationwide in the 4 reporting periods are summarised in Table 4-1.

Table 4-1: Summary of PTI occurrences recorded between 2014 and 2017 nationwide that resulted in an injury or fatality

	2014	2015	2016	2017	Totals
Fatalities	4	9	12	8	33
Injuries	400	643	561	531	2135

These results are further represented graphically in Figure 4-1 distinguishing occurrences that involved male or female passengers. It should be noted that all fatalities recorded were male passengers and were not included in the data presented in Figure 4-1.

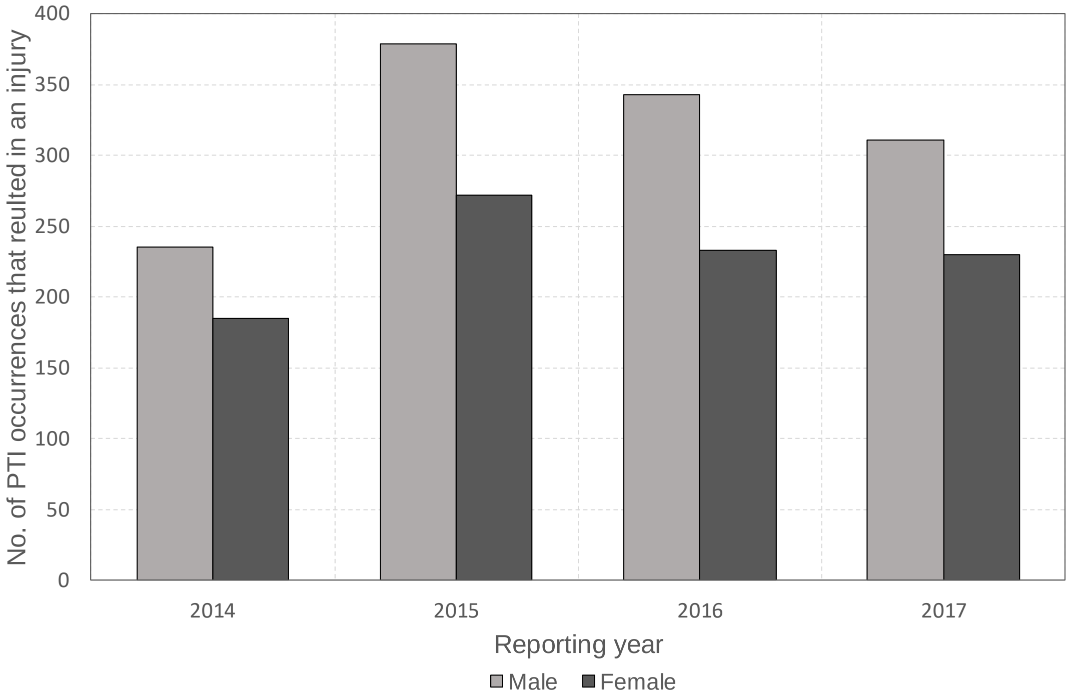

Figure 4-1: Total PTI occurrences that resulted in an injury (excluding fatalities) recorded between 2014 and 2017 nationwide

The results presented in Figure 4-1 show that 2015 had the highest number of PTI occurrences recorded. Figure 4-1 further highlights that male passengers accounted for approximately 37% more occurrences on average than female passengers during the 4-year reporting period. This observation implied that gender could be a possible factor to consider when determining the passenger profile that would be at the highest risk of experiencing a PTI occurrence. The data was further categorised into the time of day and the day of the week the occurrences took place (Figure 4-2 and 4-3).

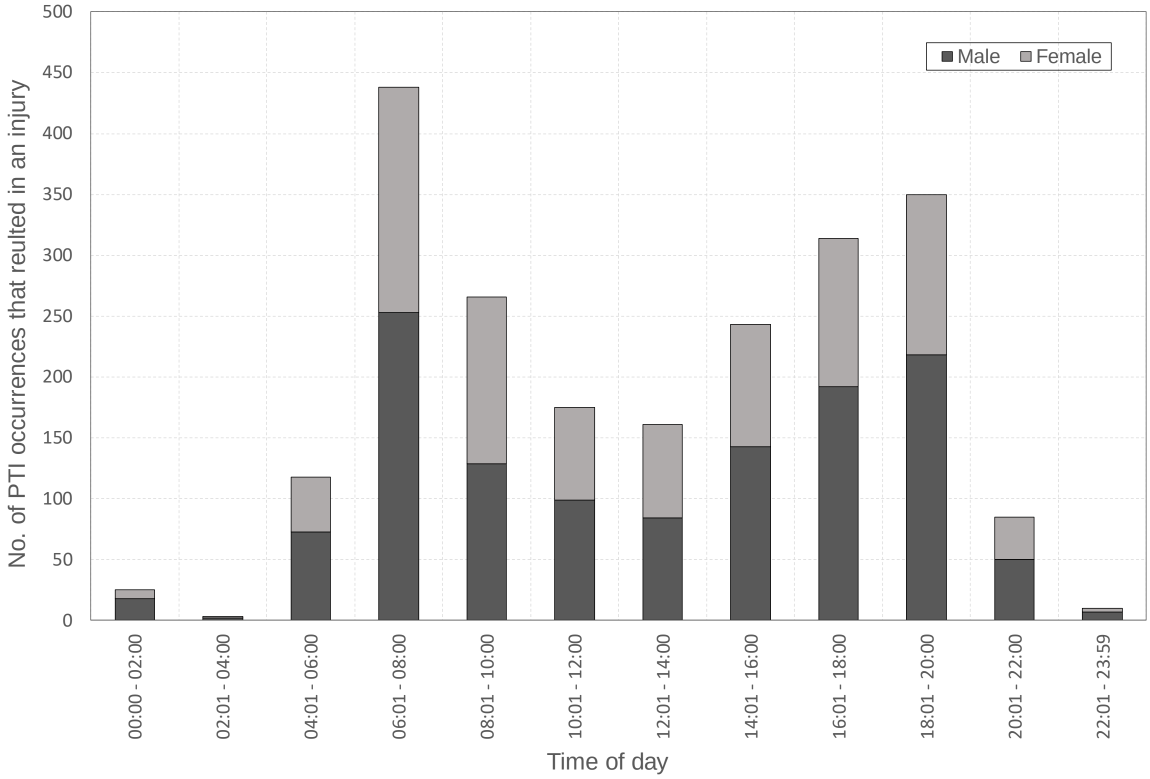

Figure 4-2: Total number of PTI occurrences recorded from 2014 to 2017 according to the time of day

Considering the results presented in Figure 4-2, the highest numbers of PTI occurrences were recorded during the morning (06h01 - 08h00) and afternoon/evening (16h01 - 20h00) peak hours. During the peak hours there is a higher volume of passengers traveling at the same time, thus leading to overcrowding on the platform and in the train. This leads to pushing and shoving as passengers are crossing over the PTI to board or alight the train. This behaviour increases the chances of a passenger falling on the platform or tripping over the gap and/or falling into the gap.

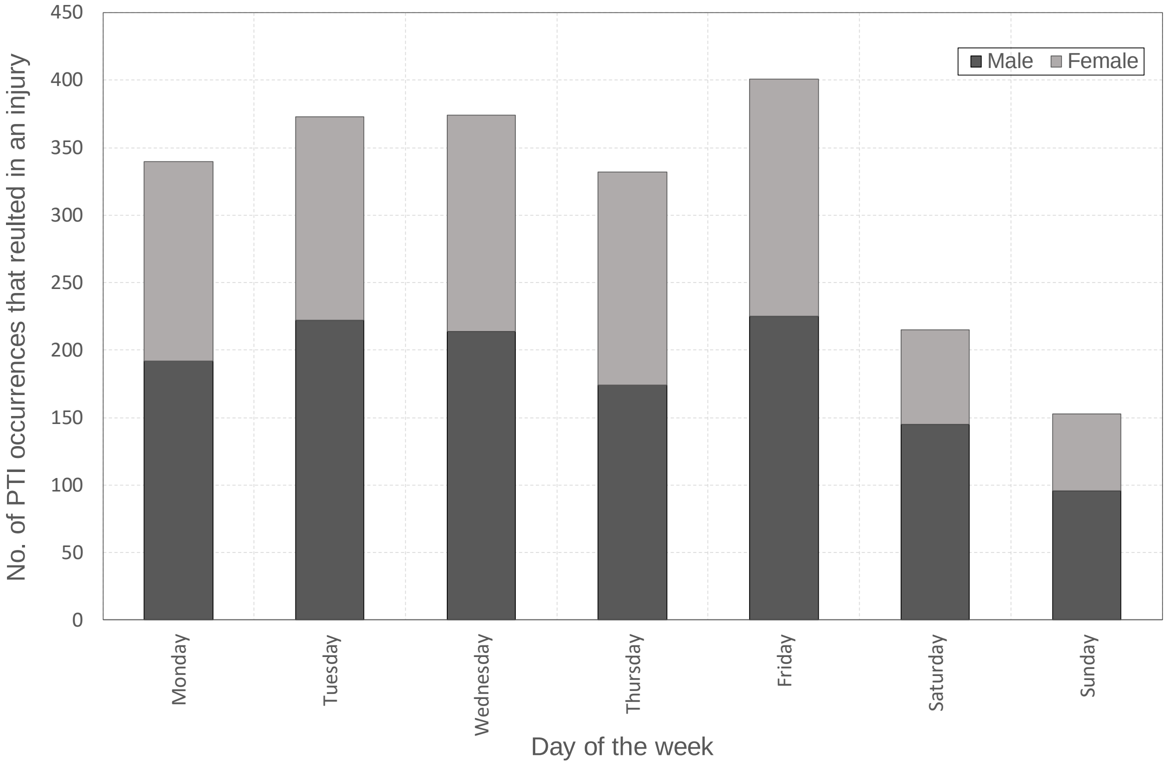

Figure 4-3: Total number of PTI occurrences recorded from 2014 to 2017 according to the day of the week

A similar explanation applies to the trend observed in Figure 4-3; on a Friday there is a higher volume of passengers traveling than on the other days of the week thereby resulting in overcrowding at the station platforms. This observation implies that passenger volumes on the platform and in the train influences the risk of a PTI occurrence, namely: overcrowding is a factor that influences passenger safety on the station platform. This conclusion relates perfectly to the conclusion drawn from the observation surveys carried out during the research, where overcrowding was recorded as one of the main stress factors that could cause a PTI occurrence and Friday had the highest recorded level of crowding during the surveys (Refer to Section 3.3).

The recorded occurrences were further categorised into the three main causes identified from the incident descriptions provided. The three main causes described were: firstly, incidents caused as a result of passengers boarding or alighting the train while it was moving; secondly, passengers having been pushed or shoved during the boarding and alighting cycle that resulted in a PTI occurrence and; lastly, staff riding, where passengers were hanging outside the train and experienced a PTI occurrence as the train was arriving at the station or leaving the station.

These three main causes were summarised as “Moving train”, “Overcrowding” and “Staff riding”. Table 4-2 presents the percentage contribution of each cause.

Table 4-2: Percentage contribution of the causes identified from occurrence descriptions recorded (RSR)

Cause identified	Moving Train	Overcrowding	Staff riding	Other (unspecified)
	86.2%	7.2%	22.6%	5.0%

Occurrences caused by the “Moving train” accounted for the largest percentage of PTI incidents followed by “Staff riding”, “Overcrowding” and lastly, “Other (unspecified)” causes. All occurrences with no indication of the cause in the incident description were classified as “Other (unspecified)”. In such instances, victims were recorded to have fallen onto the platform while getting on or off the train without any indication if the occurrence was a result of pushing or shoving, or if the train was moving.

The main aspect of train design observed to have had the highest influence on passenger safety was the train doors. This observation was based on the results from the passenger observation surveys carried out. The doors of the older Class 5M2A (“yellow”) trains were observed to be faulty and this resulted in the train doors remaining open while the train was traveling, thus creating the opportunity for passengers to attempt to board or alight the train while it was still moving.

The second most prevalent cause identified was staff riding, which involves passengers hanging outside the train as it is traveling. This method of travel is sometimes used as a means of ticket evasion or due to severe overcrowding in the trains. Traveling in this position drastically increases the likelihood of passengers falling off the train as it is arriving at or departing from the station.

Lastly, overcrowding was identified as the third most prevalent cause of PTI occurrences. Overcrowding results in pushing and shoving on the platform as passengers cross over the PTI, thus increasing the likelihood of a passenger falling onto the platform or tripping over the gap and/or falling into the gap.

It was identified that 20% of the PTI occurrences recorded were as a result of a combination of the causes identified in Table 4-2. The data was presented in the form of a Venn diagram to depict the percentages of PTI occurrences as a result of a combination of the outlined causes (see Figure 4-4).

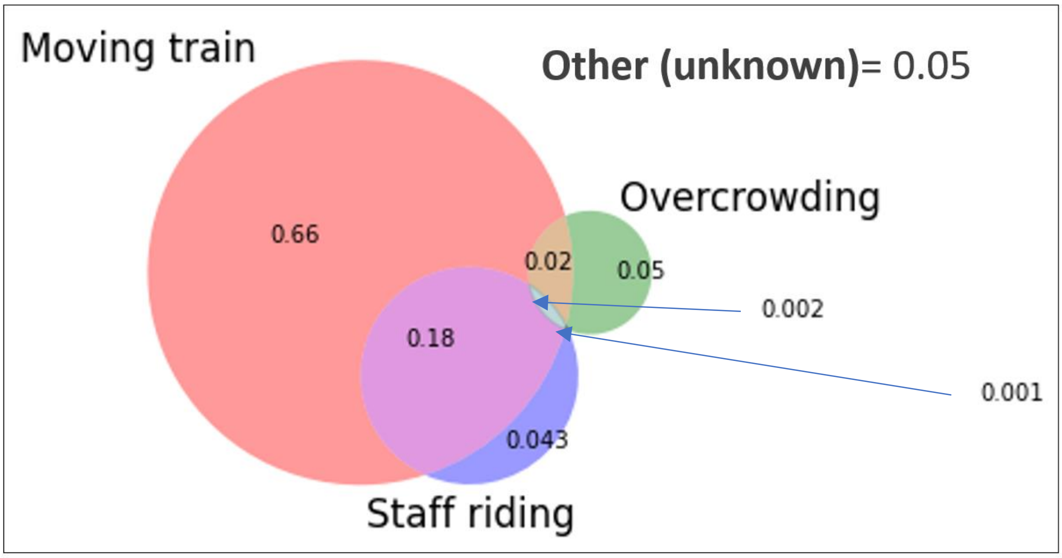

Figure 4-4: Percentage contribution of the identified causes of PTI occurrences recorded nationwide (RSR, 2014-2017)

The moving train was the main cause that accounted for the largest percentage of PTI occurrences recorded. The second largest percentage was the combination of occurrences caused by a moving train and staff riding. This observation implies that if the train design was to be modified to ensure that the doors functioned automatically, and the design modified to ensure that there was no opportunity for passengers to staff ride, approximately 84% of the PTI occurrences recorded would have been eliminated. This finding therefore reemphasizes that train design is one of the main factors to be considered when investigating PTI occurrences.

It should be noted that the gap size between the train and the platform was not explicitly specified in the occurrence description records, however the presence of a large gap influences the process of boarding and alighting as indicated by the results obtained from measuring passenger flow (refer to Section 3.4). Therefore, it can be assumed that the risk of experiencing a PTI occurrence increased in the presence of a large gap, especially if the train was moving and/or if the platform was overcrowded. For this reason, the influence of the gap size was presented by the occurrences recorded as “Other”.

The PTI occurrence data was further presented as a measure of the Fatalities and Weighted injuries (FWIs). FWIs equate injuries as 0.1 times a fatality, thus providing a standard measure of the consequences of occurrences as the fatalities caused (State of Safety Report, 2018). The equation is presented as:

$$\boldsymbol{FWIs = No.of\ Fatalities + 0.1 \times No.of\ Injuries} \qquad \textbf{Eq 4-1}$$

Figure 4-5 presents the FWIs calculated based on the causes identified over the four-year reporting period (2014-2017).

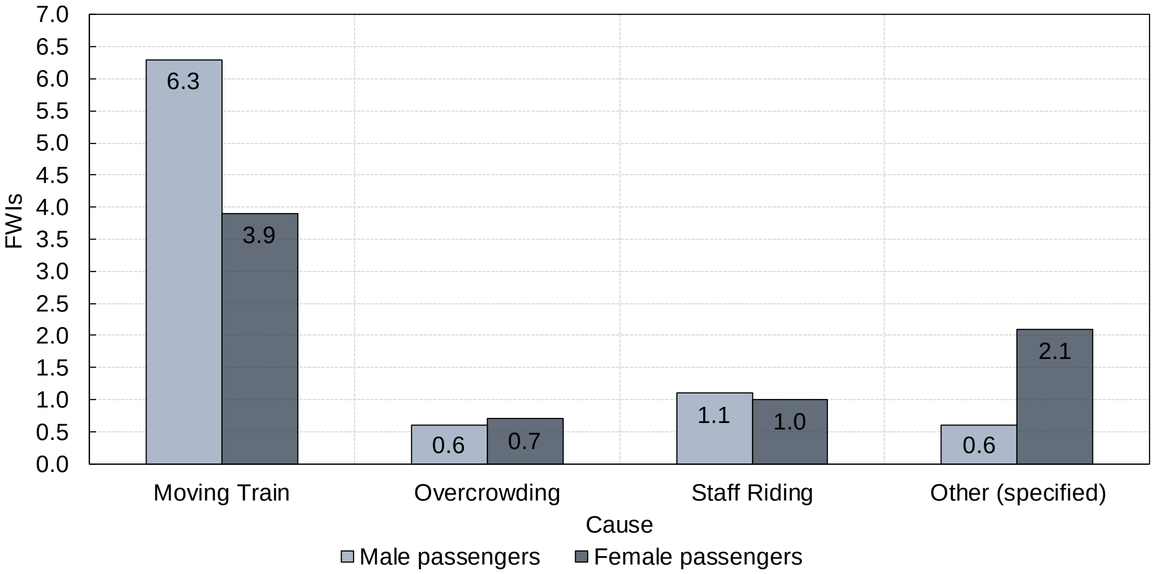

Figure 4-5: FWIs calculated from the PTI occurrences recorded nationwide according to the causes identified (RSR, 2014 2017)

The results show that occurrences as a result of a Moving train accounted for a total of at least 10.2 fatalities over the four-year period. On average, it was observed that male passengers accounted for 10% higher FWIs than female passengers. However, considering the FWIs as a result of Overcrowding and Other unspecified causes, female passengers accounted for 57 % higher FWIs than male passengers on average. This observation implied that the occurrences associated with female passengers may be predominantly as a result of external factors such as pushing and shoving, and (possibly) the gap size, while occurrences associated with male passengers may be predominantly as a result of the individual's behaviour, such as staff riding and attempting to board or alight the train while it was still moving.

Figure 4-6 presents the frequency of the different consequence classes that are expressed as FWIs for PTI occurrences. The consequence classes are defined according to the State of Safety report (RSR, 2018) as: "0" – no injury; "0-0.9" - Injury recorded and "1-1.9" – Fatality.

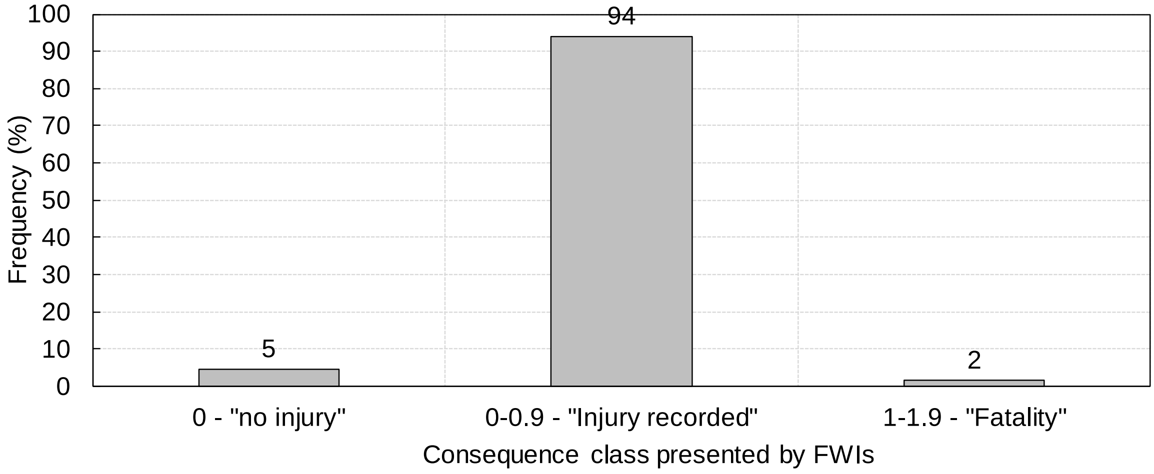

Figure 4-6: Likelihood of PTI occurrence severity based on 2014-2017 recorded incidents (RSR, 2014-2017)

The results presented in Figure 4-6 show that the number of occurrences increased significantly from the consequence class of 0 to the consequence class of 0.9. This observation can be interpreted to mean that in over 94% of all reported PTI occurrences, at least one person was injured. According to the RSR, PTI occurrence reporting is dependent on the severity of the injury. In most cases injuries such as shoulder grazes were usually not reported. Therefore, it can be assumed that only the serious injuries were reported and that the frequency of PTI occurrences is much higher than what is actually reported. This finding indicates the limitation of underreporting in the data. Considering the results presented in Figure 4-6, underreporting would result in the overestimation of the risk of occurrences that resulted in serious injuries, while occurrences with no injuries would be underestimated, therefore resulting in a biased representation of the risk of PTI occurrences, resulting in the misrepresentation of the actual PTI occurrences taking place.

4.3 MULTILINEAR REGRESSION ANALYSES

The first step in the data analysis process was to correlate the PTI occurrence frequency to the factors that were identified to have an influence on occurrences and passenger safety. Based on the results of the observation surveys and historical data, the main factors identified were overcrowding and the vertical gap size between the train and the platform. The effects of the train design could not be represented because the train type was not recorded in the incident descriptions provided. The results are presented in Table 4-3.

Table 4-3: Correlation factors determined

	Vertical gap size vs PTI occurrence frequency	Passenger volume vs PTI occurrence frequency
Correlation factor (R^2)	31.9%	97.5%

A clear correlation existed between passenger volume and PTI occurrence frequency (as expected). Furthermore, considering the size of the data available (N= 128), a correlation factor of approximately 32% between PTI occurrence frequency and the vertical gap size was considered significant and a clear indication that there was a relationship between these two factors albeit not linear. This relationship was further investigated by plotting the vertical gap size of each station along the Pretoria-Pienaarspoort corridor against the number of incidents recorded at the respective stations. Figure 4-7 presents the results.

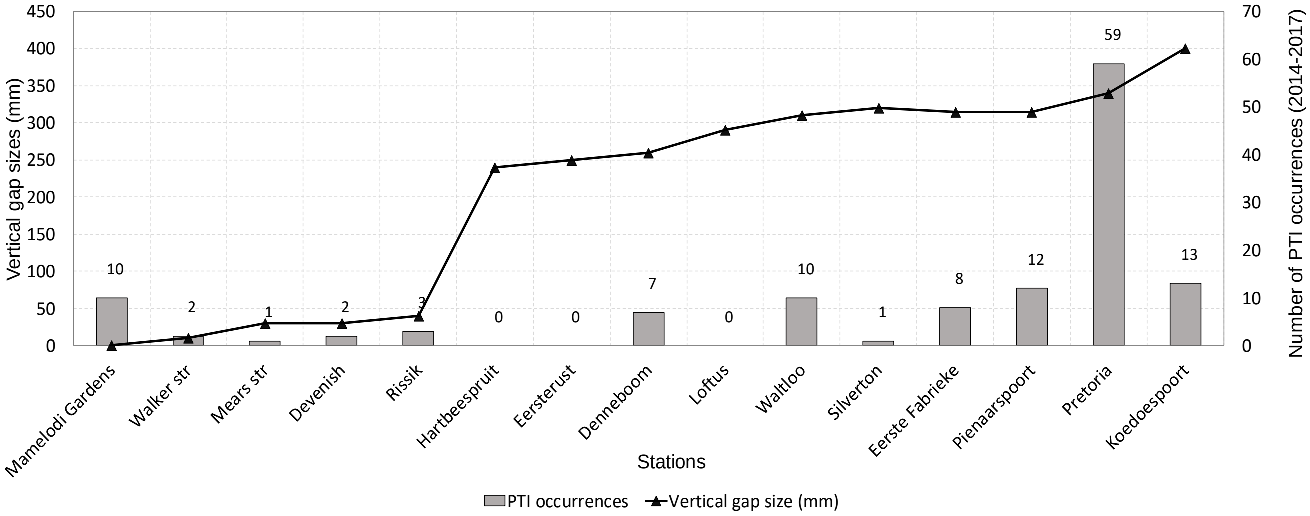

Figure 4-7: Vertical gap sizes measured at the different stations and the corresponding PTI occurrences recorded at each station (RSR, 2014-2017)

PTI occurrence frequency did not increase linearly with the increase in the vertical gap size, however, these occurrences were more prevalent and frequent in stations with a vertical gap size larger than approximately 250 mm as depicted in Figure 4-8. According to the study conducted by Wahl (2014), a station with a vertical gap size larger than 200 mm was classified as a high-risk station because it was concluded that once the vertical gap size was larger than 200 mm, there was an increase in PTI occurrences at that station. The PTI occurrence frequency was plotted against the vertical gap size to determine if a trend could be identified that would prove the conclusion made by Wahl (2014) (see Figure 4-8).

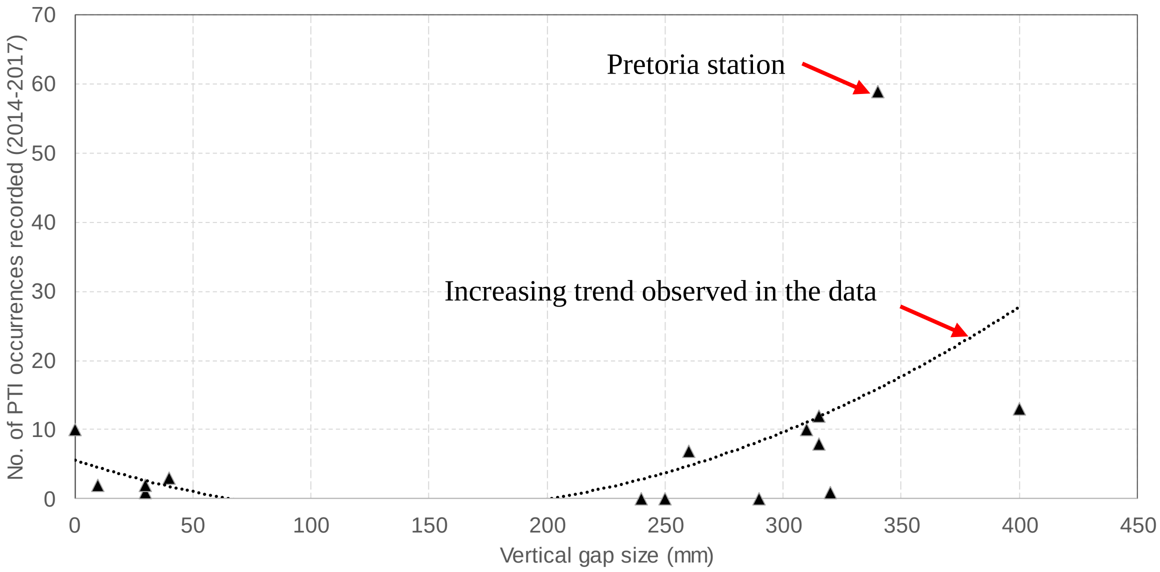

Figure 4-8: Total number of PTI occurrences obtained from the RSR (2014-2017)

The Pretoria-Pienaarspoort corridor did not have any vertical platform gaps within the range of 50 mm to 245 mm, explaining the absence of data in this range. This discontinuity in the data created a limitation in developing an effective model to describe the relationship between PTI occurrence frequency and the vertical gap size. However, an increasing trend could be deduced from the data presented, specifically considering the range of gap sizes larger than 230 mm. To further investigate this trend, the PTI occurrence frequency was normalised to represent the number of PTI occurrences per 10 000 passengers and plotted against the range of gap sizes larger than 200 mm. This was done to provide a better representation of the effects of the gap size while also considering the influence of passenger volumes at each station (see Figure 4-9).

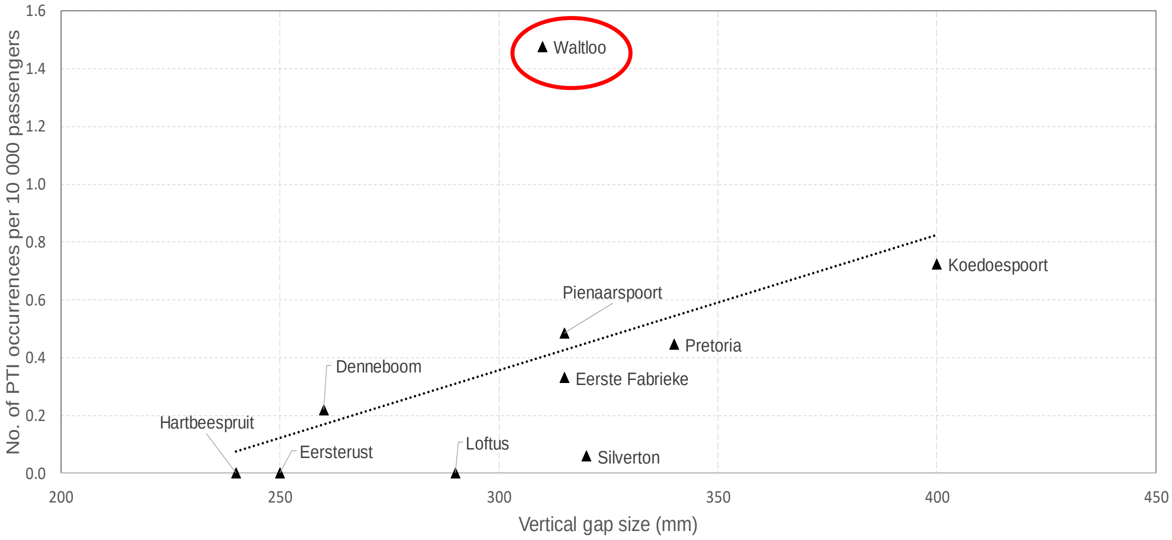

Figure 4-9: Number of PTI occurrences per 10 000 passengers vs. the vertical gap size (mm)

An approximately linear increasing relationship between the concentration of PTI occurrences at a station (PTI occurrence per 10 000 pass.) and the vertical gap size is depicted in Figure 4-9. In addition, the results revealed an outlier in the data, namely the concentration of PTI occurrences recorded at Waltloo station indicating that further investigation into the station is required. Figure 4-8 and Figure 4-9 further reveal that a threshold existed within the range of 200 mm to 230 mm where the gap size visibly influenced the frequency of PTI occurrences, thus proving Wahl’s (2014) conclusion.

The investigation was expanded further to determine the influence of other underlying factors such as gender, day of the week and time of day, on the risk of a PTI occurrence. This investigation was carried out by conducting multiple linear regression analyses on each variable and the concentration of PTI occurrences. The coefficient of determination (R^2) of the different combinations are presented in Table 4-4.

Table 4-4: Coefficient of determination of the different variable combinations

	PTI/10 000 pass.	Gender (Female vs Male)	Day of the week (Friday vs Not Friday)	Time of day (Peak vs Off-peak)
PTI/10 000 pass.	1	0.0217	0.0193	0.0214
Gender (Female vs Male)		1	0.0147E-1	0.0191E-1
Day of the week (Friday vs Not Friday)			1	0.0381E-2
Time of day (Peak vs Off-peak)				1

The R^2 values obtained from the linear regression analyses of the different variable combinations indicated that a linear regression was not adequate in representing the relationships between the different variables and the concentration of PTI occurrences. Therefore, an alternate approach was considered of determining if the concentration of PTI occurrences was station dependent. A Multilinear (ML) regression analysis was carried out and the different stations were represented using a binary variable. In addition, the vertical gap size was included in the regression, but was modified to only include gap sizes larger than 230 mm. The results obtained from the analysis are presented in Table 4-5. The stations were listed in the order of decreasing risk (based on the coefficient values) and only the statistically significant stations were included in the table.

Table 4-5: Results from ML regression analysis

	Coefficients	$\boldsymbol{p-value\ (p<0.05)}$
Intercepts(β_0)	3.88	9.67E-09
Gap size $\lvert x-230mm \rvert(\beta_9)$	0.01	1.27E-05
Waltloo (β_4)	6.71	3.16E-09
Mamelodi Gardens(β_7)	2.24	1.15E-02
Koedoespoort (β_2)	-1.95	5.80E-02
Pienaarspoort(β_8)	-3.25	4.02E-04
Pretoria station (β_1)	-3.99	1.60E-06
Eerste Fabrieke(β_6)	-4.80	3.26E-05
Denneboom station (β_5)	-5.19	2.65E-06
Silverton station (β_3)	-7.59	1.01E-04

The adjusted R^2 of the derived function was 72%, indicating a good fit of the predicted function. All the variables presented in Table 4-5 were indisputably significant except the variable presenting Koedoespoort, which was only marginally significant. It was however decided that the inclusion of the variable was necessary because Koedoespoort had the largest vertical gap size recorded along the corridor (≈ 400 mm). Considering the coefficients obtained from the model estimation, the relationship between the vertical gap size and PTI concentration is positive. This result, therefore, proves that increasing the vertical gap size results in the increased likelihood of a PTI occurrence.

The following results are highlighted:

- Waltloo station was the highest risk station based on the prediction model (corresponding to the results presented in (Figure 4-9).
- Mamelodi Gardens ranked second even though the vertical gap size measured at this station was approximately 10 mm (well below the hypothesised threshold).
- Pretoria station was ranked fourth even though it was the busiest station along the corridor, and for this reason, was expected to be the highest risk station.

These observations implied that the vertical gap size and passenger volumes were not the only factors that could influence the risk of PTI occurrences and that the risk of an occurrence was also station dependent. For this reason, further investigation into the other underlying factors that resulted in the varied station risk classification was required, thus motivating the determination of relative risk based on underlying factors such as time of day, day of the week and passenger gender.

4.4 CALCULATING RELATIVE RISK

Relative risk was defined as the likelihood of a PTI occurrence based on the frequency of the occurrence relative to the entire corridor population and as a result of a specific influencing factor. The calculation of relative risk was carried out by applying principles of human factors research. Wilson (2014) stated that the focus of human factors studies should be primarily the study of the interconnections of the system components, namely the interaction of humans and the infrastructure around them. In the case of investigating PTI occurrences, the causes and factors identified served as representations of the components of a system that resulted in a PTI occurrence. The interactions of these components were presented using a Fault Tree similar to the approach taken by Liming et al. (2010). Figure 4-10 presents the different factors considered.

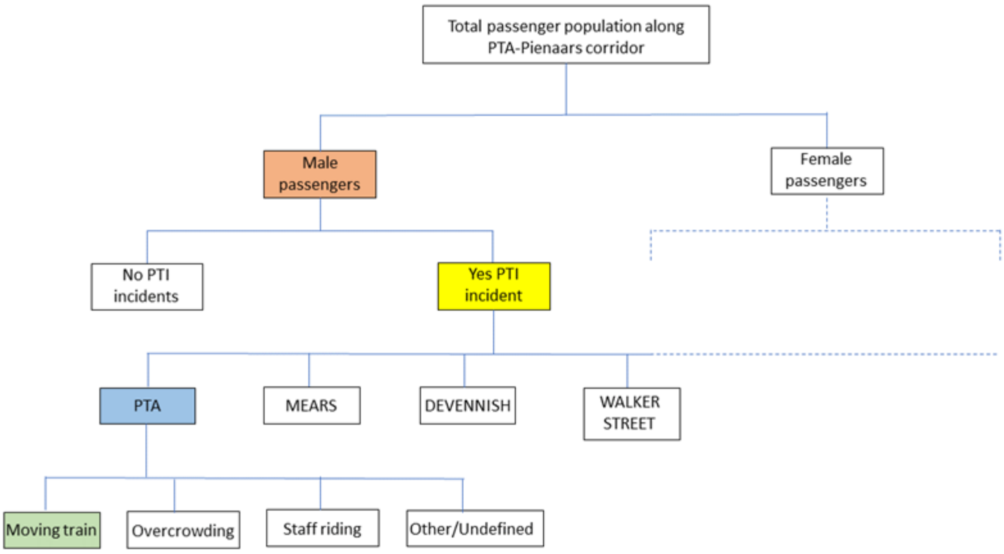

Figure 4-10: Partial schematic of the Fault tree approach

The investigation into the interaction between the different factors gave an indication of the most critical factors to consider when mitigating PTI occurrences. The different system components considered were the gender of the victims, the station along the corridor, the time of the incident and the cause of the incident.

The fault tree consisted of the entire population of commuters along the Pretoria-Pienaarspoort corridor based on the 2008 census. The population was then categorised according to gender using the percentage distribution values obtained from the observation surveys (68% male passengers and 32% female passengers).

The relative risk was calculated as follows:

- Consider determining the relative risk of a male passenger who experienced a PTI occurrence at Pretoria station because he was trying to board/alight the train while it was still moving. The relative risk of a passenger who lies in this category was determined by Equation 4.2:

$$Relative\ Risk\ (\%) = 68\% \times \frac{N_{Male}}{N} \times \frac{N_{Male_{Pretoria}}}{N_{Pretoria}} \times \frac{N_{Male_{Pretoria_{Moving\ train}}}}{N_{Male_{Pretoria}}} \qquad \textbf{Eq 4-2}$$

$$Relative\ Risk\ (\%) = 68\% \times \frac{N_{Male}}{N} \times \frac{N_{Male_{Pretoria}}}{N_{Pretoria}} \times \frac{N_{Male_{Pretoria_{Moving\ train}}}}{N_{Male_{Pretoria}}} \qquad \textbf{Eq 4-2}$$

$$Relative\ Risk\ (\%) = 68\% \times \frac{N_{Male}}{N} \times \frac{N_{Male_{Pretoria_{Moving\ train}}}}{N_{Pretoria}} \qquad \textbf{Eq 4-3}$$

The same approach was carried out when considering the female passenger population and the different scenarios outlined previously. The generalised equations applied in calculating relative risk are presented in Table 4-6. In the equations, x represents the factor of gender, y represents the time of occurrence (time, day or year), z represents the different causes of the occurrences identified and lastly, A represents the station where the incident occurred.

Table 4-6: Equations applied in carrying out the calculation of relative risk

x passenger PTI occurrence rate calculation in each recording year y	$\frac{Total\ no.\ of\ x\ occurrences\ recorded\ per\ annum}{Total\ commuter\ population\ per\ annum}$	$\frac{N_x}{N}$
PTI occurrence rate at station A in each recording year y	$\frac{N_x\ in\ station\ A\ per\ annum}{Total\ commuter\ population\ at\ station\ A\ per\ annum}$	$\frac{N_{x_A}}{N_A}$
PTI occurrence cause percentage contribution at station A	$\frac{N_x\ caused\ by\ z\ at\ station\ A}{N_x\ in\ station\ A}$	$\frac{N_{x_{A_z}}}{N_{x_A}}$

The results obtained are presented in Figures 4-12 to 4-14. The relative risk was calculated per 10 000 passengers at each station along the corridor within the 2014 to 2017 reporting periods. Figure 4-11 presents the results when only considering the gender of the passenger.

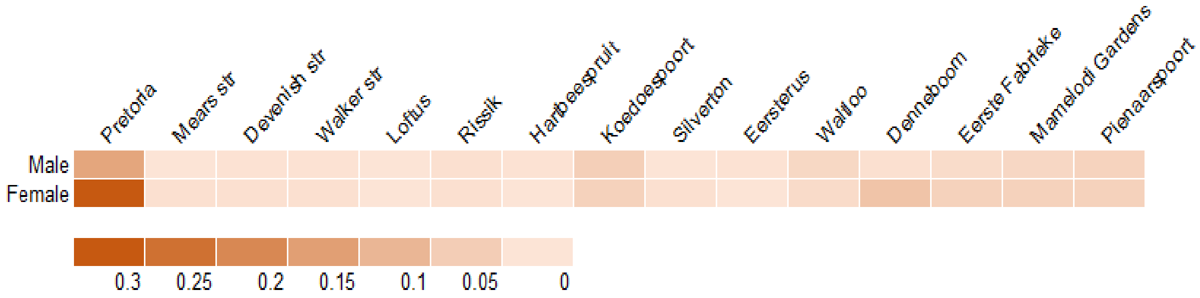

Figure 4-11: Relative risk per 10 000 passengers calculated according to stations and gender.

On average, the relative risk of a PTI occurrence calculated for female passengers was 47% higher than that calculated for male passengers. This result implied that female passengers were more susceptible to experiencing PTI occurrences (especially at Pretoria station). This finding highlights the possible unsafe conditions female passengers were exposed to at station platforms and possibly explains the reason fewer female commuters use the Metrorail service compared to male commuters. Figure 4-12 presents the results obtained when considering the day of the week.

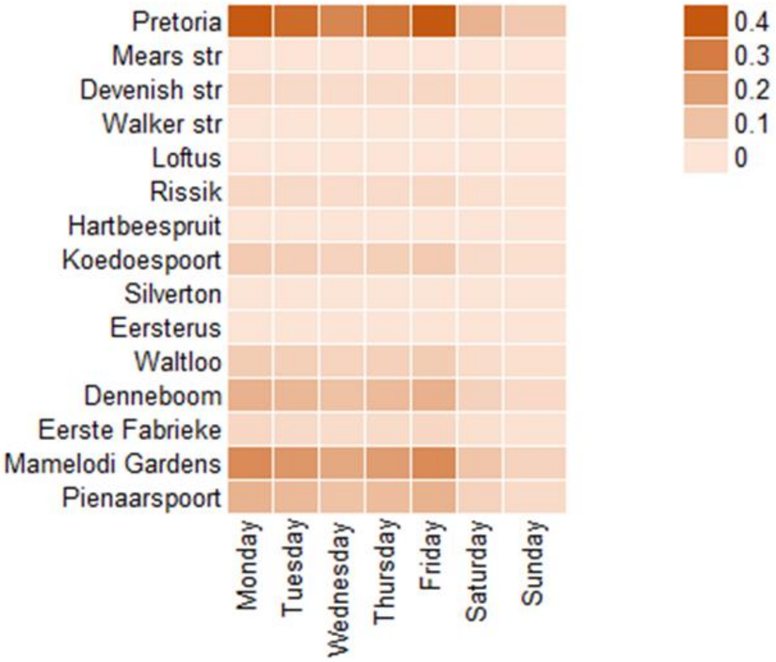

Figure 4-12: Relative risk per 10 000 passengers calculated according to stations and the day of the week.

The results presented in Figure 4-12 indicate that on average, a passenger is at high risk of experiencing a PTI occurrence on a weekday. However, the colour variation does not clearly indicate the exact day(s) of the week PTI occurrences were likely to occur. Following the approach taken by Liming (2010), the importance of each day was calculated using Equation 4-4:

$$\boldsymbol{W(X_i) = \frac{N_{\Phi(X_i)=0}}{N_{X_i=0}}} \qquad \textbf{Eq 4-4}$$

Where $N_{\Phi(X_i)=0}$ represents the cumulative relative risk of a PTI occurrence on each respective day, $N_{X_i=0}$ the percentage of PTI occurrences recorded on each respective day and $0 \leq W(X_i) \leq 1$. Additionally, the variable (day of the week) importance was calculated and defined using Equation 4-5:

$$\boldsymbol{W_M(X_i) = \frac{N_{\Phi(X_i)=0}}{N_{\Phi=0}}} \qquad \textbf{Eq 4-5}$$

Where $N_{\Phi=0}$ indicates the total PTI occurrences recorded along the corridor. By considering $W(X_i)$ and $W_M(X_i)$, the most critical days of the week were identified, and the results presented in Figure 4-13.

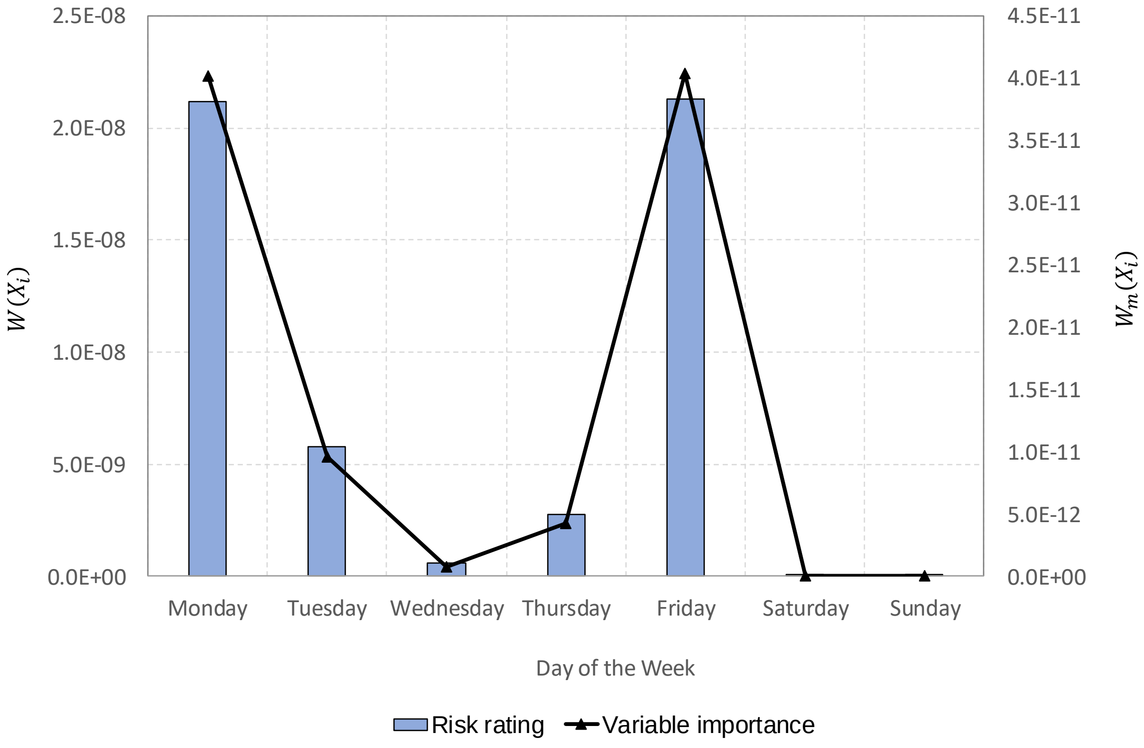

Figure 4-13: Critical day of the week calculation based on $W(X_i)$ and $W_M(X_i)$

The results in Figure 4-13 indicate that the most critical days were Monday and Friday when PTI occurrences were most likely to occur. This is explained by the fact that the highest volume of passengers has been recorded to travel on these two days (refer to the results presented in Section 3.3). Figure 4-14 presents the results obtained from calculating the relative risk based on the time of day.

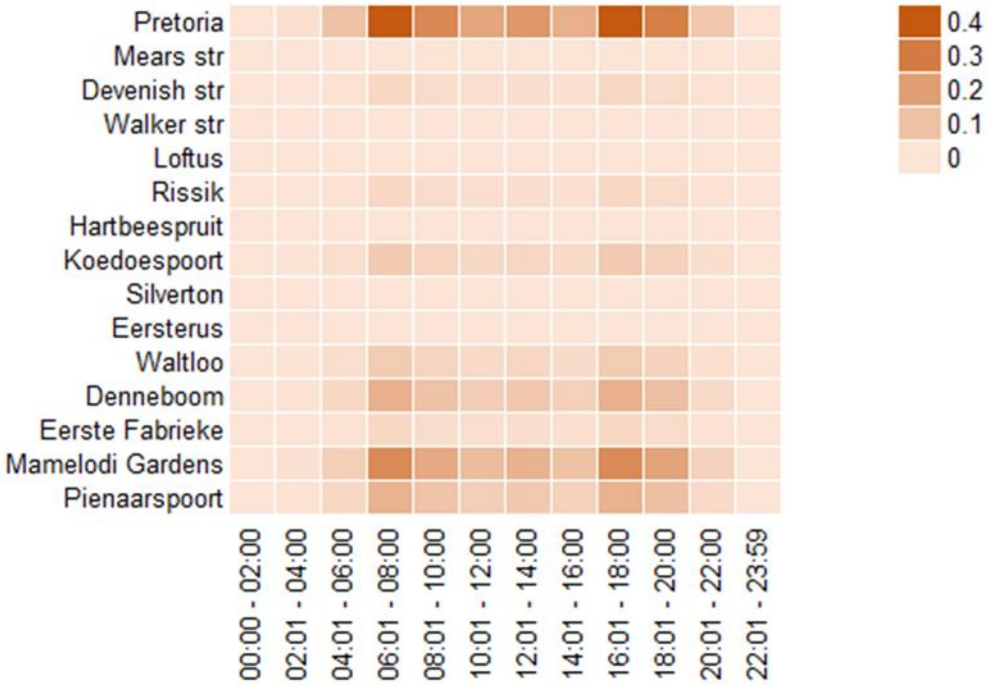

Figure 4-14: Relative risk per 10 000 passengers calculated according to stations and the time of day.

The results presented in Figure 4-14 reveal that passengers were at the highest risk of experiencing a PTI occurrence during peak hours because of overcrowding on station platforms during those times. The results were also presented in terms of $W(X_i)$ and $W_M(X_i)$(see Figure 4-15).

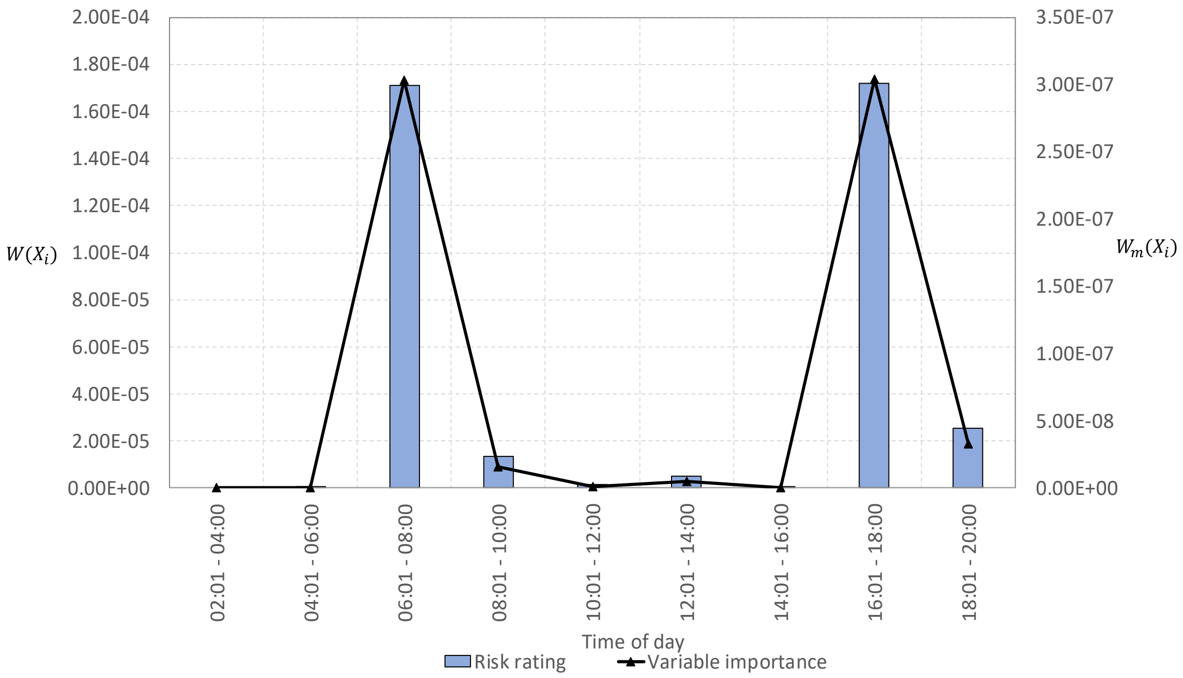

Figure 4-15: Critical time of day calculation based on ($W(X_i)$ and $W_M(X_i)$)

The results in Figure 4-15 indicate that the most critical times when a passenger is most likely to experience a PTI occurrence is at 06h01 – 08h00 and 16h01 – 18h00. Considering the measure of relative risk in Figure 4-14, it was determined that passengers are at a 17% higher risk of a PTI occurrence during the morning peak. The final calculation of relative risk was done to determine the most prevalent cause of PTI occurrences at each station along the corridor (see Figure 4-16).

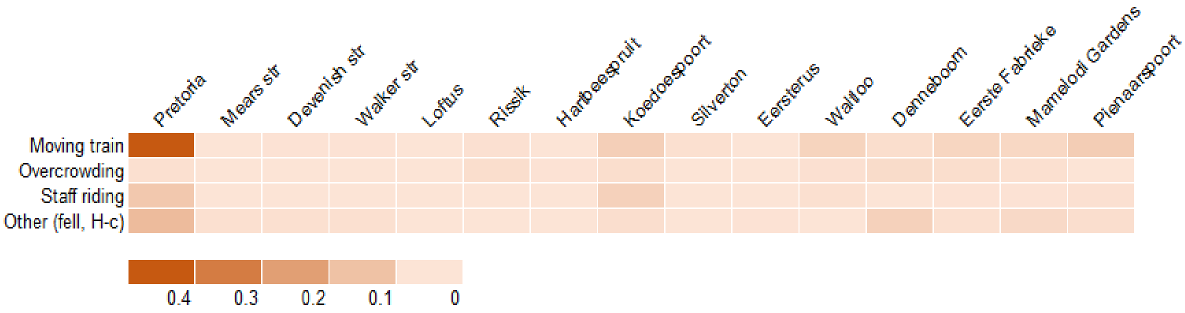

Figure 4-16: Relative risk per 10 000 passengers calculated according to the causes of PTI occurrences.

The results presented in Figure 4-16 indicate that the most prevalent cause of PTI occurrences along the corridor was the "Moving train", in addition, the results show that Pretoria station was predominantly the highest risk station along the corridor. The overall risk classification for all stations along the corridor was carried out by calculating the total $W(X_i)$ and $W_M(X_i)$ of each station. The results were plotted in relation to the gap size measured at each station (see **Error! Reference source not found.**).

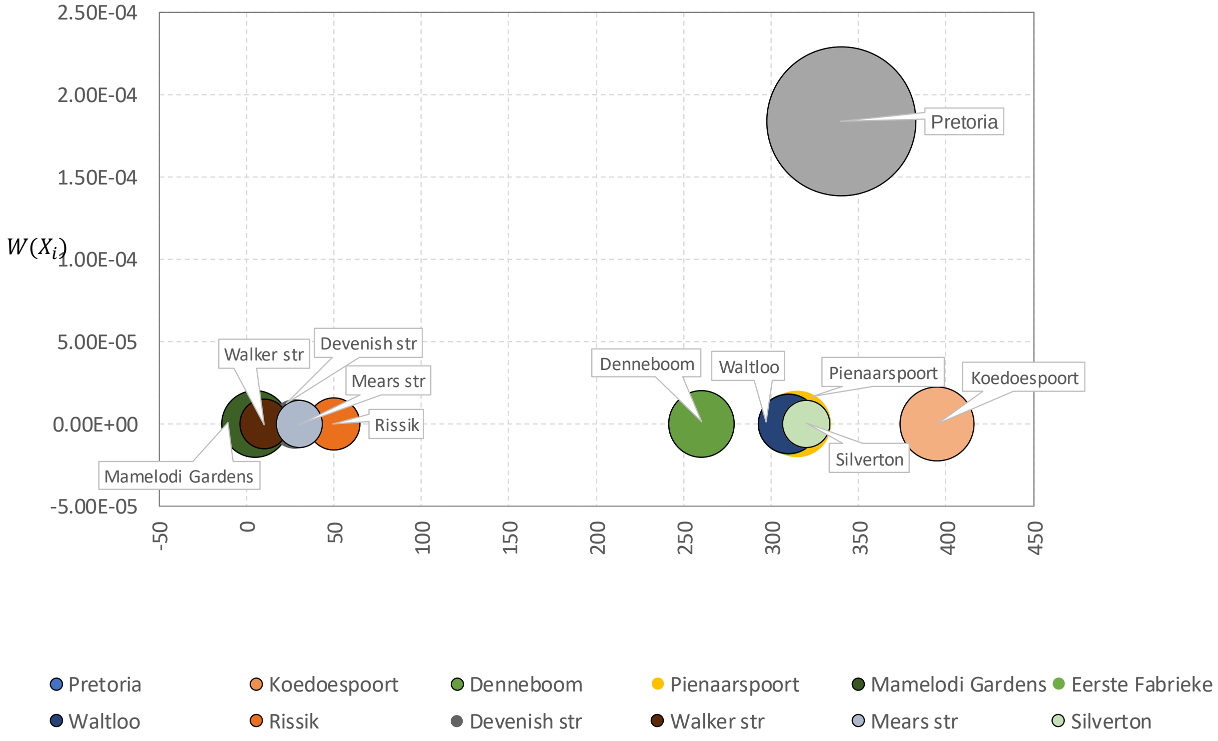

Figure 4-17: Risk classification of stations along the Pretoria-Pieanaarspoort corridor

In **Error! Reference source not found.**, the size of the circles are represented by the value of $W_M(X_i)$, therefore the higher the position of the circle in relation to the x-axis and the larger its size, the higher the risk level of that station. For example, Pretoria station is observed to be the highest risk station according to the results presented. The results in **Error! Reference source not found.** also indicate clustering of the data points, where the circle sizes are on average larger with increasing vertical gap size, thus indicating that the risk level of a station increases with increasing vertical gap size. The risk classification of the stations was thereafter carried out considering the vertical gap size and the values of $W(X_i)$ and $W_M(X_i)$ (see Table 4-7).

Table 4-7: Risk classification of stations along the Pretoria-Pieanaarspoort corridor

Risk Ranking	Station	Vertical gap size (mm)
High	Pretoria	340
	Koedoespoort	395
Medium	Denneboom	260
	Pienaarspoort	315
	Mamelodi Gardens	5
	Eerste Fabrieke	310
	Waltloo	310
Low	Rissik	50
	Devenish str	28
	Walker str	10
	Mears str	30
	Silverton	320

Considering the results presented in Table 4-7, it was concluded that the most critical stations along the corridor were Pretoria and Koedoespoort stations. The main characteristic that would explain this conclusion is that both stations serve as end stations/junctions where multiple train transfers occur daily. This characteristic of the stations results in a higher flow of passengers from one platform to another and this high volume of passengers may increase the risk of passengers experiencing a PTI occurrence.

The risk ranking in this section took precedence over the initial ranking carried out with the predicted linear function (see Section 4.3). This was done because the linear regression ranking was only dependent on the station and vertical gap size, while the relative risk ranking considered more factors (e.g. gender, time of day and the day of the week) and was therefore considered to better represent the risk of PTI occurrences.

The governing trend observed in Table 4-7 is that stations with a considerably large vertical gap size had consistent records of PTI occurrences; stations such as Pretoria, Koedoespoort, Denneboom and Pienaarspoort. However, the influence of the other underlying factors can be inferred through the inconsistency observed with stations such as Mamelodi and Silverton that did not follow this governing trend. To quantify the influence of these underlying factors a Multinomial Logistic (MNL) regression was carried out to quantify and demonstrate how all factors considered contributed to the probability of the type of PTI occurrence, i.e. the probability of an occurrence caused by a “Moving train”, “Staff riding”, “Overcrowding” or “Other (unspecified causes)”. The development of a prediction function that could adequately describe the statistical relationships between all the variables considered and the risk of a PTI occurrence would allow for the calculation of a safety rating for corridors in the long-term.

4.5 LOGISTIC REGRESSION IN RISK CALCULATION

The next step in the data analysis process was to determine the probability of the type of occurrence and develop a model that could adequately estimate the risk of a PTI occurrence based on all factors presented thus far. In this section, risk was redefined based on the formal definition, which states:

$$\boldsymbol{Risk = Probaility\ of\ failure\ \times Consequences\ of\ failure} \qquad \text{Eq 4-6}$$

In the case of investigating PTI occurrences, the probability of failure was defined as the probability of a PTI occurrence caused by, (i) “moving train”, (ii) “overcrowding”, (iii) “staff riding” or (iv) “other”. The consequences in Equation 4-6 were represented by the Fatalities and Weighted injuries (FWIs) calculated using Equation 4-1.

The choice of developing a Multinomial logistic (MNL) model was made based on the data type available for the exercise (a combination of continuous and discrete binary data and the presence of more than one outcome-type dependent variable). In the case where multiple binary variables are considered, multinomial regression analysis is commonly the most appropriate choice because logit-based models enable the identification and quantification of the effects of factors that contributed to the different outcome-types (Ye and Lord, 2013).

However, the limitation commonly encountered when applying these models is that the analysis is often outcome-based, which results in a biased estimation of the model parameters. The bias created in the model can result in the overestimation of the probability of higher consequence occurrences and underestimation of lower consequence occurrences. Adding to this limitation, the case of underreporting results in data inaccuracies that lead to the misrepresentation of the true shares of the occurrence types in the population. This limitation of underreporting can be partially addressed through the application of the proposed Weighted Exogenous Sample Maximum Likelihood Estimator (WESMLE) (Patil and Lord, 2011; Ye, 2011). However, this approach is only effective in addressing bias if the rate of underreporting is known, or at least partially known, otherwise the use of an incorrect underreporting rate may cause larger biases.

In the case of PTI occurrence reporting, the probability of reporting an incident was based on injury severity, so in most cases minor injuries were not reported (i.e. incidents with a lower consequence class were more likely to be underreported) (RSR, 2018). The consequence class distribution of the type of PTI occurrences was plotted to determine if underreporting of any type was prevalent in the incidents recorded from 2014 to 2017 (Figure 4-18).

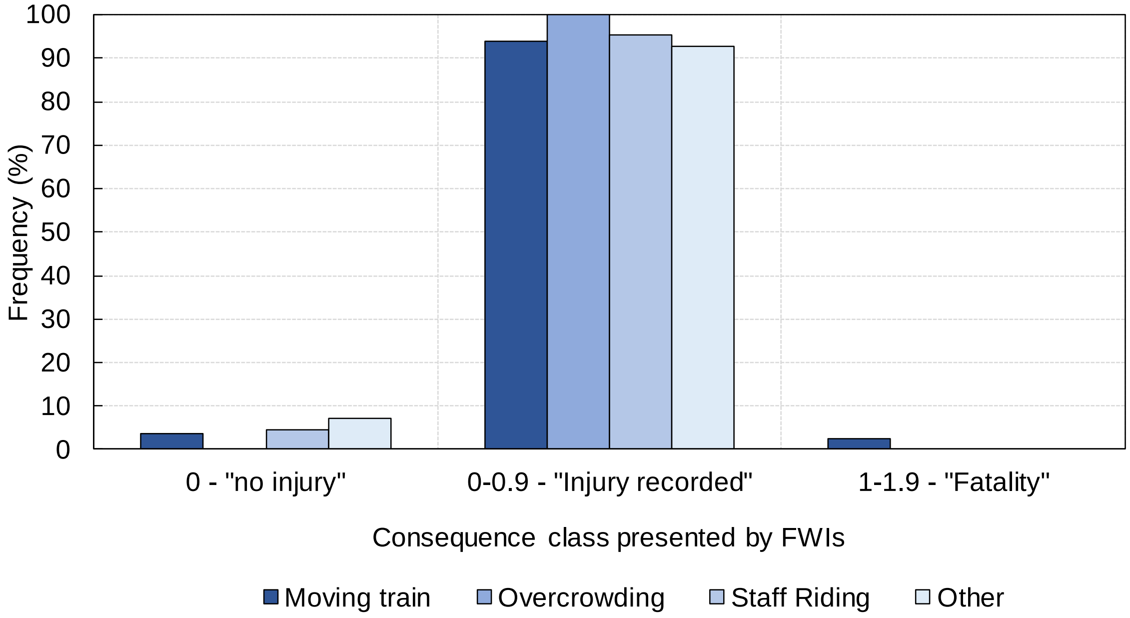

Figure 4-18: Consequence class distribution based on occurrence type

Considering the four causes identified, it can be deduced from Figure 4-18 that occurrences caused by overcrowding were the most underreported because 100% of the reported occurrences resulted in an injury. Incidents caused by a moving train had the highest reporting rate and for this reason, the moving train outcome was used as the baseline outcome to minimise the bias in the model prediction. The Multinomial Logistic function was approximated using the IBM SPSS Statistics program based on the variables presented in Table 4-8.

Table 4-8: Variables applied in estimating the PTI occurrence Multinomial Logit model

Variable no.	Variable Description
1	Type of PTI occurrence (Outcome): 1 if Overcrowding, 2 if Staff riding, 3 if Other, 4 if Moving Train
2	Gap sizes corresponding to occurrences caused by Overcrowding
3	Gap sizes corresponding to occurrences caused by Staff riding
4	Gap sizes corresponding to occurrences caused by Other
5	Gap sizes corresponding to occurrences caused by Moving train
6	Station risk 1 classification: 1 if high risk, 0 if medium/low risk
7	Station risk 2 classification: 1 if high/med risk, 0 if low risk
8	Time of day: 1 if peak hour, 0 if off-peak
9	Day of the week 1 classification: 1 if Monday, 0 if not Monday
10	Day of the week 2 classification: 1 if Friday, 0 if not Friday
11	Gender classification: 1 if Male, 0 if Female

To investigate the appropriateness of the choice of variables in Table 4-8, the following tests were conducted to determine if the assumptions required for the application of the MNL regression were met:

1. **Testing the collinearity of the independent variables**: Based on the data analysis carried out in Section 4.3, the vertical gap size and the station risk classification were observed to be correlated. A simple regression analysis was carried out to support this observation and the results revealed that a perfect correlation of 1.00 existed between the vertical gap size and the station risk classification.

2. **Ensuring mutual exclusivity of the dependent variable**: Based on the results presented in Figure 4-4, 20% of the occurrences recorded were as a result of a combination of the causes identified. To ensure mutual exclusivity of the dependent variables, all occurrences as a result of more than one cause were excluded from the model.

3. **Testing the linear relationship between the vertical gap size and the logit transformation of the dependent variable**: A logit transformation was carried out on the dependent variable and thereafter correlated with the vertical gap size. The results revealed a coefficient of determination (R^2) of 7.5%, indicating that a linear relationship was not prevalent between the continuous variable (the vertical gap size) and the dependent variable.

4. **Removal of outliers**: Outliers such as Loftus station were excluded from the model. Loftus station had no records of PTI occurrences even though it had one of the largest gap sizes. The results from the observation survey also motivated its classification as a high-risk station. The lack of PTI occurrence records at the station was assumed to be as a result of the lack of supervision.

5. **Testing the assumption of a large sample size**: Lastly, the logistic regression typically requires a large sample size of at least 2 000 data points (Ye and Lord, 2013). Small sample sizes significantly affect the development of outcome-based models. Also, bias is produced when using an insufficient sample size. The general criteria followed is that a minimum of 10 cases is required with the least infrequent outcome for each independent variable in the model. This meant that, when considering the 6-independent variable and the least frequent outcome of 0.08, it was required that the minimum sample size be 720 = (10*6 / 0.08) (Schreiber-Gregory, 2018). However, during the research of the Pretoria-Pienaarspoort corridor the data available was a sample size of 96 recorded occurrences.

Considering the assumptions outlined by Schreiber-Gregory (2018) and the results from the tests conducted to ensure the assumptions were met, the limitation of the sample size was acknowledged, however it was decided that the Multinomial logistic regression would still be carried out to determine any potential underlying trends that could prove valuable to the study, while taking into account the limitations. In addition, the vertical gap size was excluded from the model due to the collinearity observed between the station risk classification and the gap size. The station risk classification 1 was used instead to represent a proxy for the vertical gap size, where a high-risk station was assigned a value of 1 and a low/medium risk station was assigned a value of 0 (Refer to Table 4-8).

The estimated model parameters are presented in Table 4-9 and only include the intercepts and the significant variables. The detailed results including all the variables are attached in Appendix A.

Table 4-9: Variable outcomes of the MNL regression analysis

Type outcome		β	**Sig.** $(p - values)$
Overcrowding (OC)	Intercept	-1.755	0.214
	Station risk 1 (Med/Low risk)	1.865	0.086
Staff riding (SR)	Intercept	-4.082	0.004
	Day of the week 2 (not Friday)	1.944	0.081
Other	Intercept	-2.964	0.007
	Gender (Female)	1.629	0.009
The reference category is the outcome: Moving train (MT)			

Considering the significance level of $p < 0.05$ the results presented in Table 4-9 were interpreted as follows:

1. All else being equal, occurrences caused by a Moving train were the most likely to occur. This conclusion was drawn based on the negative values estimated for the intercepts of the other variable outcomes.

2. Secondly, considering the outcome estimated for Overcrowding, the parameter estimate indicates that a PTI occurrence is more likely to occur due to Overcrowding (as compared to Moving train) at a low risk station, inversely implying that a passenger would be more likely to experience a PTI occurrence as a result of a Moving train at a high-risk station.

3. In addition, it was estimated that a passenger was more likely to experience a PTI occurrence as a result of staff riding on any day that was not Friday. Thus, implying that a passenger was more likely to experience a PTI occurrence as a result of a Moving train on Friday.

4. Lastly, according to the predicted model female passengers were more likely to experience a PTI occurrence as a result of Other causes as compared to a Moving train. Therefore, implying that male passengers are predicted to be more likely to experience a PTI occurrence caused by a Moving train.

Based on the model outputs, only the station risk classification, day of the week and gender were significant in influencing the model outcomes. Considering the overall significance of the model, the Likelihood Ratio Tests table presented in Table 4-10, highlights the same conclusion regarding variable significance.

Table 4-10: Likelihood Ratio Tests of the overall estimated model

Effect	Model Fitting Criteria -2 Log Likelihood of Reduced Model	Likelihood Ratio Tests Chi-Square	df	Sig.
Intercept	62.497	0.000	0	
Station_risk1(High)	68.942	**6.444**	**3**	**0.092**
Time_of_day	66.458	3.961	3	0.266
Day_week1(Monday)	65.448	2.951	3	0.399
Day_week2(Friday)	70.677	**8.180**	**3**	**0.042**
Gender	70.176	**7.679**	**3**	**0.053**

The statistical characteristics of the estimated model are presented in Table 4-10. The table includes the significance measure, accuracy and the goodness-of-fit of the model measured by the Pearson and Deviance coefficient.

Table 4-11: Model Fit information

Model	Model Fitting Criteria -2 Log Likelihood	Likelihood Ratio Tests Chi-Square	df	Sig.
Intercept only	89.824			
Final	62.497	27.327	15	0.026
		Goodness-of-Fit		
		Chi-Square	df	Sig.
Pearson		19.376	33	0.971
Deviance		22.203	33	0.923

Based on the results presented in Table 4-11, the prediction model was significant and had a Pearson coefficient of 97.1%. This was considered an excellent fit, however the inability to fully describe the influence of all the factors presented a significant limitation in the accuracy of the model. In addition, the limitation of considering each outcome independently was an added limitation because the results presented in Section 3 implied that the incident outcomes were not mutually exclusive as assumed in the development of the estimated model. For example, the likelihood of an incident caused by a moving train is increased in the presence of a large gap and overcrowding (Refer to Figure 4-16). This observation implies that the different outcomes influence each other. Therefore, the estimated model can serve as only an indicator of the expected percentage distribution of the causes of PTI occurrences. Furthermore, the estimated model can provide an indication of the most necessary interventions to address the most serious cause of PTI occurrences.

The total expected values were calculated from the ML model estimated in Section 4.3 and the results plotted against the actual values recorded to determine the percentage difference between the two data sets and test the model effectiveness in estimating the expected FWIs per 10 000 passengers assuming all predicted PTI occurrences resulted in an injury. The results are presented in Figure 4-19.

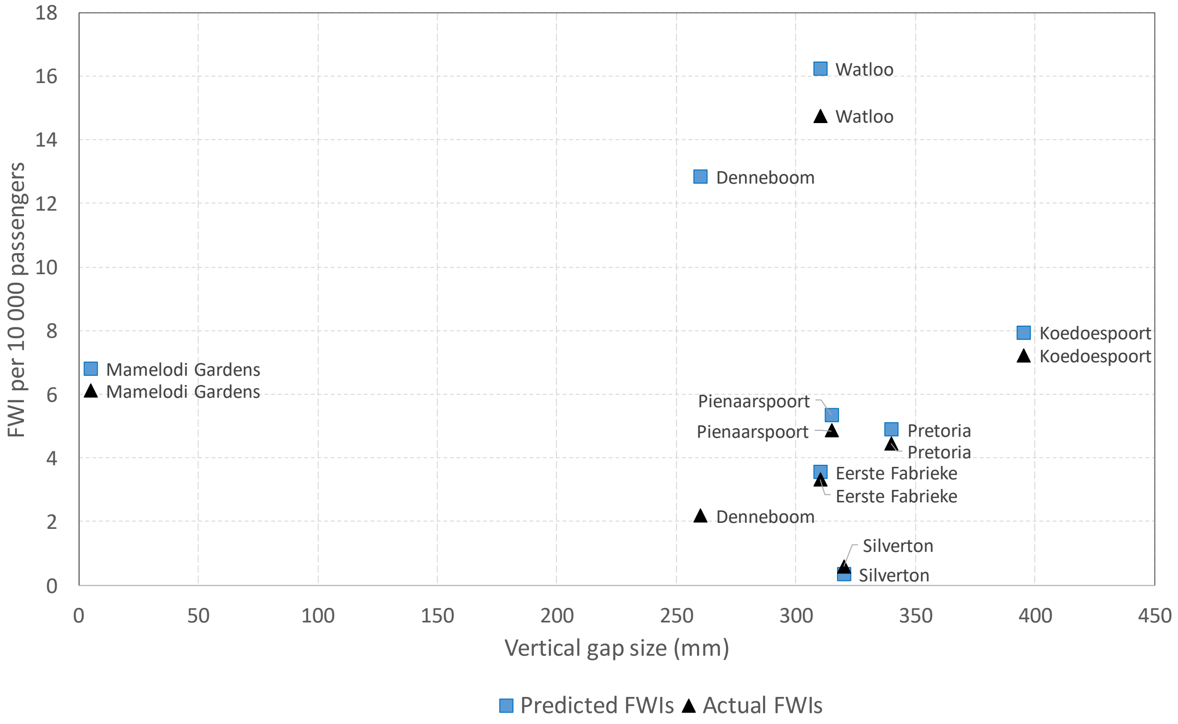

Figure 4-19: Estimated FWIs from the ML model

In general, the model overestimated the expected PTI occurrences, especially considering Denneboom station, where the model predicted approximately 83% more FWIs than the actual FWIs recorded. This finding indicates that Denneboom may be an outlier and that further investigation into the platform conditions at the station is required. To further explain the results obtained from the MNL model, the different scenarios that could result in a PTI occurrence are presented in a Fault Tree format in Figure 4-20. The total estimated probability values are presented in Table A4 and A5 in Appendix A.

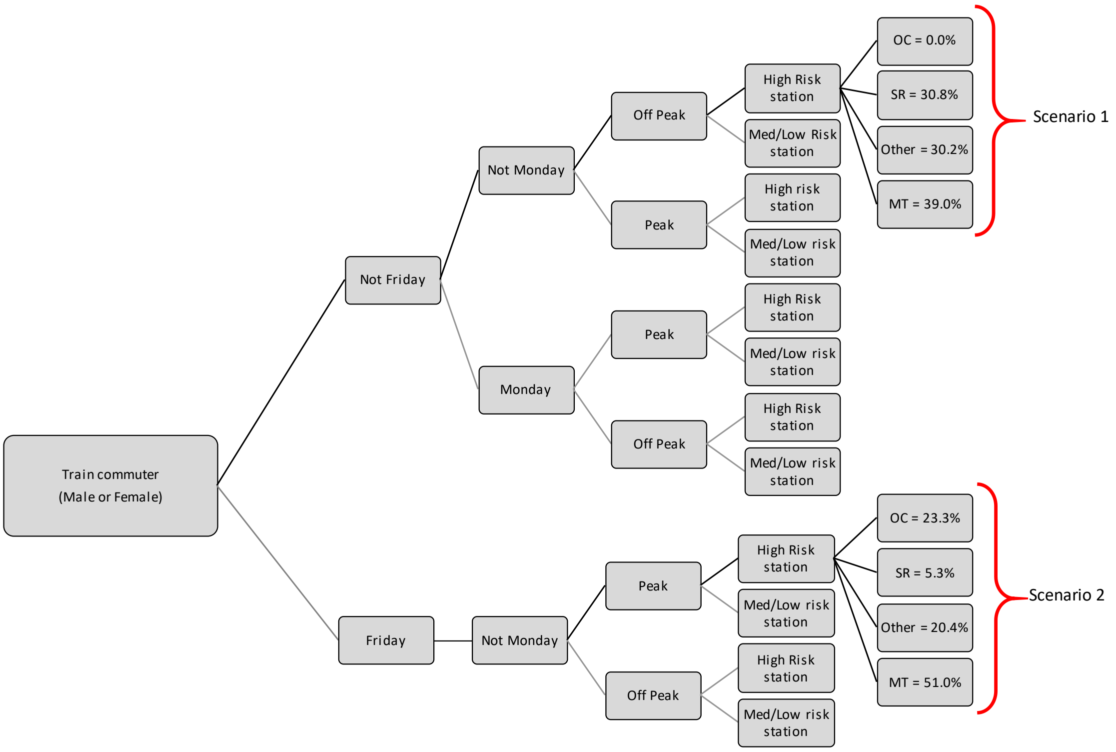

Figure 4-20: Fault tree diagram presenting all scenarios considered that could result in a PTI occurrence.

Each scenario presented in Figure 4-20 resulted in an estimated percentage cause distribution. To consider the overall expected percentage distribution, the average estimated FWIs per 10 000 passengers for each cause in all scenarios considered along the corridor was calculated and the results presented in Figure 4-21. The calculation was carried out by considering the estimated total FWIs (presented in Figure 4-19) calculated from the ML model and the percentage distribution of each cause was then calculated and presented as the number of FWIs per cause identified (Figure 4-21). The actual recorded FWIs are presented in Figure **4-22** .

Figure 4-21: Estimated FWIs obtained from MNL function

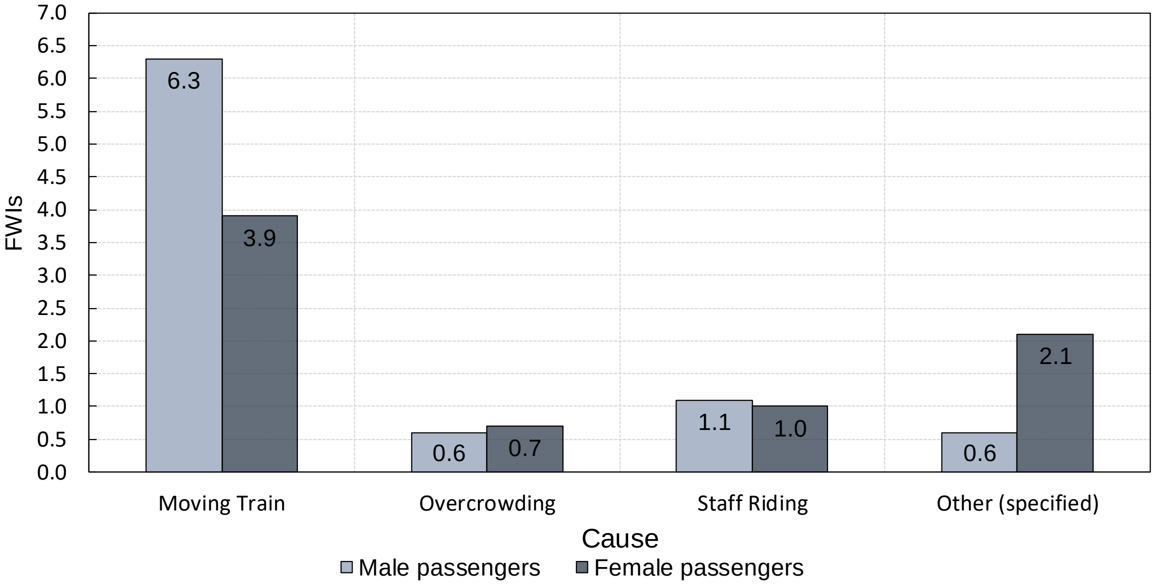

Figure 4-22: Actual recorded FWIs per cause identified along the corridor

The results presented indicate that the developed MNL model grossly underestimated the expected number of FWIs per cause when compared to the actual numbers recorded, however, the trend observed from the calculated FWIs was similar to the reported FWIs. The inaccuracies in the estimations can be attributed to the data manipulation carried out before the application of the MNL regression analysis. Furthermore, this finding indicates that further research is required to improve the quality of data available on incidents in order to improve the investigation process.

Increasing the sample size by conducting the investigation on a larger scale is one option, alternatively, identifying other underlying variables may address the randomness in the data and improve the model estimations. The introduction of other variables such as the train type and age of the victims would add depth to the analysis, but this can only be achieved if the occurrence reporting process is improved and includes these factors.

In summary, the data analysis proved that platform design (i.e. vertical gap size and station risk classification), human behaviour influenced by the time of day and day of the week, individual factors such as passenger gender, and the train design (i.e. manual train doors and train design that allows for staff riding) can be considered the main factors that influence the risk of a passenger experiencing a PTI occurrence. The data analysis can be further expanded to determine the effect of the risk of a PTI occurrence on the corridor capacity. In this way, the cost to the service provider can be hypothesised, and the value of implementing safety interventions in stations can be motivated.

4.6 EFFECTS OF PTI OCCURRENCES ON CORRIDOR CAPACITY

Considering the results obtained from the measure of passenger flow rate, the improvement of safety is directly linked to the improvement of service efficiency. The implementation of interventions that would streamline the boarding and alighting process of passengers would not only reduce the risk of a PTI occurrence, but also result in the reduction of the required dwell time at stations. This reduction in the dwell time would result in shorter lead times between trains and therefore create the opportunity of adding more trains into the network and lead to increased train availability (Carey and Carville, 2000). In addition, the increased safety at the PTI would reduce the risk of an occurrence, thus avoiding the delays caused by occurrences.

According to the results obtained from the observation surveys, the service reliability of Metrorail was 60% (on average) with the operation of the 10M4 (“blue”) trains. This low reliability caused excessive delays that resulted in increased crowding on the platform, further increasing the risk of PTI occurrences as proven in the previous sections. Furthermore, the unsafe conditions of travel also result in reduced train ridership and the inability to attract new users.

The effects of improving service efficiency and safety was tested by applying a Monte Carlo simulation over 300 hours to determine the average line capacity that can be achieved considering the following scenarios:

Table 4-12: Scenarios considered in the calculation of operation capacity

	Assumptions made in the calculation
Scenario 1	Only 5M2A (“yellow”) train operations at a 60% reliability rate
Scenario 2	Only 10M4 (“blue”) train operations at a 60% reliability rate
Scenario 3	Mixed operation of the 5M2A (“yellow”) and 10M4 (“blue”) trains at an 80% and 20% operation share

In each scenario the likelihood of a PTI occurrence was considered by including the probability of an injury occurring in the event of an occurrence that would result in the disruption of normal operations. The probability of an injury was determined by carrying out a simple logistic regression on the data of recorded PTI incidents along the Pretoria-Pienaarspoort corridor from 2014 to 2017.**Figure 4-23** presents the different scenarios considered in the regression analysis.

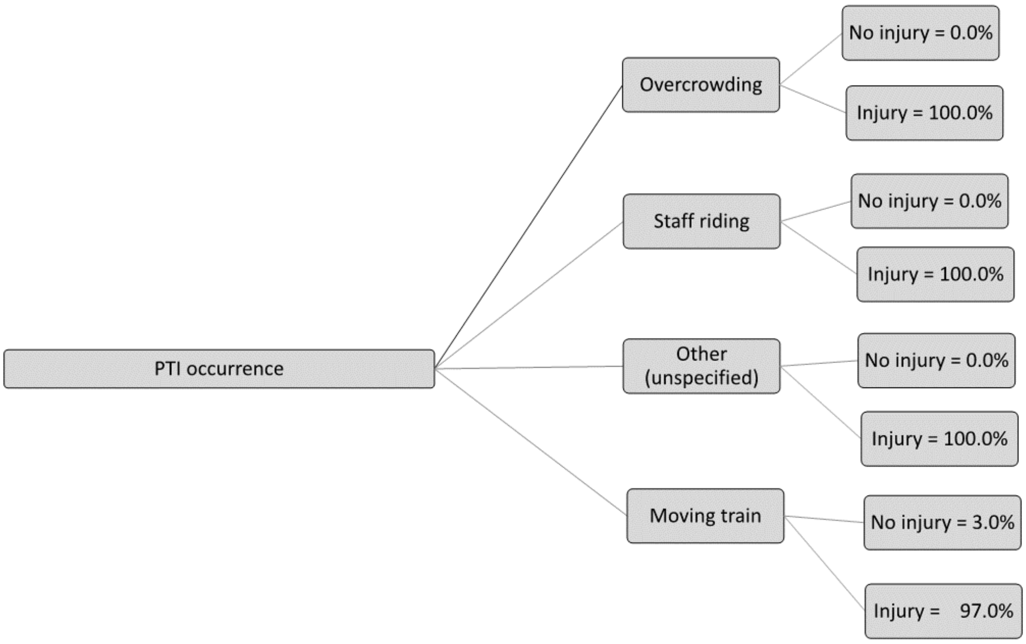

Figure 4-23: Fault tree representing the estimated probability of an injury caused by a PTI occurrence

The results from the regression analysis yielded the probability of an injury occurring in the event of an occurrence to be 99.2% on average. This result can be attributed to the fact that only incidents that resulted in an injury were most likely to be recorded. When considering occurrences as a result of overcrowding, staff riding and other unspecified causes, only incidents that resulted in an injury were recorded. However, considering incidents as a result of the moving train both injury and non-injury incidents were recorded.

To compare the difference in operations between the different train types, the probability of operation disruptions in the case of the 5M2A (yellow) trains was taken to be 99.2% and the effect of operations disruption was that the capacity of the line was reduced by half due to the delays caused. In the case of operations with the 10M4 (blue) trains, the probability of disruptions was taken as 50% because incidents as a result of staff riding and the moving train were assumed to be impossible due to the train design. However, the probability of an injury occurring as a result of overcrowding or other causes was unknown because of the underreporting of non-injury occurrences as a result of these two causes. The effect of operation disruptions was also assumed to result in half the capacity of passengers being catered for.

The line capacity is dependent on the available spaces in the coach, the total number of coaches forming one train, the flow rate of passengers boarding and alighting, the minimum average headways, reliability of the service and the train operation speed. The assumptions made for each train type are presented in Table 4-13.

Table 4-13: Assumptions made in capacity calculation

		Class 10M4	**Class 5M2A**			**Source of information**
		Rissik station	**Rissik station**	**Walker street station**	**Loftus station**	
T_o	**Doors opening**	2	3.4			*Observation surveys
	PTI crossing delay	0	0	0.5	1	
τ_a **(pass/sec)**	**Morning**	1.950	1.950	1.917	-	
	Afternoon	1.697	-	-	-	
τ_b **(pass/sec)**	**Morning**	0.938	-	1.917	-	
	Afternoon	2.133	2.708	-	2.400	
Vehicle capacity	C_s **(sps/coach)**	224	150			*PRASA
	No of Train Units (TUs)	8	12			*PRASA
Operating speed	V_0**(km/h)**	120	90			*PRASA
Loading factor (α_{max})	**Morning**	0.40				*PRASA travel chart
	Afternoon	1.00				
Reliability	**R (%)**	60				*Rissik station study
TU frequency per hour	f_{max}	2.00	2.00	1.00	1.00	*Assumed operation under ideal/stipulated timetable (based on Rissik station study)

Equation 4-9 was applied in the capacity calculation (Vuchic, 2005).

$$C\left[\frac{passengers}{hour}\right] = \alpha_{max} \cdot f_{max} \cdot C_{TU} = \alpha_{max} \times \frac{3600}{Max\ h_{min}} \times C_v \cdot n_{max} \qquad \text{Eq.4-7}$$

The values in Table 4-13 were applied in Equation 4-9 in the calculation of C_{TU}:

$$\frac{TUs \times 3600}{[T_o + (\tau_a \times no. of\ passengers\ per\ channel) + (\tau_b \times no. of\ passengers\ per\ channel) + (f_{max} \times 3600)]}$$

The expected capacity values for each scenario is presented in Table 4-14 with the corresponding probability of operations disruption.

Table 4-14: Expected value of capacity (passengers/hour) along the Pretoria-Pienaarspoort corridor per scenario.

		Train type	Operation share	Capacity (pass/hr)	Probability of disruption	Expected values (pass/hr)
Scenario 1	(non-injury)	Yellow	100%	1764.08	0.80%	2648.13
	(injury)			882.04	99.2%	
Scenario 2	(non-injury)	Blue	100%	1756.91	50%	2637.37
	(injury)			878.46	50%	
Scenario 3	(non-injury)	Yellow	20%	1764.08	0.80%	2638.52
	(injury)			882.04	99.2%	
	(non-injury)	Blue	80%	1756.91	50%	
	(injury)			878.46	50%	

The results in Table 4-14 show that the highest expected capacity is obtained with the mixed operation configuration, however the effect of the probability of a disruption comes into account when considering operations over time. To depict this effect a Monte Carlo simulation was carried out over a period of 300 hours to determine the change in capacity over time while taking into account the issue of safety. The results are presented in Figure 4-24.

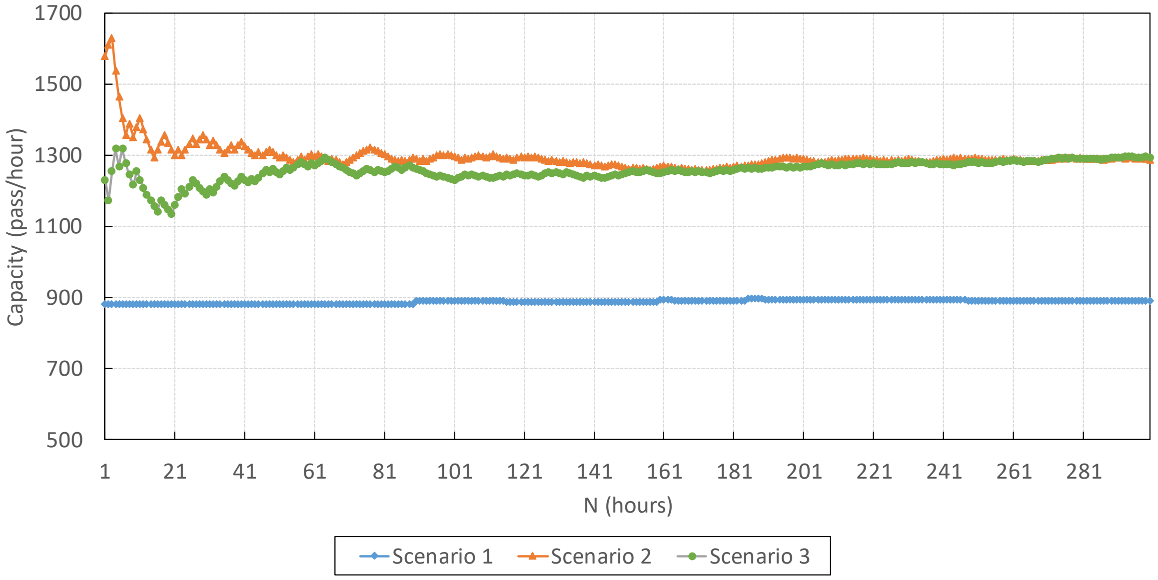

Figure 4-24: Train capacity in passengers per hour calculated for the different scenarios considered in the simulation

The results in Figure 4-24 show the detrimental effect of unsafe conditions over time on the capacity of the service, consequently showing the effect of unsafe conditions on the efficiency of the service. These results emphasise the importance of ensuring safety at station platforms, not only for the commuters but also for the benefit of effective operations.

The low probability of an injury associated with Scenario 2 resulted in the highest average corridor capacity, thus indicating the potential benefits of improving the overall safety of the service by upgrading the train design to ensure that occurrences caused by staff riding and a moving train are eliminated. An increase in capacity is the first step in ensuring the sustainability in the increased ridership that would result due to the improvement in the safety of the service. This perceived benefit to the service ties back with the argument made by Martens and Di Ciommo (2017) that was centred in the inadequacy of the elements considered in the measurement of the cost-benefit analysis. It was argued that the cost-benefit analysis as traditionally employed in practice is primarily a systematic method of measuring benefits and costs using the single denominator of money. They concluded that it would be more accurate and beneficial to measure the benefits of a transport system using the method of measuring accessibility gains. In this way persons who were initially excluded from a transport service would have higher accessibility gains than those who had access to the service, thus motivating the benefits of implementing transport interventions that increase accessibility to commuters and not just revenue.

This argument is especially relevant when considering the Metrorail service because the increase in capacity of the service would not necessarily result in increased revenue to the service provider, because the service was designed to be low priced to ensure service accessibility to the lower income earners. Therefore, it would be more effective to measure the benefits of improving service safety by adopting a cost-benefit model that considered potential accessibility gains for passengers. For example, the measure of walkability to the stations, considering the safety at the PTI, convenience, and comfort levels would provide better insight into the level of accessibility of the service (especially for special needs passengers) (Litman, 2017).

4.7 DISCUSSION

This chapter outlined the data analysis carried out in the determination of the risk of PTI occurrences along the Pretoria-Pienaarspoort corridor. The process involved the determination of the statistical relationships that existed between the factors identified to have an influence on the risk of PTI occurrences and the concentration of PTI occurrences recorded.

The results obtained from analysing the historic data from the RSR showed that occurrences as a result of a Moving train accounted for 86.2% of all recorded occurrences nationwide and resulted in at least 10.2 fatalities over the four-year period (2014-2017) according to the measure of FWIs. On average, it was observed that male passengers accounted for 10% higher number of FWIs than female passengers. Furthermore, the measure of FWIs revealed that in over 94% of all reported occurrences, at least one person was injured. The results also indicated the concern of underreporting of occurrences, where only occurrences that resulted in serious injuries were recorded. This finding therefore implies that the frequency of PTI occurrences is much higher than what is actually reported.

The historical data was then represented as the concentration of PTI occurrences (i.e. PTI occurrences per 10 000 passengers) and plotted against the vertical gap size. The results in Figure 4-8 and 4-9 reveal a vertical gap size threshold within the range of 200 mm to 230 mm where the gap size visibly influenced the frequency of PTI occurrences. To further investigate this observation, multiple linear regression analyses were carried out to determine the main factors that affected the risk of PTI occurrences. The concentration of PTI occurrences was not linearly correlated to factors such as passenger gender, the day of the week or the time of day, therefore an investigation was carried out to determine if the PTI occurrence concentration was instead station dependent. A Multilinear (ML) regression analysis was then carried out between the stations (presented by a proxy binary variable) and the concentration of PTI occurrences.

The vertical gap size was also included in the regression analysis but was modified to only consider gap sizes larger than 230 mm (as per the hypothesised threshold). The results of the regression analysis revealed that increasing the vertical gap size would increase the PTI occurrence concentration. The results also provided an indication of the station risk ranking based on the PTI occurrence concentration. However, the ML prediction model was inadequate in quantifying all the factors that were believed to influence the risk of a PTI occurrence, thus motivating the calculation of the relative risk of occurrences based on all factors considered thus far (i.e. the vertical gap size, station type, passenger gender, time of day and the day of the week).

Relative risk was defined as the likelihood of a PTI occurrence based on the frequency of an occurrence caused by a specific factor relative to the entire corridor population. A Fault tree analysis was carried out in the calculation and the significance of each factor was determined and ranked using the approach developed by Liming (2010). The results indicated that Pretoria station was the highest risk station in the corridor and occurrences caused by a moving train were most prevalent in this station. The second ranking station was Koedoespoort and occurrences as a result of overcrowding were most prevalent at this station. These two stations serve as end stations, and for this reason, large volumes of passengers transfer between platforms constantly, and as proven in the analysis, the level of overcrowding increases the risk of PTI occurrences. The results also revealed that a person was most likely to experience a PTI occurrence on a Monday or Friday during the morning and afternoon peak hours, and lastly, female passengers were at the highest risk of experiencing an occurrence along the corridor.

To further quantify and describe the influence of all factors considered, the development of a Multinomial Logistic (MNL) model to estimate the likelihood of the type of PTI occurrence was carried out. This approach was similar to the method used to determine accident severity levels in traffic accident investigations. By applying a Logit based model, the identification and quantification of the effects of the different factors mentioned earlier was possible and their contribution clearly represented for each occurrence-type. The occurrence types were based on the causes identified by the RSR.

The expected value of FWIs as a result of the different causes of PTI occurrences along the Pretoria-Pienaarspoort corridor was calculated by taking the estimated number of occurrences calculated from the ML model and determining the percentage contribution of each cause as estimated from the MNL model. The results obtained from the calculation grossly underestimated the expected number of FWIs per cause when compared to the actual numbers recorded, however, the trend observed from the calculated FWIs was similar to the reported

FWIs. This finding indicates that further research is required to improve the quality of data available on incidents in order to improve the investigation process.

Lastly, a Monte Carlo simulation was carried out to hypothesise the effect of the risk of a PTI occurrence on the corridor capacity over a 300-hour period. This simulation was carried out with the aim of measuring the cost to the service provider (in the case of a PTI occurrence) and the value of implementing safety interventions in stations. The three scenarios considered were, (i) exclusive operation of the older class 5M2A trains, (ii) the exclusive operation of the newer class 10M4 trains, and (iii) the mixed operation of the two train types. The effect of a PTI occurrence was represented in each scenario by assuming that if an incident were to occur, only half the capacity would be achieved on the corridor due to the delays caused. The results from the simulation showed the potential detrimental effect on the capacity of the Metrorail service when operating in unsafe conditions. The low probability of an injury associated with the second scenario resulted in the highest average corridor capacity, thus indicating the potential benefits of improving the overall safety of the service by upgrading the train design to ensure that occurrences caused by staff riding and a moving train are eliminated.

The assumptions made in the calculations carried out in this section were based on informal unstructured interviews carried out with PRASA officials and observations made during the surveys carried out throughout the research. However, considering the available sample sizes, the assumptions made are limited and raises the concern of the widespread application of the estimated model across any corridor on the Metrorail network considering that the assumptions made were largely based on the observations along the Pretoria-Pienaarspoort corridor. Further research is required on a larger scale to obtain larger sample sizes that would result in more adequate model estimations with reduced biases. The value in determining the rate of underreporting would also result in reduced bias in model prediction, especially if the WESMLE approach is applied as proposed by Patil and Lord (2011), and Ye (2011). The improvement of the reporting process would also result in higher quality data and better model predictions can be developed. The inclusion of information such as age, train type and the level of disruption of the system would broaden the investigation, thus improving the measurement and prediction of the effects of a PTI occurrence on the whole service more accurately.

5 CONCLUSIONS AND RECOMMENDATIONS

5.1 CONCLUSIONS

The objectives of the research were met through the development of a method of calculating the relative risk of a PTI occurrence along a specified corridor based on historical data and the estimation of the expected risk using ML and MNL regression analysis. Factors such as station parameters (the size of the gap at the PTI and the level of crowding at a station), the time of day, the day of the week and the individual characteristics such as gender were used as inputs in the estimated function to calculate the expected risk of a PTI occurrence. The results from the calculations showed that:

- Increasing the vertical gap size also increased the likelihood of a PTI occurrence (especially considering vertical gap sizes larger than 230 mm).
- A person was most likely to experience a PTI occurrence on a Monday or Friday during the morning and afternoon peak hours because of the high volume of passengers traveling at those specific periods.
- The two end stations (Pretoria and Koedoespoort stations) were the highest risk stations along the Pretoria-Pienaarspoort corridor because of the large vertical gap sizes measured at the PTI at each station and the large volumes of passengers transferring between platforms constantly.
- Female passengers traveling along the Pretoria-Pienaarpoort corridor were 47 % more likely to experience a PTI occurrence than male passengers.

When considering the level of accessibility into the Metrorail service for special needs passengers, only 7% of the total population observed during the observation surveys formed part of this category. This low percentage of special needs commuters using the service was attributed to the prevalent incidents of theft, vandalism, service unreliability and inaccessibility into the train because of the large gap size. Furthermore, the results from the manual measurement of the gap sizes along the Pretoria-Pienaarspoort corridor revealed that none of the stations were accessible to PRMs (especially wheelchair users). The gap sizes were either too large for a PRM to navigate unassisted, or access to the station platform was not possible because the elevators were not operational and there were no wheelchair ramps available.

In addition, the effects of PTI occurrences were hypothesised through the calculation of the estimated corridor capacity when incorporating the PTI risk factor. It was seen that the corridor capacity over a period of 300 hours was 31 % higher with the implementation of the new 10 M4 trains that eliminated the risk of a PTI occurrence as a result of passengers staff riding or attempting to board or alight the train while it was still moving. Thus, ranking the intervention of introducing a safer train design as one of the most effective mitigation strategies for reducing the risk of PTI occurrences. However, a holistic approach to redesigning the station environment is required to effectively eliminate PTI occurrences from the rail system.

Improving service reliability and effective communication at stations would reduce crowd frustration and agitation that translates into passengers rushing to board the train on platforms and pushing their way into or out of the train because of the desperation induced from the service unreliability. In the presence of a large gap size, this unsafe behaviour increases the risk of PTI occurrences.

5.2 RECOMMENDATIONS

The following recommendations can be drawn from this research:

- Developing a standard form of occurrence reporting that accurately specifies passenger age and gender, as well as accurately recording the station and the time of occurrence would improve the process of data analysis. Also, the inclusion of the possible causes of occurrences would further improve the investigation process.

- Conducting passenger observation surveys on a larger scale and regularly would ensure the availability of adequate data on passenger behaviour. This type of data could be applied in station design planning and in determining the most effective interventions to improve passenger safety and service delivery.

- Develop methods of measuring accessibility gains when motivating the implementation of transport services or interventions that would benefit the lower income bracket. In addition to the measurement of accessibility gains, methods of defining the “real” cost of an occurrence to the victim of the occurrence should also be developed. By doing this, the benefits of safety interventions can be further motivated.

- Further research into the successful implementation of continuous gap measurements would provide data on the rate at which the gap size at the PTI changes and in the long term would provide an indication of the deterioration rate of the track structure at stations. Developing a method of continuous gap measurement would create the opportunity of adopting a proactive maintenance approach that would, in the long term, reduce operational costs for the train service provider.

- The implementation of The National Rail Policy would address the systemic shortcomings addressed in the study conducted by Hutchings (2017). These shortcomings being financial constraints, insufficient resources and inadequate training of personnel in the rail industry. By addressing these shortcomings, potential access to resources that would ensure effective investigation into safety incidents would be possible and subsequently, the implementation of essential safety interventions would be possible.

6 REFERENCES

Ala, V., 2008. *Road Safety as a Public Health Priority*. Department of Health. Available at: https://www.doh.gov.ph/sites/default/files/publications/Vol3Issue3November2008.pdf

Arduino.cc., 2018. *Arduino*. [Online]. Available at: https://www.arduino.cc/

Banner Engineering, 2019. *Ultrasonic Sensors: Answers to Frequently Asked Questions, Leader in Industrial Automation.* [Online]. Available at: https://www.bannerengineering.com/za/en/company/expert-insights/ultrasonic-sensors-101.html [Accessed 1 Sep. 2018]

Bezruczko, N. & Linacre, M. J., 2005. Measurement theory foundations. *In: Rasch measurement in health sciences.*

Carey, M. & Carville, S., 2000. Testing schedule performance and reliability for train stations. *Journal of the Operational Research Society*, Volume 51, pp. 666-682.

Čebašek, T.M., Esen, A.F., Woodward, P.K., Laghrouche, O. & Connolly, D.P., 2018. Full scale laboratory testing of ballast and concrete slab tracks under phased cyclic loading. *Transportation Geotechnics*, Volume 17, pp.33-40.

Cheng, Y.H., 2010. Exploring passenger anxiety associated with train travel. *Transportation*, 37(6), pp. 875 - 896

Daniel, J., Rotter, N. G. & Nichnadowicz, V. F., 2009. *Customer behavior relative to gap between platform and train*, s.l.: New Jersey Department of Transportation.

Department of Transport, 1999. *Moving South Africa Action Agenda 1999*. Department of Transport. South Africa.

Department of Transport, 2017. *National Rail Policy Draft White Paper 2017*. Department of Transport. South Africa.

De Villiers, W. J., 1986. Strategic planning, management practices and systems of the South African Transport Services. *Pretoria: South African Transport Services.*

Directive, R.S., 2004. Directive 2004/49/ec of 29 April 2004 on Safety on the Community's Railways and Amending Council Directive 95/18/ec on the Licensing of Railway Undertakings and Directive 2001/14/ec on the Allocation of Railway Infrastructure Capacity and the Levying of Charges for the Use of Railway Infrastructure and Safety Certification. *Official Journal of the European Union L, 220*, pp.16-39.

Fearnley, N., Leiren, M. D., Skollerud, K. H. & Aarhaug, J., 2009. Benefit of measures for universal design in public transport. *European Transport Conference*, 2009Association for European Transport (AET), Citeseer.

Fernández, R., Valencia, A. & Seriani, S., 2015. On passenger saturation flow in public transport doors. *Transportation Research Part A: Policy and Practice,* Volume 78, pp. 102-112.

Fischer, G. H. & Molenaar, I. W., 2012. *Rasch models: Foundations, recent developments, and applications*. Springer Science & Business Media

Fujiyama, T., Thoreau, R. & Tyler, N., 2014. The effects of the design factors of the train-platform interface on pedestrian flow rates. In: *Pedestrian and Evacuation Dynamics 2012.* s.l.:Springer, pp. 1163-1173.

Harris, N. G. & Anderson, R. J., 2007. An international comparison of urban rail boarding and alighting rates. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit,* Volume 221, pp. 521-526.

Holloway, C., Thoreau, R., Roan, T.R., Boampong, D., Clarke, T., Watts, D. & Tyler, N., 2016. Effect of vertical step height on boarding and alighting time of train passengers. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Volume 230(4), pp.1234-1241.

Hutchings, J., 2017. *Systemic factors in the investigation of South African railway occurrences.* (Doctoral dissertation)

Jara-Díaz, S., 2007. *Transport production and cost structure', Transport Economic Theory* (pp. 11-49). Emerald Group Publishing Limited.

Levinson, D. M., & Krizek, K. J., 2008. *Planning for place and plexus: Metropolitan land use and transport*. New York: Taylor & Francis.

Lewin, J., 1963. Road Research Laboratory. *Diatoms. Ann. Rev. Micro,* Volume 17, pp. 373-414.

Liming, Z., Jianwei, Y., Guoqiang, C. and Limin, J., 2010, February. Monte-Carlo simulation based on FTA in reliability analysis of door system. In *2010 The 2nd International Conference on Computer and Automation Engineering (ICCAE)* (Vol. 5, pp. 713-717). IEEE.

Lionjanga, N. & Venter, C., 2018. Does public transport accessibility enhance subjective well-being? A study of the City of Johannesburg. *Research in Transportation Economics,* Volume 69, pp. 523-535.

Litman, T., 2012. Evaluating accessibility for transport planning. *Measuring People's Ability to Reach Desired Goods and Activities. Victoria Transport Policy Institute.*

Martens, K. & Di Ciommo, F., 2017. Travel time savings, accessibility gains and equity effects in cost–benefit analysis. *Transport reviews,* Volume 37, pp. 152-169.

Marqués, R. and Hernández-Herrador, V., 2017. On the effect of networks of cycle-tracks on the risk of cycling. The case of Seville. *Accident Analysis & Prevention*, 102, pp. 181-190.

Mathebula, M. & Sopazi, P., 2016. Degraded modes of operation: antecedents for railway accidents. *Sixth SARA International Conference & Exhibition,* Midrand, South Africa.

Metrorail, 2018. *Metrorail.* [Online]. Available at: www.metrorail.co.za

Mondliwa, P. & Mondliwa, P., 2019. South Africa's Public Transport Market Inquiry: Integrating Modes. [Online] *CCRED - A leading university research group*. Available at: https://www.competition.org.za/review/2017/8/29/south-africas-public-transport-market-inquiry-integrating-modes [Accessed 3 Aug. 2019].

Moug, V. & Coxon, S., 2013. An analysis of station infrastructure design to improve accessibility between the platform and suburban train carriages. *Australasian Transport Research Forum (ATRF)*, 36th, 2013, Brisbane, Queensland, Australia.

Patil, S., Geedipally, S. R. & Lord, D., 2012. Analysis of crash severities using nested logit model—accounting for the underreporting of crashes. *Accident Analysis & Prevention,* Volume 45, pp. 646-653.

Politzer-Ables, S., 2016. *Interpretation of coefficients in logistic regression*. Available at: http://www.mypolyuweb.hk/~sjpolit/logisticregression.htm

Rail Safety and Standard Board, 2015. Platform train interface strategy. *Rail Safety and Standard Board. UK.* [Online]. Available at: https://www.rssb.co.uk/en/Insights-and-News/Industry-Topics/Platform-Train-Interface

Railway Safety Regulator, 2015. Technical requirements for systemic engineering and operational standards – Railway stations. *Regulator Standard*, South Africa.

Railway Safety Regulator, 2016. *State of Safety Report*. 2015/2016. Railway Safety Regulator. South Africa.

Railway Safety Regulator, 2017. *State of Safety Report*. 2016/2017. Railway Safety Regulator. South Africa.

Railway Safety Regulator, 2018. *State of Safety Report*. 2016/2017. Railway Safety Regulator. South Africa.

Road Research Laboratory, 1963. A method for measuring saturation flow at traffic signals. *Road Note 34*, Road Research Laboratory, Crowthorne

Schreiber-Gregory, D., 2018. Logistic and Linear Regression Assumptions: Violation Recognition and Control. *SAS Institute Inc.*, pp.130-150. Available at: https://www.lexjansen.com/wuss/2018/130_Final_Paper_PDF.pdf [Accessed 6 Aug. 2019].

Shimizu, M., Oizumi, J., Matsuoka, R., Takeda, H., Okukura, H., Ooya, A. & Koike, A., 2016. Development of a novel system to measure a clearance of a passenger platform. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, Volume 41.

SparkFun Electronics., 2017. *SparkFun Electronics*. [Online]. Available at: https://www.sparkfun.com/categories/23

Spanjaard, M., Reeves, N.D., Van Dieen, J.H., Baltzopoulos, V. & Maganaris, C.N., 2008. Influence of step-height and body mass on gastrocnemius muscle fascicle behavior during stair ascent. *Journal of biomechanics*, Volume 41(5), pp.937-944.

Statistics Solutions, 2018. *Assumptions of Logistic Regression*. Available at www.statisticssolutions.com/assumptions-of-logistic-regression/

Statistics Solutions, 2018. *Transforming variables to meet an assumption*. Available at: http://www.statisticssolutions.com/transforming-variables-to-meet-an-assumption/

Statistics South Africa, 2013. *Mid-year population estimates, 2013,* Pretoria : Statistics South Africa .

Tang, T.-Q., Shao, Y.-X., Chen, L. & Shang, H.Y., 2017. Modeling passengers' boarding behavior at the platform of high speed railway station. *Journal of Advanced Transportation*, Volume 2017.

Tenøy, A. & Leiren, M. D., 2008. Accessible public transport. *A view of Europe today–policies, laws and guidelines. TØI-rapport,* Volume 952, p. 2008.

Urrutia, J.D., Bobihis, S.D., Serrano, C.D., Mercado, J. & Bernardino, R.R., 2018. A logistic regression analysis on the influence of accident factors on the fatalities of road accidents in metro Manila. *Journal of Fundamental and Applied Sciences*, Volume 10(3S), pp.32-44.

Van Der Merwe, J., Malan, C. and Havenga, J., 2017. Systemic cost of risk for heavy haul operations in South Africa. *11th International Heavy Haul Association Conference (2017)*, pp 1299-1305.

Vuchic, V. R., 2017. Urban transit: operations, planning, and economics. John Wiley & Sons.

Washington, S. P., Karlaftis, M. G. & Mannering, F., 2010. *Statistical and econometric methods for transportation data analysis.* Chapman and Hall/CRC.

Wahl, A.J., 2014. Technical study report on the causes of occurrences at the platform-train interface on the Mabopane-Naledi corridor. *Railway Safety Regulator*, South Africa.

Wilson, J. R., 2014. Fundamentals of systems ergonomics/human factors. *Applied ergonomics*, Volume 45, pp. 5-13.

World Bank, 2013. Improving Accessibility to Transport for People with Limited Mobility (PLM). *A Practical Guidance Note*. May 2013. World Bank.

Ye, F. & Lord, D., 2011. Investigation of effects of underreporting crash data on three commonly used traffic crash severity models: multinomial logit, ordered probit, and mixed logit. *Transportation Research Record*, Volume 2241, pp. 51-58.

APPENDIX A

RESULTS OBTAINED FROM THE RESEARCH

APPENDIX A NOTES ON APPENDICES

A.1 INTRODUCTION

Appendix A contains the following data:

f) Data collected during manual gap size measurement along the Pretoria-Pienaarspoort corridor

g) Data collected during the observation surveys carried out along the Pretoria-Pienaarspoort corridor.

h) The results obtained from the Multinomial logistic regression

A.2 CONTENTS OF APPENDICES

Table A1 presents the data collected during manual gap size measurement along the Pretoria-Pienaarspoort corridor

					Crowding on platform									Gap size (mm)					Crowding in the train					Overall perception					Comments	Is there access to PRM in station?
Station	Type of station	Number of platforms	Platform no.	Security personnel (on platform)	1	2	3	4	5	Train no.	Train type	Door number	Time of arrival	Horizontal	Vertical	Time of departure	Delay	Security personnel (in train)	1	2	3	4	5	1	2	3	4	5		
Rissik	Through	2	2	1		1				-	blue	6	10:41	50	130	10:42						1			1				narrow platform	No, lifts not working
Loftus	Through	2	2	-	1					-	blue		10:43	0	280	10:44			1							1				No, no lifts, large V gap
Loftus	Through	2	2	-	1					-	blue		11:18	0	290	11:19			1								1			
Walker str	Through	2	2	1	1					-	blue	4	11:21	300	0	11:22			1								1		On a curve	No, no lifts
Walker str	Through	2	2	1		1				8156	blue	-	11:58	120	0	11:58					1					1				
Devenish str	Through	2	2	-	1					8156	blue	-	12:00	110	5	12:00					1					1				No, no lifts
Devenish str	Through	2	2	-							yellow	-	12:13	-	-	12:13														
Devenish str	Through	2	2	-	1					8158	blue	-	12:57	120	30	12:58				1					1				The wait was long!!	
Mears	Through	2	2	-	1					8158	blue	-	13:01	150	30	13:01		-		1					1					No, no lifts
Mears	Through	2	2	1	1					8163	blue	-	13:20	120	30	13:20		-	1								1			
Pretoria	End	3	3	hard to locate			1			8163	blue	-	13:25	0	350	13:34		-	1						1				Station seems disorderly	Yes (main entrance); however large V gap
Pretoria	End	3	1	1			1			8165	blue	-	-	0	330	13:37		1	1								1			
Mears	Through	2	1	-						8165	blue	-	13:40	130	30	13:40		1	1								1			
Devenish str	Through	2	1	-						8165	blue	-	13:42	170	30	13:42		1	1								1			
Walker str	Through	2	1	-						8165	blue	-	13:45	110	20	13:45		1	1								1			
Loftus	Through	2	1	-						8165	blue	-	13:47	50	330	13:47		1	1								1			
Rissik	Through	2	1	-						8165	blue	-	13:50	130	30	13:50		1	1									1		

					Crowding									Gap size (mm)					Crowding					Overall					Comments	Is there access to PRM in station?
Station	Type of station	Number of platforms	Platform no.	Security personnel (on platform)	1	2	3	4	5	Train no.	Train type	Door number	Time of arrival	Horizontal	Vertical	Time of departure	Delay	Security personnel (in train)	1	2	3	4	5	1	2	3	4	5		
Rissik	Through	2	1	1	1					8147	blue	-	9:46	130	30	9:47				1						1				No
Hartbeespruit	Through	2	1	1	1					8147	blue	-	9:48	30	260	9:49		-		1						1				Yes, ramp available; however large V gap
Hartbeespruit	Through	2	1	1		1				-	blue	-	10:20	50	220	10:21		-	1							1				
Koedoespoort	Through	4	3	3		1				-	blue	-	10:26	0	410	10:26		-	1							1				No, no lifts/ramps
Koedoespoort	Through	4	3	3	1					8153	blue	-	10:57	0	390	10:58		-	1						1				Bumpy ride	
Silverton	Through	2	2	1	1					8153	blue	-	10:59	60	360	10:59		-	1							1			On a curve	Yes, ramp available; however PTI large
Silverton	Through	2	2	1		1				8157	blue	-	11:57	70	360	11:58		-		1							1			
Eersterust	Through	2	2	1	1					8157	blue	-	11:59	50	260	11:59		-		1							1			No, only stairs avaible
Eersterust	Through	2	2	1	1					8159	blue	-	12:28	100	270	12:29		-			1						1			
Waltloo	Through	2	1	-		1				8159	blue	-	12:31	50	340	12:31		-			1					1				
Denneboom	Through	2	1	-				1		8159	blue	-	12:33	-	-	12:33		-			1				1					
Eerste Fabrieke										8159	blue	-	12:35								1				1				Train stopped just before the station	
Eerste Fabrieke	Through	4	3	-		1				8159	blue	-	12:37	0	300	12:46		-			1				1					
Mamelodi Gardens	Through	2	-	-				1		8159	blue	-	12:49	150	5	12:50		-				1			1				Noise, couldn't hear the announcements clearly	
Green View										8159	blue	-										1			1				Did not stop	
Pienaarspoort	End	1	1	1				1		8159	blue	-	12:54	0	320	13:08						1			1				Bumpy ride	No access to platform
Pienaarspoort	End	1	1	1		1				8159	blue	-	12:54	0	310	13:08		-	1							1				
Green View										8159	blue	-																	Did not stop	
Mamelodi Gardens	Through	2	-	-			1			8159	blue	-	13:13	130	5	13:13		-		1					1				Noise	
Eerste Fabrieke	Through									8159	blue	-	13:14	-	-	13:16													Train stopped just before the station	
Eerste Fabrieke	Through	4	2	-		1				8159	blue	-	13:17	0	330	13:18		-		1					1				bumpy ride	
Denneboom	Through	2	2	-		1				8159	blue	-	13:20	0	260	13:21		-		1						1				
Waltloo	Through	2	2	-		1				8159	blue	-	13:23	0	280	13:23		-		1						1				
Eersterust	Through	2	1	-	1					8159	blue	-	13:25	5	240	13:26		-		1						1				
Silverton	Through	2	1	-		1				8159	blue	-	13:27	0	240	13:27		-	1								1			
Koedoespoort	Through	4	2	-		1				8159	blue	-	13:28	0	390	13:28		-	1								1			
Hartbeespruit	Through	2	2	-	1					8159	blue	-	13:33	0	250	13:34		-	1								1			
Rissik	Through	2	2	1	1					8159	blue	-	13:35	120	30	13:36		-	1								1			
Loftus	Through	2	2	-	1					8159	blue	-	13:37	0	260	13:38		-	1									1		

Table A2 presents the data collected during the observation surveys carried out along the Pretoria- Pienaarspoort corridor.

Table A2: Data collected during the observation surveys

Track No. – 01- Heading East		Detraining		Boarding		Notes	No. of doors	Train no.	Dwell time (seconds)
Train time	Video number	1	2	1	2				
		Carrying luggage	Using Cell phone	Carrying luggage	Using Cell phone				
07:50	Pienaarspoort	2	0	1	0	Yellow train	2		19
08:33	Pienaarspoort	-	-	-	-	Not recorded			-
08:46	Pienaarspoort	1	0	1	0		3		-
09:21	Pienaarspoort	5	0	8	0	People standing right infront of the door	2		21
09:58	Pienaarspoort	3	0	1	1		4		23
10:48	Pienaarspoort	1	0	3	0		4		25
11:18	Pienaarspoort	2	0	2	0		4		20
11:46	Pienaarspoort	0	0	4	0		3		23
12:53	Pienaarspoort	1	0	2 (carrying child) 6	0	Student blocks camera	4		27
13:17	Pienaarspoort	0	0	1 (carrying child) 4	0	Person with child running on platform and people running into train	4		22
13:45	Pienaarspoort	0	0	1	0	Short train	3		20
14:16	Pienaarspoort	0	0	9	0	Rush hour begin	4		21
14:49	Pienaarspoort	1	0	12	0	Yellow trains One person hanging at the door Door did not open People don’t wait till train stopped	4		13
15:10	Pienaarspoort	0	0	9	1	People ran into train	4		23
15:22	Pienaarspoort	1	0	9	0	People ran into train Passenger stopped door using foot	4		27

Track No.		Detraining		Boarding		Number of Doors	Train number	Vid #	Notes
Time of arrival	Destination	1	2	3	4				
		Total	Female	Total	Female				
07:50	Pretoria East	2	0	0	0	1		48	
08:32		5	2	0	0			49	
08:46		0	0	1	0			50	
09:21		5	0	0	0			51	
09:59		0	0	1	0			52	
10:48		1	1	6	0			53	
11:18		0	0	3	3			54	
11:46		3	1	1	0	2	8157	“55	
12:52		0	0	4	1	1		56	
13:17		2	1	6(1)	3(1)	2		57	
13:46		0	0	4	1	2		58	Vid started late, possible missing data
14:15		1	0	6	3	1		59	
14:50		1	0	10	7	1		60	
15:09		1	0	1	1	2		61	
15:22		3	0	1	1	2		62	
15:52		2	0	17	8	2		63	
16:02		0	0	5	2	1		64	

16:34		3	0	18	7	2		65	

Track No.			Detraining		Boarding		Notes	Number of doors counted
Time of arrival	Train number	Video number (M2U00XXX)	1	2	3	4		
			Female looking down	Male looking down	Female looking down	Male looking down		
7:50		053	0	1	0	0	Monday Unclear	1
8:34		054	0	0	0	0	No Footage	
8:46		055	0	1	0	1		3
9:23		056	1	2	5	3	2nd door unclear	2
9:59		057	1	1	0	0		2
10:48		058	1	0	1	7		4
11:18		059	1	2	0	2		3
11:47		060	0	0	2	3		3
12:53		061	3	1	3	2		2
12:53		062	0	0	0	0	No train or data	
13:17		063	0	0	3	1		2
13:46		064	0	0	0	2		2
14:15		065	0	0	3	1		2
14:49		066	0	0	2	2	Passenger had to open one of the doors	1
15:10		067	0	0	7	2		2
15:21		068	0	1	2	9	1 Runner looked	3

Track No.			Detraining		Boarding		Notes	Number of doors counted
Time of arrival	Train number	Video number (M2U00XXX)	1	2	3	4		
			Female looking down	Male looking down	Female looking down	Male looking down		
							down into train	
15:52		069	0	3	9	3		2

Table A3: Multinomial logistic regression analysis results

TYPE OUTCOME		B	STD. ERROR	SIG.	EXP(B)
OC	Intercept	-1.755	1.412	0.214	
	[Station_risk1=0]	1.865	1.086	0.086	6.458
	[Station_risk1=1]	0^b			
	[Time_of_day=0]	-19.516	9913.503	0.998	3.344E-09
	[Time_of_day=1]	0^b			
	[Day_week1=0]	0.418	1.245	0.737	1.519
	[Day_week1=1]	0B			
	[Day_week2=0]	-1.258	0.918	0.171	0.284
	[Day_week2=1]	0^b			
	[Gender=0]	0.555	0.818	0.497	1.742
	[Gender=1]	0^b	1.412	0.214	
SR	Intercept	-4.082	1.409	0.004	
	[Station_risk1=0]	-19.277	0.000		4.249E-09
	[Station_risk1=1]	0^b			
	[Time_of_day=0]	0.079	0.778	0.919	1.083
	[Time_of_day=1]	0^b			
	[Day_week1=0]	1.305	0.840	0.120	3.688
	[Day_week1=1]	0^b			
	[Day_week2=0]	1.944	1.113	0.081	6.986
	[Day_week2=1]	0^b			
	[Gender=0]	0.518	0.629	0.410	1.679
	[Gender=1]	0^b		0.004	
OTHER	Intercept	-2.964	1.105	0.007	
	[Station_risk1=0]	0.769	0.901	0.394	2.157
	[Station_risk1=1]	0^b			
	[Time_of_day=0]	-0.044	0.779	0.955	0.957
	[Time_of_day=1]	0^b			
	[Day_week1=0]	0.417	0.716	0.560	1.518
	[Day_week1=1]	0^b			
	[Day_week2=0]	0.703	0.782	0.369	2.020
	[Day_week2=1]	0^b			
	[Gender=0]	1.629	0.619	0.009	5.101
	[Gender=1]	0^b			
*The reference category is: MT					
*b – This parameter is set to zero because it is redundant					

Table A4: Multinomial logistic regression analysis results

OBSERVED AND PREDICTED FREQUENCIES										
						frequency			percentage	
gender						Observed	Predicted	Pearson Residual	Observed	Predicted
male passenger	not friday	not monday	offpeak	high risk	OC	0	0	0	0.00%	0.00%
					SR	1	1.741	-0.666	16.70%	29.00%
					Other	1	0.56	0.617	16.70%	9.30%
					MT	4	3.699	0.253	66.70%	61.70%
			peak	med/low risk	OC	0	0.264	-0.6	0.00%	26.40%
					SR	0	0	0	0.00%	0.00%
					Other	0	0.187	-0.48	0.00%	18.70%
					MT	1	0.548	0.908	100.00%	54.80%
				high risk	OC	2	1.164	0.793	7.70%	4.50%
					SR	8	6.777	0.546	30.80%	26.10%
					Other	3	2.468	0.356	11.50%	9.50%
					MT	13	15.591	-1.037	50.00%	60.00%
		monday	peak	med/low risk	OC	0	0.206	-0.509	0.00%	20.60%
					SR	0	0	0	0.00%	0.00%
					Other	0	0.146	-0.413	0.00%	14.60%
					MT	1	0.648	0.737	100.00%	64.80%
				high risk	OC	1	0.464	0.802	8.30%	3.90%
					SR	1	1.113	-0.112	8.30%	9.30%
					Other	1	0.984	0.017	8.30%	8.20%
					MT	9	9.439	-0.309	75.00%	78.70%
	friday	not monday	offpeak	med/low risk	OC	0	0	0	0.00%	0.00%
					SR	0	0	0	0.00%	0.00%
					Other	0	0.139	-0.402	0.00%	13.90%
					MT	1	0.861	0.402	100.00%	86.10%
				high risk	OC	0	0	0	0.00%	0.00%
					SR	0	0.059	-0.25	0.00%	5.90%
					Other	0	0.066	-0.265	0.00%	6.60%
					MT	1	0.875	0.377	100.00%	87.50%
			peak	med/low risk	OC	1	0.592	0.83	100.00%	59.20%
					SR	0	0	0	0.00%	0.00%
					Other	0	0.059	-0.25	0.00%	5.90%
					MT	0	0.349	-0.732	0.00%	34.90%
				high risk	OC	0	1.31	-1.269	0.00%	18.70%
					SR	0	0.31	-0.57	0.00%	4.40%
					Other	0	0.391	-0.643	0.00%	5.60%
					MT	7	4.989	1.68	100.00%	71.30%

Table A5: Multinomial logistic regression analysis results

OBSERVED AND PREDICTED FREQUENCIES										
						frequency			percentage	
gender						Observed	Predicted	Pearson Residual	Observed	Predicted
female passenger	not friday	not monday	offpeak	high risk	OC	0	0.000	0.000	0.0%	0.0%
					SR	1	0.925	0.094	33.3%	30.8%
					Other	1	0.905	0.120	33.3%	30.2%
					MT	4	1.171	-0.202	33.3%	39.0%
			peak	med/low risk	OC	0	0.938	0.073	25.0%	23.4%
					SR	0	0.000	0.000	0.0%	0.0%
					Other	0	1.945	0.055	50.0%	48.6%
					MT	1	1.117	-0.130	25.0%	27.9%
				high risk	OC	2	0.634	-0.816	0.0%	4.9%
					SR	8	3.557	-0.347	23.1%	27.4%
					Other	3	3.935	-0.564	23.1%	30.3%
					MT	13	4.874	1.218	53.8%	37.5%
		monday	offpeak	med/low risk	OC	0	0.000	0.000	0.0%	0.0%
					SR	0	0.000	0.000	0.0%	0.0%
					Other	0	0.523	0.954	100.0%	52.3%
					MT	1	0.477	-0.954	0.0%	47.7%
				high risk	OC	1	0.000	0.000	0.0%	0.0%
					SR	1	0.124	-0.377	0.0%	12.4%
					Other	1	0.295	-0.648	0.0%	29.5%
					MT	9	0.580	0.850	100.0%	58.0%
			peak	high risk	OC	0	0.330	-0.589	0.0%	4.7%
					SR	1	0.763	0.287	14.3%	10.9%
					Other	2	2.051	-0.042	28.6%	29.3%
					MT	4	3.856	0.110	57.1%	55.1%
	friday	not monday	offpeak	high risk	OC	0	0.000	0.000	0.0%	0.0%
					SR	0	0.151	2.270	50.0%	7.6%
					Other	0	0.511	-0.829	0.0%	25.6%
					MT	1	1.337	-0.507	50.0%	66.9%
			peak	high risk	OC	0	2.098	0.711	33.3%	23.3%
					SR	0	0.479	-0.712	0.0%	5.3%
					Other	0	1.834	0.965	33.3%	20.4%
					MT	1	4.589	-1.059	33.3%	51.0%