University of Pretoria

MODELLING THE TRIP LENGTH DISTRIBUTION OF SHOPPING TRIPS FROM GPS DATA

NICOLAAS JOHANNES JONKER

A dissertation submitted in fulfilment of the requirements for the degree of MASTERS OF ENGINEERING (TRANSPORTATION ENGINEERING)

In the

FACULTY OF ENGINEERING, THE BUILT ENVIRONMENT AND INFORMATION TECHNOLOGY

UNIVERSITY OF PRETORIA

February 2016

DISSERTATION

MODELLING THE TRIP LENGTH DISTRIBUTION OF SHOPPING TRIPS FROM GPS DATA

NJ JONKER

Supervisor:	Professor Christo Venter
Department:	Civil Engineering
University:	University of Pretoria
Degree:	Masters of Engineering (Transportation Engineering)

The newly proposed approach to the calculation of bulk service contributions in Gauteng uses not only trip generation, but also Vehicle-Kilometres of Travel (VKT) generated by a development as the basis for estimating traffic impact. This presents an empirical problem as data on VKT or trip lengths, linked to specific types and sizes of developments, are scarce and difficult to measure.

Previous approaches to measuring trip lengths have used travel surveys with and without Global Positioning Systems (GPS) data. South Africa has relied only on travel surveys without GPS data to estimate trip length information. The recommended trip length information for different developments in South Africa is provided in the TMH17 document which is used in the bulk service contributions calculations. However, the trip lengths provided in the TMH17 is based on limited South African data and is supplemented by studies that were done in Florida in the United States of America.

The advances made in GPS technology over the last decade have created new opportunities that can be used to collect GPS data for travel surveys. In this research a novel approach to collecting and analysing trip length data using passive GPS loggers distributed to a sample of 726 drivers in Gauteng was tested. A stop time of 110 seconds and repeated use of road links were used to detect trip ends in the GPS data set. The shopping centre trips were extracted using Geographic Information System (GIS) data of the locations of shopping centres compared to the trip end positions. The average trip lengths to and from shopping centre were then calculated. It was found that the average trip length per shopping centre size is longer by approximately 4.8 km compared to the prescribed TMH17 average trip lengths. These results need to be confirmed with further research.

The GPS data also provided the opportunity to calculate the percentage of travel per road Class to and from shopping centres. This is important, owing to the bulk contributions calculations only using the half adjusted average trip length. This is the average trip length halved and then only using the distance travelled on roads under the jurisdiction of the municipality excluding travel on Class 4 and Class 5 roads. It was found that 43% of the trip length distance is travelled on Class 2 and Class 3 roads. The 43% was compared to the TMH17 method of reducing the half average trip length to estimate the halve adjusted average trip length. The 43% was found to give similar results than the TMH17 method.

Owing to the significant difference in average trip lengths between TMH17 and the GPS data results an alternative method of estimating average trip lengths was proposed. It was proposed that average trip lengths be estimated based on shopping centre type and not Gross Leasable Area (GLA). It was also proposed that the 43% reducing factor be used instead of the TMH17 method of estimating the halve adjusted average trip length as the 43% reducing factor is far less complicated. The proposed alternative method is subjected to further research and confirming of the average trip lengths results.

ACKNOWLEDGEMENT

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

a) This dissertation is based on a research project of the University of Pretoria. Permission to use the material is gratefully acknowledged. The opinions expressed are those of the author and do not necessarily represent the policy of the University of Pretoria.
b) The University of Pretoria for the provision of the GPS data.
c) The South African National Road Agency for sponsoring original data collection.
d) The following persons are gratefully acknowledged for their assistance during the course of the study:
    i. DJ Jackson
    ii. E Cronje
e) Professor C Venter, my supervisor for his guidance and support.
f) My wife for her encouragement and support during this study.

Table of Contents

List of Tables

List of Figures

List of Equations

Appendixes

Glossary

A	Size of the urbanised area (km^2)
$AADT_D$	AADT trip generation rate (trips per day per size unit)
A_D	Size of the land use rights in appropriate size units
C_A	External road contribution
C_{AH}	Strength component for total development (total rights)
C_{AHD}	Strength component for a particular land use
C_{AQ}	Capacity component for total development (total rights)
C_{AQD}	Capacity component for a particular land use
C_B	Boundary road contribution
CoJ	City of Johannesburg Metropolitan Municipality
CoT	City of Tshwane Metropolitan Municipality
COTO	Committee of Transport Officials
C'_{AH}	Strength component for existing land use rights
C'_{AQ}	Capacity component for existing land use rights
DOT	Department of Transport
E_{HD}	Average number of E80 axles per heavy vehicle
EMM	Ekurhuleni Metropolitan Municipality
F_{LA}	Parameters of the formula
F_{LB}	Parameters of the formula
F_{QD}	Traffic factor to convert AADT to an impact trip rate
F_T	Adjustment for size of the municipality
GIS	Geographic Information System
GLA	Gross Leasable Area
GPS	Global Positioning Systems
K_B	Cost of one kilometre of boundary road to nominal standards
L_{45}	Length of travel on Class 4/5 roads (subject to 50% reduction)
L_B	Length of boundary road segment (km)
$L_D/2$	Half average trip length (km) on external roads only
LL	Link-to-Link method
$L_T/2$	Half total average trip length from origin to destination
L_T	Total average trip length from origin to destination
P_B	A factor which is either 0.5 or 1
P_{HD}	Proportion of heavy vehicles (of $AADT_D$)
P_N	Proportion travel on roads not under jurisdiction of municipality

PP	Point-to-Point method
PPMCC	Pearson Product-Moment Correlation Coefficient
R_H	Cost rate per E80-km/hour for the strength component
RMS	Root Mean Square
R_Q	Cost rate per veh-km/hour for the capacity component
SACSA	South African Council of Shopping Centres
Stats SA	Statistics South Africa
T_D	Impact trip rate (trips per hour per size unit)
USA	United States of America
VKT	Vehicle-Kilometres of Travel

1. Introduction and Background

Bulk services contributions are the contributions paid by developers to the municipality for the funding of engineering services required as a result of their development. The newly proposed approach to the calculation of bulk service contributions for roads in Gauteng uses not only trip generation, but also Vehicle-Kilometres of Travel (VKT) generated by a development as the basis for estimating traffic impact. This presents an empirical problem as data on VKT or trip lengths, linked to specific types and sizes of developments, are scarce and difficult to measure.

The traffic impact of a development will be incorrectly estimated, if inaccurate values are used for the average trip lengths. The consequence of underestimating the impact is that the contribution paid to the municipality will not be sufficient to fund the required improvements on the road network. If the traffic impact is overestimated, the contribution that should be paid to the municipality would be excessive. The additional contribution as a result of over estimating the impact could result in some developments not being feasible as the contribution for roads can be a significant amount of several million Rand.

One of the fundamental principles of engineering service contributions is equity and fairness. Thus the determination and utilisation of engineering service contributions should be equitable and fair to all involved (COTO, 2012). For the road contributions determination to be equitable and fair the best relevant available data should be used to estimate values (average trip lengths, etc.) required in the contribution calculations.

Previous approaches to measuring trip lengths have used travel surveys with and without Global Positioning Systems (GPS) data. South Africa has relied only on travel surveys without GPS data to estimate trip length information. The recommended trip length information for different developments in South Africa is provided in the TMH17 document, South African Trip Data Manual, which is used for the bulk service contributions calculations. However, the trip lengths provided in the TMH17 is based on limited South African data and is supplemented by studies that were done in Florida in the United States of America.

The advances made in GPS technology over the last decade have created new opportunities that can be used to collect GPS data for travel surveys. In this research a novel approach to collecting and analysing trip length data using passive GPS loggers distributed to a sample of 726 drivers in Gauteng was tested. The shopping centre trips were extracted and verified using Geographic Information System (GIS) data of the locations of shopping centres. This dissertation explains the data collection method and data properties.

The main objective of this dissertation is to provide an initial trip length distribution model calibrated on the data for different sizes of shopping centres. This could potentially be a useful method for deriving empirically validated trip length distributions for use in the calculation of bulk service contributions in Gauteng.

1.1 Objectives of the Study

The objectives of the study will be the following:

- To develop and test a method for extracting trip length data from vehicle-based GPS data;
- To estimate trip length distributions for different sizes and types of shopping centres;
- To examine and compare shopping centre average trip lengths of different user types, e.g. by gender and income level, and
- To examine the implications of the derived trip length distributions for calculation of bulk service distributions in South Africa and propose improvements to the procedure in light of the findings.

1.2 Scope of the Study

The scope of the study is defined as, determining the average shopping centre trip length per shopping centre size within the Gauteng province based on the collected GPS data in November 2011 to March 2012. The results of the average trip length estimated from the GPS data can then be used for the bulk service contribution calculations which use a halved adjusted average trip length.

It should be noted that only trips made to identifiable shopping centres were considered, excluding stand-alone shops. This was done to prevent possible non-shopping-centre trips to be included in the shopping centre trip length calculations.

A further limitation is the fact that the research uses a limited dataset as follows:

- The GPS data of 726 drivers for approximately 4 days;
- The residential suburb of all the 726 participants, and
- The location and size of shopping centres within Gauteng.

1.3 Methodology

The GPS data are retrieved from the GPS devices. A trip end identification model runs through the data to identify the trip ends. The trip ends are then compared to the shopping centre GIS information. If a trip end is located near a shopping centre it is identified as shopping centre trip end. These shopping centre trip ends are then used to calculate the length of travel to and from the shopping centre. Trip lengths are then compared to shopping centre size.

1.4 Overview of the Dissertation

This dissertation will consist of the following chapters:

- Chapter 1 serves as introduction to the report;
- Chapter 2 contains a technical introduction based on a literature study;
- Chapter 3 describes the methodology and observations undertaken during study;
- Chapter 4 describes the analysis and results;
- Chapter 5 discusses the implications of the results;
- Chapter 6 contains the conclusion and recommendations of the study;
- List of references, and
- Appendixes.

2. Literature Study

2.1 Introduction

This chapter discusses the existing literature available on the new proposed method of calculating bulk contributions and the average trip length variable used in these calculations for shopping centres. Because trip length data linked to specific types and sizes of developments are scarce and difficult to measure, the use of GPS loggers to measure trip lengths was researched as well.

2.2 Bulk Contributions for Engineering Services

Bulk contributions are the contributions paid by developers to the municipality for the funding of engineering services required as a result of their developments. The engineering services for which bulk contributions are compulsory in South Africa can be listed as follows: water, electricity, sewerage, stormwater and roads. An amount for each engineering service is calculated by the municipality (COTO, 2012).

The recommended formulas according to COTO 2012 to calculate the bulk contribution for road infrastructure is shown in Equation 2-1 to Equation 2-7:

Equation 2-1: Total Contribution for Roads

Total contribution for roads = $C_A + C_B$

Where:

C_A = External road contribution

C_B = Boundary road contribution

The equations used to calculate External road contribution (C_A) and Boundary road contribution (C_B) are shown in Equation 2-2 and Equation 2-3. The External road contribution (C_A) is influenced by the average trip length variable, because C_{AQ} and C_{AH} are both functions of the average trip length. The Boundary road contribution (C_B) is not influenced by the average trip length variable.

Equation 2-2: Basic External Road Contribution (C_A)

$C_A = (C_{AQ} – C'_{AQ}) + (C_{AH} – C'_{AH})$

Where:

C_A = Basic external road contribution

C_{AQ} = Capacity component for total development (total rights)

C'_{AQ} = Capacity component for existing land use rights

C_{AH} = Strength component for total development (total rights)

C'_{AH} = Strength component for existing land use rights

Equation 2-3: Boundary road contribution (C_B)

C_B = sum of P_B*K_B*L_B for boundary road segments B

Where:

C_B = Boundary road contribution

P_B = A factor which is either 0.5 or 1

K_B = Cost of one kilometre of boundary road to nominal standards

L_B = Length of boundary road segment (km)

Equation 2-4 to Equation 2-7 show where the trip length variable is used to calculate C_{AQ} and C_{AH}. The trip length (L_D) variable used in these equations is the distance travelled between two developments excluding the distance travelled on the following roads (COTO, 2012):

- Roads not under the jurisdiction of the Municipality; and
- Class 4 and Class 5 roads.

Equation 2-4: Capacity Component for Total Development (C_{AQ})

C_{AQ} = Sum of C_{AQD} for different land uses (D)

In which:

$C_{AQD} = A_D * T_D * (L_D/2) * R_Q$

Equation 2-5: Impact Trip Rate (T_D)

$T_D = F_{QD} * AADT_D$

Where:

C_{AQ} = Capacity component for total development

C_{AQD} = Capacity component for a particular land use

A_D = Size of the land use rights in appropriate size units

T_D = Impact trip rate (trips per hour per size unit)

F_{QD} = Traffic factor to convert AADT to an impact trip rate

$AADT_D$ = AADT trip generation rate (trips per day per size unit)

$L_D/2$ = Half average trip length (km) on external roads (roads not within development) only

R_Q = Cost rate per veh-km/hour for the capacity component

Equation 2-6: Strength Component for Total Development (C_{AH})

C_{AH} = Sum of C_{AHD} for different land uses (D)

Equation 2-7: Strength Component for a Particular Land Use (C_{AHD})

$C_{AHD} = A_D * AADT_D * P_{HD} * E_{HD} * (L_D/2) * R_H$

Where:

C_{AH}	=	Strength component for total development
C_{AHD}	=	Strength component for a particular land use
A_D	=	Size of the land use rights in appropriate size units
P_{HD}	=	Proportion of heavy vehicles (of $AADT_D$)
E_{HD}	=	Average number of E80 axles per heavy vehicle
$AADT_D$	=	AADT trip generation rate (trips per day per size unit)
$L_D/2$	=	Half average trip length (km) on external roads only
R_H	=	Cost rate per E80-km/hour for the strength component

The average trip length on external roads (L_D) variable is required in both formulas for calculating the Capacity component for the development (C_{AQ}) and the Strength component for the development (C_{AH}). It should be noted that the average trip length used in calculating C_{AH} is the average trip lengths of delivery vehicles only. Thus the value for L_D could differ for the calculation of C_{AQ} and C_{AH} (COTO, 2012). The bulk services contribution for roads are directly proportional to the average trip length variable (L_D) as shown in Equation 2-4 and Equation 2-7. Thus the bulk services contribution for roads can be significantly influenced by the average trip length value.

The type and location of the development will have an influence on the average trip length (L_D) variable. This dissertation focusses on the average trip length (L_D) variable for shopping centres within the Gauteng province.

2.3 Shopping Centres

A shopping centre or shopping mall is generally accepted to be a complex of shops located within one or more buildings in close vicinity of each other, with parking provided on the property for the shopping centres' customers. Noteworthy characteristics of shopping centres are discussed below.

2.3.1 Zoning for a Shopping Centre

The zoning of a property is defined as: "a category of directions regulating the development of land and setting out the purposes for which the land may be used and the land use or development rules applicable in respect of the said category of directions, as determined by the Town Planning Scheme" (City of Johannesburg, 2010). The required property zoning for a shopping centre is usually Business 1 to Business 4. These zonings contain the following primary and secondary land use rights associated with shopping centres (City of Johannesburg, 2010):

- Business purposes;
- Shops;
- Social halls;
- Restaurants;
- Car sales lot;
- Motor showrooms;
- Offices, and
- Place of amusement.

2.3.2 Trip Generation of Shopping Centres

The trip generation of a shopping centre, is the number of trips generated by the shopping centre during a peak hour (TMH17, 2013). For planning purposes, trip generation rates are used to estimate the number of trips that could be generated by a development. The trip generation rates used for shopping centres in South Africa are shown in Table 2-1 and are based on the Gross Leasable Area (GLA).

Table 2-1: Trip Generation Rates for Shopping Centres in South Africa

No	Land Use	Unit	Recommended Trip Generation Rates		
			Period	Rate per 100 m^2	Split In/Out
14	Shopping Centre	100 m^2	Week PM	224.5 $GLA^{-0.34}$	50 : 50
			Saturday	250.2 $GLA^{-0.30}$	50 : 50

Source: DOT, 1995

2.3.3 Parking at Shopping Centres

Shopping centres usually provide parking in close vicinity to the shopping centre entrances. Generally the municipalities require a shopping centre to provide 6 parking bays for every 100 m^2 GLA on the premises of the shopping centre.

2.3.4 Classification of Shopping Centres

Shopping centres can be classified based on characteristics as shown in Table 2-2. The classification is mainly based on the GLA of the shopping centres. As the size of the shopping centres increase the travel time, average radius of primary trade area, size of land and number of shops increases.

Table 2-2: Shopping Centre Classification

Classification	Size of Centre (m2) (GLA)	Number of stores	Size of land (ha)	Average radius of primary trade area	Median travel time to the centre
Convenience Centres	500 to 5 000	5 to 25	0.15 to 1.5	1 to 1.5 km	2 to 3 minutes
Neighbourhood Centres	5 000 to 12 000	25 to 50	1.5 to 3.6	1.5 to 2.0 km	4 to 9 minutes
Community Centres	12 000 to 25 000	50 to 100	3.6 to 7.5	2.5 to 3.0 km	6 to 14 minutes
Small Regional Centres	25 000 to 50 000	75 to 150	7.5 to 15.0	3.0 to 5.0 km	10 to 16 minutes
Regional Centres	50 000 to 100 000	150 to 250	More than 15.0	5.0 to 8.0 km	14 to 20 minutes
Super Regional Centres	More than 100 000	More than 250	-	More than 10.0 km	24 to 30 minutes

Source: Prinsloo, 2010

Intuitively one expects that the average trip length of a shopping centre will increase as the GLA increases. Table 2-2 supports this expectation as the travel time increases with the size of the shopping centre.

2.3.5 The Feasibility of a Shopping Centre

The feasibility of a shopping centre depends on whether a shopping centre is within reach of potential customers and can attract the required number of customers. A shopping centre needs to meet a particular threshold of sales to be feasible. Retail sales are influenced by the spatial distribution of demand and supply (Borgers 2011). Other factors influencing sales are: accessibility, visibility, shopping centre size and placement of retailers within the shopping centre (Borgers 2011). The consumer’s decision of where to go for his shopping, is not only influenced by the shops at a particular retail development or the travel distance but is also influenced by the access and parking arrangements. A retail development which is easy to access and has more than enough parking, is more attractive to the consumer than a retail development which is difficult to enter and has limited parking (Reimers 2013).

The attraction of a sufficient number of customers is critical for a shopping centre's feasibility. Larger types of shopping centres require a larger number of customers to be feasible. This leads one to expect larger shopping centres to have longer average trip lengths.

2.4 Shopping Centre Trips and Trip Lengths

A trip can be defined as the movement from an origin to a destination. Trip chains comprise of many individual trips with the last destination being the original starting point (Banks, 2002). An example of a trip chain is traveling from home to work to shops to home. Primary trips are trips made with a specific purpose of visiting a trip generator. Primary trips consist of two trips, one trip to the desired destination and the next trip back to the origin of the first trip (Stover, 2002). Shopping centre trips can occur within trip chains or as primary trips.

A trip length is the distance travelled between an origin and destination. Limited data on trip lengths for shopping centres are available. The available information regarding trip lengths are discussed in more detail below.

2.4.1 TMH17 Trip Lengths

The TMH17 recommends Equation 2-8 to Equation 2-10 with the values shown in Table 2-3 and Table 2-4 to calculate the average trip lengths for shopping centres. These average calculated trip lengths are specifically used for the bulk services contribution calculations as discussed in section 2.2.

Equation 2-8: Half Adjusted Average Trip Length ($L_D/2$)

$$(L_D/2) = F_T * [(1 - P_N) * (L_T / 2) - L_{45}]$$

Where:

$L_D/2$	=	Half adjusted average trip length (km)
$L_T/2$	=	Half total average trip length from origin to destination
F_T	=	Adjustment for size of the municipality
P_N	=	Proportion travel on roads not under jurisdiction of municipality
L_{45}	=	Length of travel on Class 4/5 roads (subject to 50% reduction)

Equation 2-9: Trip Length Urban Size Adjustment Factor

$F_T = 1 - F_{LA} * e^{-A * F_{LB}}$

Where:

F_T = Adjustment for size of the urbanised area

A = Size of the urbanised area (km^2)

F_{LA}, F_{LB} = Parameters of the formula

Table 2-3: Trip Length Urban Size Adjustment Parameters

Adjustment Parameter	Parameter
Parameter F_{LA}	± 0.500
Parameter F_{LB}	± 0.050

Equation 2-8 calculates half the adjusted average trip length. Only half of the L_T variable is taken into account since it is reasoned that the other half of the trip length is the responsibility of the (origin or destination) other development. This equation adjusts the half average trip length further with variables F_T, P_N and L_{45}. F_T adjusts the half average trip length with a factor to account for the urbanised area of the municipality. Trip lengths can be expected to be shorter within smaller urbanised areas compared to larger urbanised areas. Equation 2-9 shows how to calculate F_T and Table 2-3 gives the ranges of the F_{LA} and F_{LB} parameters. These parameters are set by the municipality within the given ranges. P_N decreases the half average trip length with a percentage to allow for travel on roads not under the jurisdiction of the municipality. L_{45} decreases the half average trip length to account for travel on Class 4/5 roads. Kilometres travelled on Class 4/5 roads are excluded from the bulk services contribution as these roads are either provided by the developer or by previous developers as boundary or internal roads (COTO, 2012).

Table 2-4 provides average trip length (L_T) values for different retail developments. Shopping centres have an average trip length of 10 km which needs to be adjusted with Equation 2-10 for the size of the shopping centre.

Table 2-4: Trip Lengths from the TMH17 for Retail Developments

Retail Developments	Size Units	Total Trip Length (L_T) (km)	Half Trip Length (km)	Class 4/5 Half Trip (km)	Non-Municipal (%)	Adjusted Trip Length	Factor A	Factor B
Building materials	100 m^2 GLA	8.00	4.00	1.00	40%	1.40		
Hardware and Paint store	100 m^2 GLA	7.00	3.50	1.00	40%	1.10		
Nursery (Garden Centre	100 m^2 GLA	6.50	3.25	1.00	30%	1.27		
Shopping Centre	**100 m^2 GLA**	**10.00**	**5.00**	**1.00**	**40%**	**2.00**	**0.740**	**148 000**
Bulk Trade Centre	100 m^2 GLA	10.00	5.00	1.00	50%	1.50		
Motor dealership	100 m^2 GLA	6.50	3.25	0.75	40%	1.20		
Furniture Store	100 m^2 GLA	8.00	4.00	1.00	40%	1.40		

Source: TMH17, 2013

The size adjustment factor is calculated with Equation 2-10 with the values for variables A and B as are shown in Table 2-4. The average trip length (10 km) is multiplied with the adjustment factor to provide an average trip length for a particular shopping centre (TMH17, 2013). The average trip lengths given in the TMH17 are based on South African and United States of America (USA) data. Owing to the insufficient amount of local average trip length data, the data was supplemented with studies done in Florida in the USA (TMH17, 2013). How applicable the USA data are to the South African environment is unclear, as no specific validation studies seem to have been performed locally.

Equation 2-10: Size Adjustment Factor for Developments

Size adjustment factor = 1 – (A / (1 + (GLA / B))

The size adjustment factor (Equation 2-10) was used to draw Figure 2-1 which shows the average trip length for shopping centres according to the TMH17. The TMH17's average trip lengths increase as the shopping centre's GLA increases. However, the size adjustment factor effectively decreases the average trip length to a maximum of less than 7 km, even for the largest shopping centres. Thus the 10 km value shown in Table 2-4 is not used in practice.

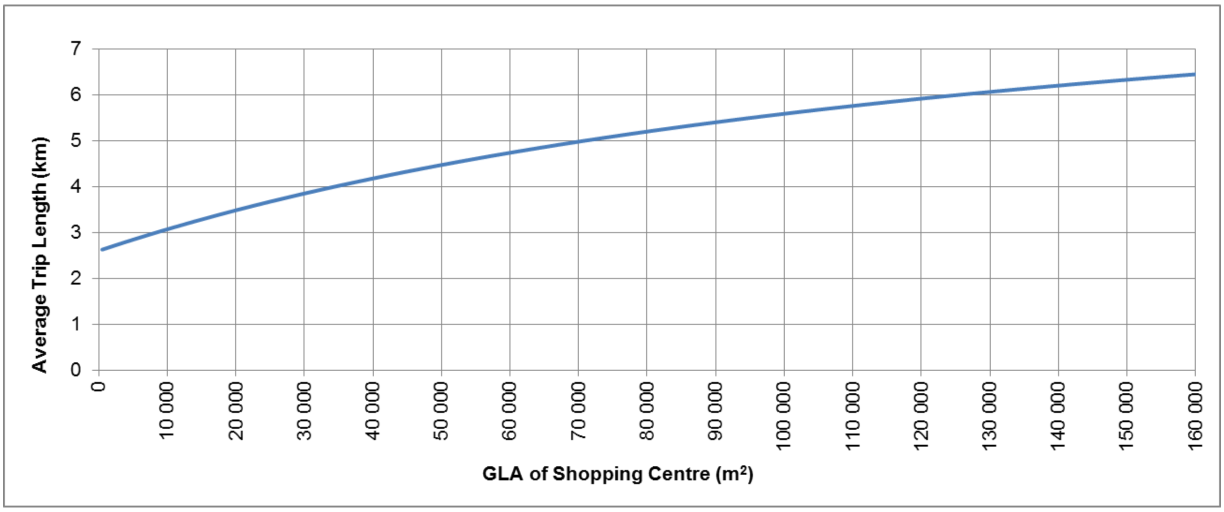


Figure 2-1: TMH17 Average Trip Lengths for Shopping Centres (L_T)

2.4.2 Ekurhuleni Metropolitan Municipality Trip Length Data

Table 2-5 shows the trip length data for the Ekurhuleni Municipality within Gauteng. The average trip length to a shop is estimated to be 14.27 km according to this table. It should be noted that this table does not differentiate between shop sizes and the data source used for setting up this table is not clear from the source document (TTT Africa, 2008).

Table 2-5: Average Trip Lengths (km) on Roads in Ekurhuleni

Development Type	No of Trips	Municipal Cl.1-3	Provincial Cl.1-3	National Cl.1-3	Total Cl.1-3	USA Total	Municipal Cl. 4	Total Cl.1-4
Any to Education	27 883	0.92	0.96	0.39	2.27		3.14	5.41
Education to Any	13 338	2.36	1.52	1.34	5.22		3.48	8.7
Any to Other	3 331	3.42	4.53	3.31	11.26		6.99	18.25
Any to Shop	2 868	2.91	3.19	2.24	8.34	7.27	5.93	14.27
Any to Work External	22 302	2.42	8.08	13.04	23.54		5.37	28.91
Any to Work Internal	81 796	2.46	3.67	3.49	9.62	9	4.45	14.07

Source: TTT Africa, 2008

2.4.3 Simulated Trip Length Data from the Netherlands

An alternative way of obtaining trip length data are from activity-travel simulation models. Timmermans used an application of Albatross, an agent based model, in the city of Rotterdam, the Netherlands. This model simulates households' and individuals' activities for an entire day, including shopping, work and recreational activities. This model provides locations of where these activities will take place and what modes of transport are used to get to these activities. The simulation is based on an individual's socio-demographic profiles (Timmermans, 2013). The results of the model focusing on shopping indicators are shown in Table 2-6. The average trip length to multi-store shops (similar to the shopping centres considered in this study) was found to be 5.5 km, which is shorter than trip lengths in the USA and South Africa, probably due to the more compact land use in the Netherlands.

Table 2-6: Albatross Model for City of Rotterdam - Shopping Indicators

Single Store Shopping								
Travel Distance (km)			**Travel Time (Minutes)**			**Duration of shopping (Minutes)**		
Average	CV (%)	Confidence interval	Average	CV (%)	Confidence interval	Average	CV (%)	Confidence interval
3.51	1.93	3.38-3.65	10.72	0.95	10.52-10.92	35.10	0.80	34.54-35.65
Multi Store Shopping								
5.50	3.93	5.07-5.92	12.92	2.20	12.37-13.49	62.14	1.30	60.56-63.73

Source: Timmermans, 2013

The Table 2-6 also suggests that for multi-store shopping trips people are willing to travel further and longer than for single-store shopping.

2.4.4 Household Travel Surveys

Household travel surveys could be a source of data to estimate average trip lengths, however, the level of accuracy is a significant short coming. A household travel survey is a survey conducted to obtain data on the travel habits of individuals within these households during a given period (DOT, 2005a). Trip diaries are most commonly used in household travel surveys to report trips made by each member of the household. According to Stopher trip diaries reports 20% to 30% less trips than the number of trips actually made. Most of these unreported trips are short trips. The under reporting of trips is a major shortcoming of household surveys (Stopher, 2007). According to Bricka there is a correlation between trip purpose and the probability of the trip being reported in trip diaries. Important repeated trips like home to work and home to school trips are reported more accurately than a quick trip to the shop (Bricka et al., 2012). Respondents to household travel surveys are mostly unable to provide accurate data on travel times as well as on the trip lengths of their trips (Stopher, 2007).

Thus, with traditional travel surveys the number of retail trips might be significantly underreported. Furthermore, the travel time and distance travelled to retail developments are inaccurate as reported in the household travel surveys. Travel distances in travel surveys are usually surveyed using a set of options from which the participant can choose from, for example 0 km to 1 km, 1 km to 5 km, 5 km to 10 km and so on. The consequence of this surveying method is inaccurate travel distance information. Nevertheless, household travel surveys are still a source of trip length data. The trip length data gathered through previous household travel surveys are discussed below.

2.4.4.1 South Africa

South Africa conducted its first national household travel survey in 2003. Some of the results of the travel survey relating to shopping trips are shown in Table 2-7 and Table 2-8. It is shown in Table 2-7 that, most people can reach a "Food Shop" within 1 to 15 minutes and an "Other Shop" within 16 to 30 minutes. This provides some indication of travel time, but no indication of travel distance. Since these times were given for all modes of travel, it would be impractical to assume an average speed to estimate a travel distance.

Table 2-7: Travel Time to Various Services

Facility	Percentage of Households				
	1 – 15 mins	16 – 30 mins	31 – 60 mins	> 60 mins	Cannot get there
Food shop	81.6	12.3	4.6	1.3	0.1
Other shop	33.2	35.1	22.9	8.3	0.3
Traditional healer	25.8	23.4	18.6	9.7	22.5
Medical services	44.2	34.1	16.1	5.2	0.5
Post office	45.9	33.8	14.4	4.5	1.4
Welfare office	31.9	36.3	21.1	6.9	3.8
Police station	40.2	35.3	17.7	5.7	1
Municipal office	38.1	35.8	17	5.3	3.8
Tribal authority	27.2	24	16.9	7.5	24.5

Source: DOT, 2005b

As shown in Table 2-8, residents in the Gauteng province named shopping as the main purpose for their trip making. The percentages for each province add up to more than 100% because respondents may name more than one main purpose for their trips. This table clearly shows that retail developments are major trip generators in Gauteng.

Table 2-8: Main Purposes of Trips Made by Household Members, by Province

Province	% of Household Members Naming Purpose			
	Education	Shopping	Visiting	Work
Western Cape	33.4	26	19.8	41
Eastern Cape	49.2	26.9	27.7	16.5
Northern Cape	34.3	28.3	39.7	30.9
Free state	38.2	33.3	42.9	26.6
KwaZulu-Natal	46.4	24.4	20.1	22.7
North West	39.6	23.7	29.5	26.7
Gauteng	29.9	44.1	33.9	39.3
Mpumalanga	41.4	34	36.7	23.7
Limpopo	51.2	22	27	15.9

Source: DOT, 2005b

2.4.4.2 Gauteng Transportation Model

The Gauteng Transportation Model only modelled the morning peak period during which shopping centres don't generate a significant number of trips. The trip length distribution used in the Gauteng Transportation Model for shopping trips is shown in Figure 2-2. The trip length distribution is also categorised for income groups. The majority of trips for all income groups are in the 0-5 km range. The trip length distribution was based on a household travel survey data.

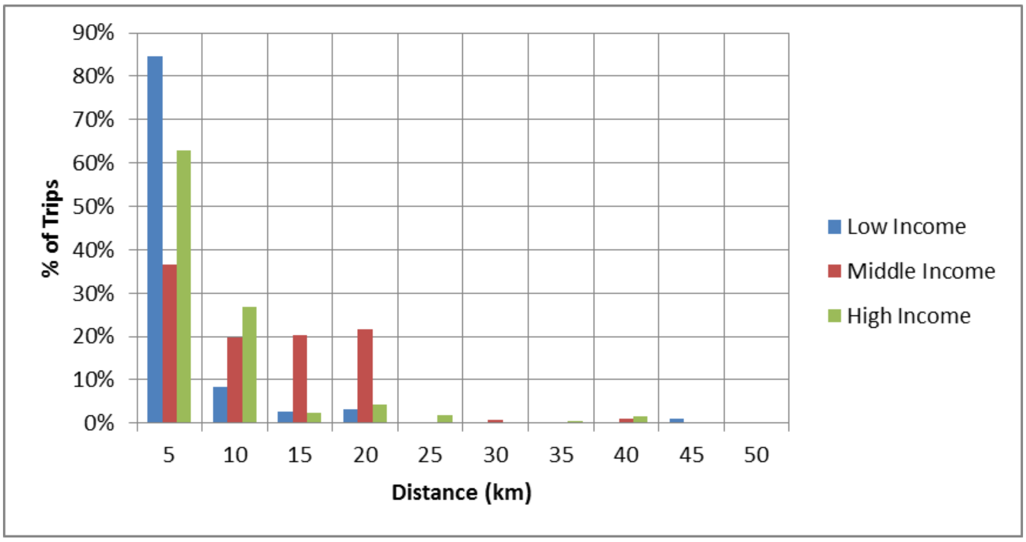


Figure 2-2: Shopping Trip Lengths per Income Group from the Gauteng Transportation Model

From the three distributions the low income group has the shortest trip lengths as a result of financial constraints. The medium income group has longer trip lengths compared to the high income group. The reason for this is that the high income group is located in well-developed areas with shopping centres in close vicinity. The medium income group is located in areas with fewer shopping centres, while still having the financial means to travel to their preferred shopping centres.

2.4.4.3 Istanbul, Turkey

A study done in 2010, showed the observed trip length distribution for Istanbul as shown in Figure 2-3. The trip length distributions for the following trip purposes are shown:

- HBW (Home Based Work trip);
- HBS (Home Based Shopping trip);
- HBO (Home Based Other trip), and
- NHB (Not Home Based trip).

These trip length frequency distributions were derived from the 2006 Household Travel Survey conducted by the Transportation Department of the Metropolitan Municipality of Istanbul.

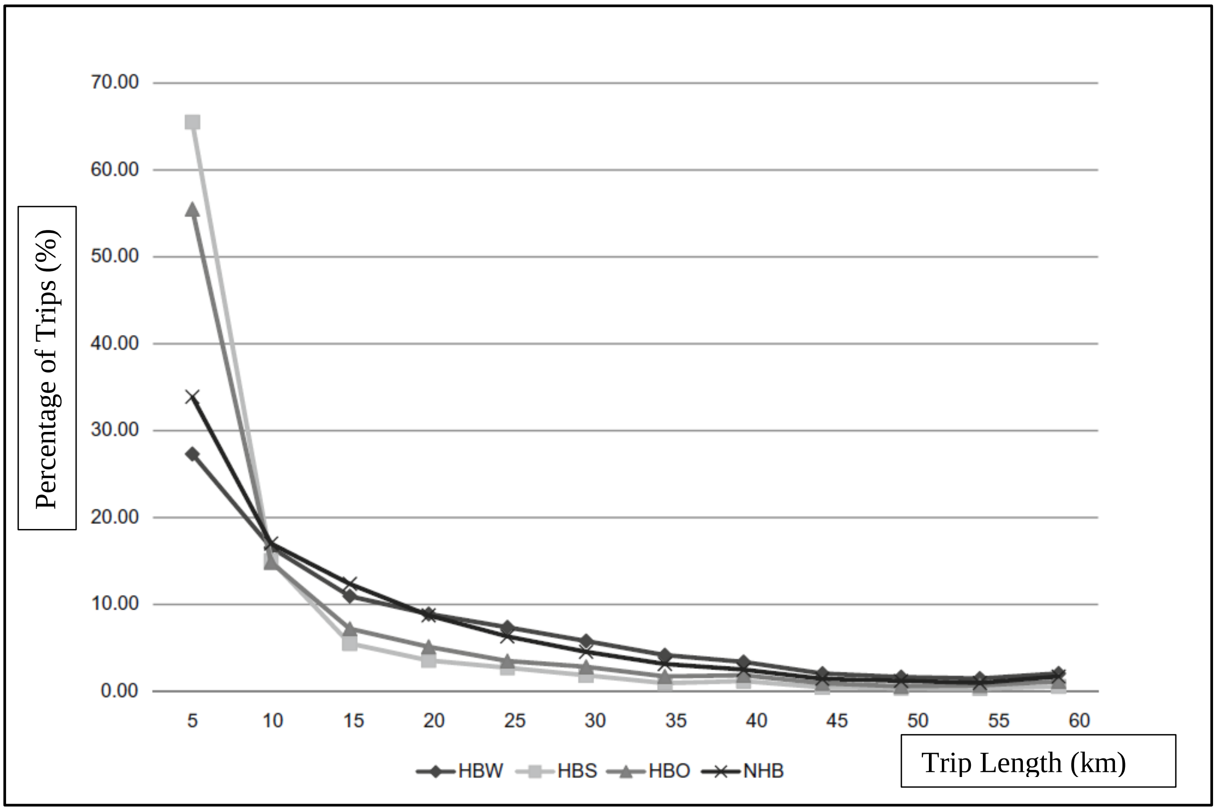


Figure 2-3: Observed Trip Length Frequency Distributions by Trip Purposes – Istanbul in Turkey (Murat, 2010)

According to Figure 2-3, approximately 85% (65% + 15% + 5%) of Home Based Shopping (HBS) trips made in Istanbul are equal to or less than 15 km. All the trip types has a similar trip length frequency distribution shapes. In the next Section 2.5, the type of distribution used for trip length frequency distributions.

2.5 Trip Length Frequency Distributions

Trip Length Frequency Distributions (TLFD) are normally presented as percentages of a trip distance, similarly to probability distributions. According to Pearson the Gamma and the Weibull distributions have the best fit for TLFDs (Pearson, 1974).

2.5.1 Distributions

2.5.1.1 The Gamma Distribution

The Gamma distribution has two parameters which defines the distribution. These parameters are the shape (α) and scale (β) parameter. The Gamma distribution function is shown in Equation 2-11. In Figure 2-4 variations of the Gamma distributions are shown with different α and β values.

Equation 2-11: Gamma Distribution (Johnson, 2005).

$$f(x) = \frac{1}{\beta^{\alpha} * \Gamma(\alpha)} * x^{\alpha-1} * e^{\frac{-x}{\beta}}$$

Where:

α	=	Shape parameter
β	=	Scale parameter
e	=	Base of natural logarithms (2.71828….)
Γ(α)	=	(α - 1) !
f(x)	=	Relative density of occurrence trip length x

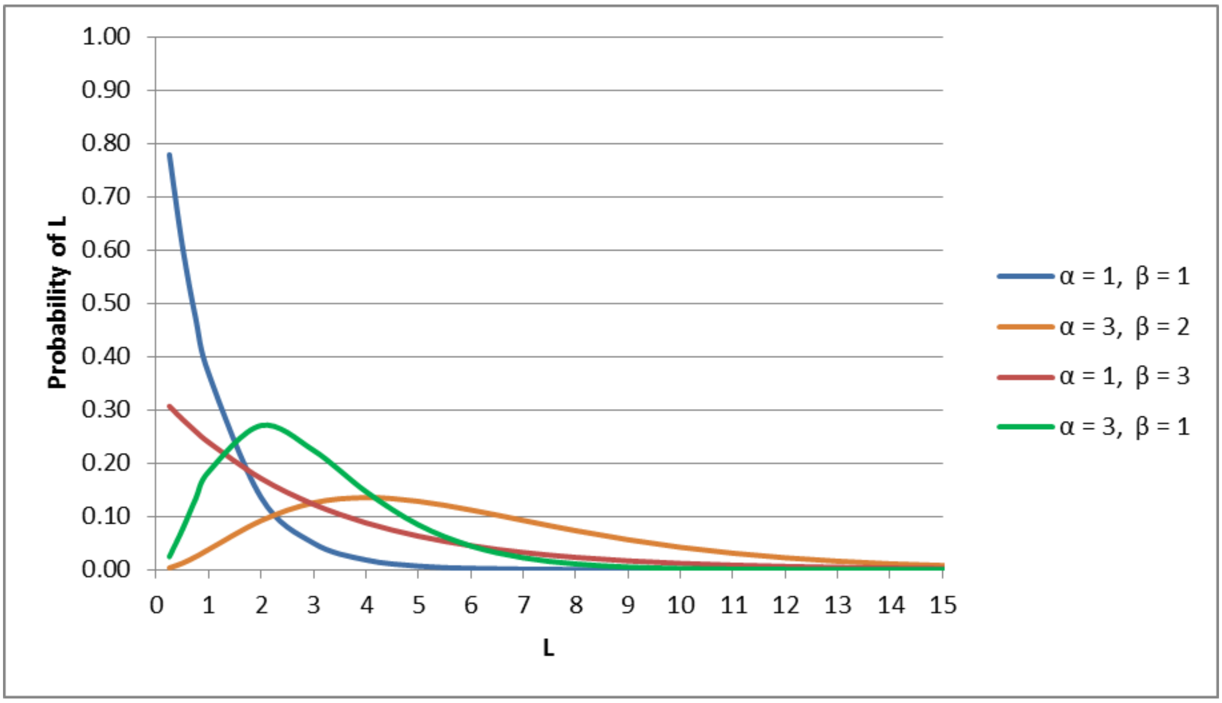


Figure 2-4: Variations of Gamma Distributions

For the Gamma distribution the mean and the variance can be calculated as shown in Equation 2-12 and Equation 2-13.

Equation 2-12: Mean of Gamma Distribution (Johnson, 2005).

$\mu = \alpha \beta$

Where:

α	=	Shape parameter
β	=	Scale parameter
μ	=	Mean

Equation 2-13: Variance of Gamma Distribution (Johnson, 2005).

$\sigma^2 = \alpha \beta^2$

Where:

α	=	Shape parameter
β	=	Scale parameter
σ	=	Standard Deviation
σ^2	=	Variance

2.5.1.2 The Weibull Distribution

The Weibull distribution also has two parameters which define the distribution. These parameters are the shape (α) and scale (β) parameter similar to the Gamma distribution. The Weibull distribution function is shown in Equation 2-14 and in Figure 2-5 variations of the Weibull distributions are shown with different parameters.

Equation 2-14: Weibull Distribution (Van As, 2008)

$$f(x) = \left(\frac{\beta}{\alpha}\right) * \left(\frac{x}{\alpha}\right)^{(\beta-1)} * e^{\left(\frac{-x}{\alpha}\right)^{\beta}}$$

Where:

α	=	Shape parameter
β	=	Scale parameter
e	=	Base of natural logarithms (2.71828….)
f(x)	=	Relative density of occurrence trip length x

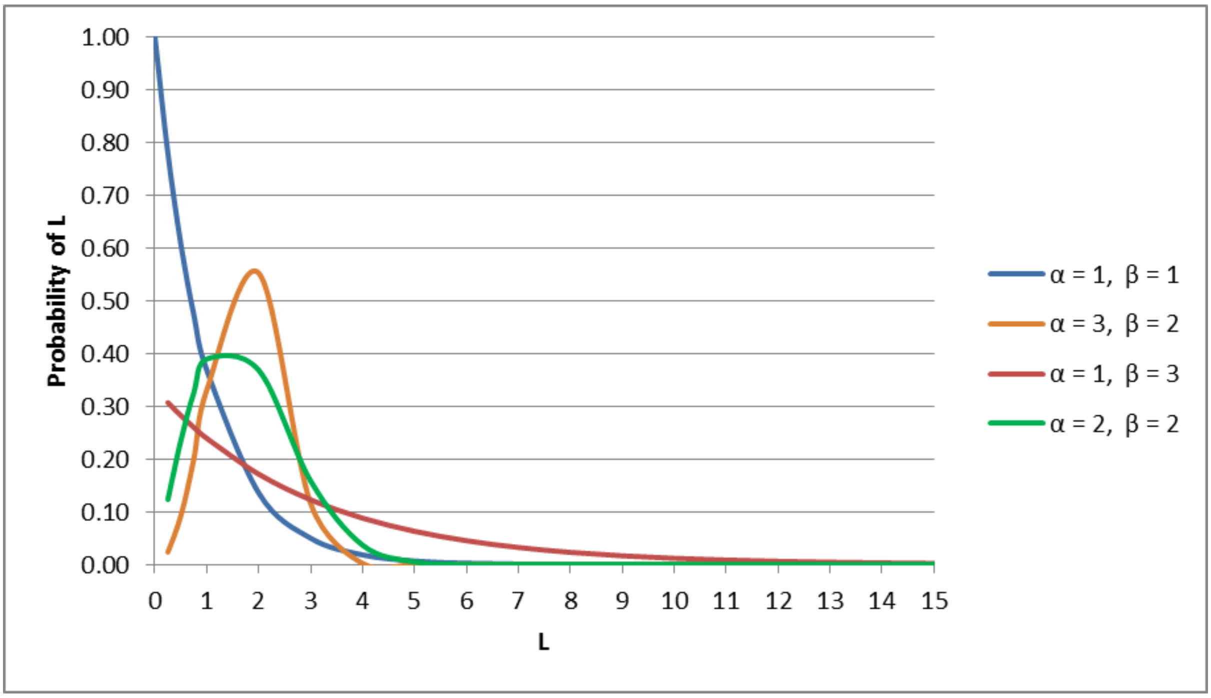


Figure 2-5: Variations of Weibull Distributions

For the Weibull distribution the mean, variance and standard deviation can be calculated as shown in Equation 2-15 and Equation 2-16.

Equation 2-15: Mean of Weibull Distribution (Van As, 2008)

$$\mu = \alpha * \Gamma(1 + \frac{1}{\beta})$$

Where:

α = Shape parameter

β = Scale parameter

μ = Mean

Equation 2-16: Variance of Weibull Distribution (Van As, 2008)

$$\sigma^2 = \alpha^2 * (\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2(1 + \frac{1}{\beta})$$

Where:

α = Shape parameter

β = Scale parameter

σ = Standard Deviation

σ^2 = Variance

2.5.2 Curve Fitting

Pearson recommended that the least squares method be used to estimate values for α and β to fit both distributions through data points. The R^2 value can then be used to evaluate the fit. The closer the R^2 value is to one, the better the fit (Pearson, 1974). Generally if the R^2 value is above 0.7 the fit is acceptable.

2.5.3 Curve Fitting through Household Travel Survey Trip Length Data

Since travel distances data is usually gathered in categories in household travel surveys (0 km to 1 km, 1 km to 5 km, 5 km to 10 km and so), some detail is lost in the data with this method. The distribution curve fitted through this data will be simplified subsequently as well. Exact trip length data is required for a comprehensive trip length distribution. The data gathering method would need to be adjusted to gather exact trip length data. This can be done with GPS technology.

2.6 GPS Devices Used in Travel Surveys

GPS devices are being used more often in travel surveys during the last decade. GPS equipment has the capability of giving precise data on an individual’s travel movements over a period of time. This makes the GPS survey method very attractive when comparing the GPS data to questionnaires or travel diaries which lacks accuracy and detail (Stopher, 2008). Since 2002 more than 25 household travel surveys conducted within the USA have used GPS supplements to assess the underreporting or misreporting by trip diaries of the travel surveys (NCHRP, 2014a). GPS data has the potential to provide accurate information on the following: number of trips made, duration of trips, when trips are made, trip lengths and routes used (Bricka et al., 2012).

2.6.1 GPS Devices

The GPS devices currently available can be divided into two categories, namely "Passive" and "Active" GPS systems. A "Passive GPS" system is switched on by the surveyor and given to the participant. The GPS requires no input from the participant (Stopher, 2008). The GPS will record the position of the vehicle or person at pre-set time intervals (for example 5 seconds or 3 seconds). From this data the traveling speed and heading can be calculated. However, most GPS devices calculate the speed and heading internally using the Doppler measurement (Shen et al., 2014). The Doppler measurement is more accurate than using the change in position because with each change in position also comes a change in error and the stored position is usually rounded to some extent compared to the actual value computed within the device.

The "Active GPS" system requires the participant to type in information about the trip he or she will be making before or after each trip. The "Active GPS" then functions in the same manner as the "Passive GPS". With the "Active GPS" system more data can be gathered from participants. However, as the typing in of information into the GPS is a tedious task, participants usually skip the task near the end of the survey or type in very vague information regarding their trips (Stopher, 2008).

2.6.2 GPS-Based Prompted-Recall Travel Surveys

This survey method uses passive GPS devices together with a follow-up survey that is based on the trips identified within the GPS data. The GPS data are used to compile an activity-travel pattern of the respondent. The respondent is presented the activity-travel pattern and prompted to confirm or reject trip ends identified through the GPS data. Trip ends can be added as well by the respondent if trip ends were missed by the trip end identification algorithm. Additional information can also be confirmed or collected from respondent for example, trip purpose, mode of travel, parking costs and vehicle occupancy. Recent GPS-based prompted recall surveys used web-based data collection platforms to retrieve the GPS data and allow the respondent to complete the follow up survey (NCHRP, 2014a).

One of the earliest prompted recall studies in Lexington used 100 households and compared their trip diaries with GPS trip data. As shown in Figure 2-6, only 34.5% of the households' trips could be matched with both sources. There were 31.0% of the household's with less trips in the GPS data as reported in the diaries and 34.5% of the households with more trips in the GPS data than in the diaries (Murakami, 1999).

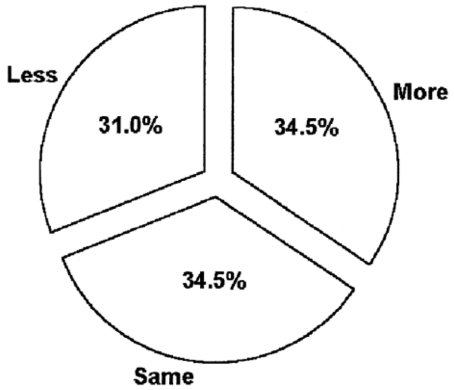


Figure 2-6: Comparison of Matched Recorded Trips with Reported Trips (Murakami, 1999)

In the same study the recall of travel distances was compared to the calculated distances from the GPS data. The graph of this comparison is shown in Figure 2-7. The straight line indicates where the recall distance is equal to the measured distance. The calculated median, 75 and 25 percentile distances from GPS data are shown according to the recall distance. It should be noted from this graph that the participants' recall distances are higher than the calculated median value for all but one instance.

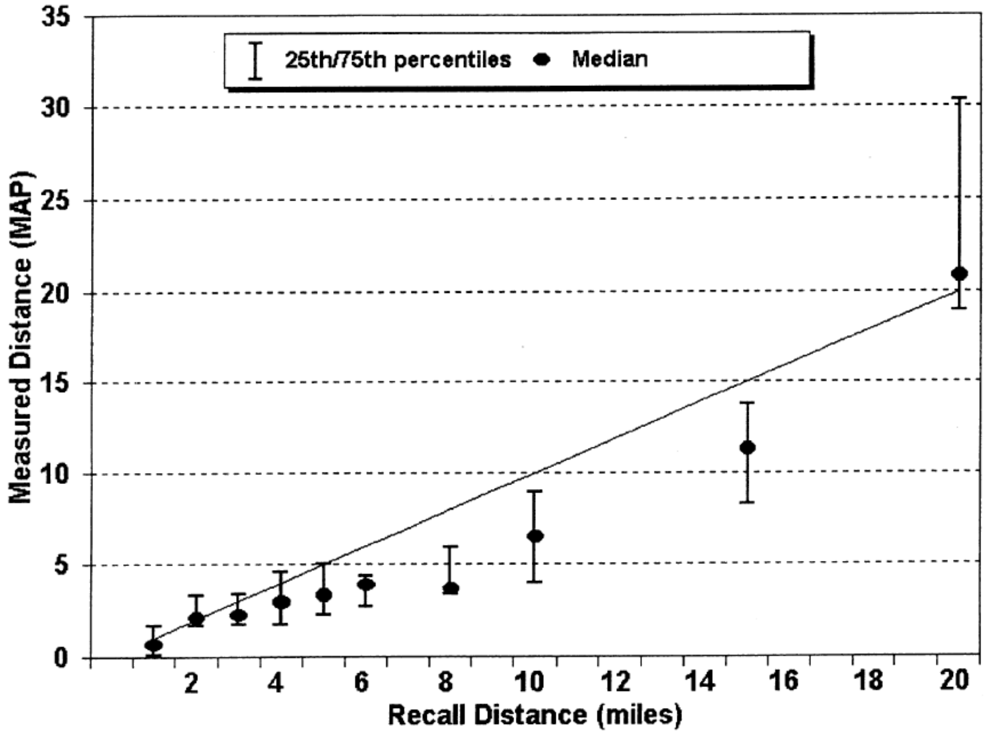


Figure 2-7: Comparison of Recall and Measured Travel Distances (Murakami, 1999)

2.6.3 Sources of GPS Data

Sources of GPS data can be divided in two categories, primary data and third party data. Primary data are GPS data collected by surveys for specific research or study. Third party data are GPS data gathered by companies for its own purposes and then sells their GPS data to research entities. In Table 2-9 the factors to consider are given when the decision needs to be made if primary or third party data will be used for research or study (NCHRP, 2014b).

Table 2-9: Collecting Compared to Purchasing GPS Data

Factor	Primary Data Collection	Third-Party Data
Cost and influencing variables	Likely Higher Cost • Sample size • Response rates • Deployment period • Languages support • Incentive amounts	Likely Lower Cost • Sample size • Geographic coverage • Spatial and granularity • Temporal span • Variables reported
Ease of data acquisition	More Labour Intensive • Requires experience in project management, field data collection management and computer programming	Less Labour Intensive • Purchase
Ease of data processing	More Intensive • Intimate knowledge of when, where and how the data were collected • Reduced uncertainty as to what processing steps need to be taken	Less Intensive • Unknown, substantial probing and guesswork required to determine the extent of processing that is necessary
Usability of data	More Certain • Biases are known • Con collect demographic characteristics of persons and households • Wider array of travel behaviour data can be collected • Sample size requirements can be monitored to ensure statistical significance	Less Certain • Limited knowledge of when, where and how data were collected • Little to no knowledge of how data were treated/processed/weighted before they were delivered to the client • Biases are unknown or have to be uncovered through investigation

Source: NCHRP, 2014b

2.6.4 Problems with GPS Data

However, GPS data has some problems of inaccurate reporting as well. The main problems that can occur with the devices are signal loss, device malfunctions, battery dies during the survey period, urban canyon effect and cold start problem. The urban canyon effect occurs when tall buildings are located on both sides of a street creating a "urban canyon", which can distort the GPS signal. The cold start problem is the time it takes the GPS to find a signal after being switched off. During the search for a signal the GPS can record several inaccurate points which could show the origin of trip at an incorrect location (Chen et al., 2010).

The data processing method used can cause some inaccuracy as well. Trip ends in the GPS data can be over - or under-estimated which leads to inaccurate number of trips reported. This is due to the complexity of determining trip ends. According to Bricka if one decides to use only GPS data as one's survey data, the study should be done with caution. The trip purpose, trip ends and mode detection algorithms are not always accurate (Bricka et al., 2012). Other challenges with GPS data are: cost, integration into existing modelling paradigms, privacy, sample bias and data management (NCHRP, 2014a).

During the planning phase of the GPS survey, mitigating measures should be implemented to limit the impact of these problems as far possible. The type of GPS survey will influence the approach and the problems that can be expected. For example, if the in-vehicle GPS is powered from the vehicles battery, a cold start will occur every time the vehicle is switched on. Compared to the GPS being carried by a person, the device will have its own power source which will limit cold starts. However, determining the trip ends and modes will be challenging.

2.6.5 Smartphones as an Alternative to GPS Devices for Household Travel Surveys

In the USA it is estimated by the Pew Research Centre that 46% of adults own a smartphone with 76% of them getting real time GPS data. This is significant number of people that could contribute to a travel survey. Smartphones are capable of running custom software applications ("apps") and this provides an opportunity to develop a software application for a travel survey. Participants in the survey can then download the application and participate in the survey. Depending on the app developed for the survey the smartphones can be used as active or passive GPS loggers (NCHRP, 2014a).

Using the participant's own smartphone addresses some common implementation challenges with GPS-based travel surveys (NCHRP, 2014a), such as:

- Eliminating the need to deliver and retrieve GPS loggers;
- Reducing the time between data collection and data review, and
- Reducing costs associated with equipment loss.

However, smartphones have their own challenges. The most prominent challenges are (NCHRP, 2014a):

- Market fragmentation (software platforms);
- Power management (Battery life);
- Data plans and associated costs;
- Self-selection, and
- Capture mode biases.

The first issue is that smartphone software platforms are fragmented. The five main software platforms are: Android, iOS, Blackberry, Windows and Symbian. This makes the development of an app for a travel survey costly and difficult. The second issue of power management of the smartphone is that the participant has to be able to use his phone as normal while logging GPS data. Currently the batteries of smartphones get rapidly depleted when logging GPS data constantly. This discourages people to participate in these surveys. The impact on the battery life can be reduced by using Wi-Fi/network location as an alternative to GPS in urban areas (Woehrle, 2013). However, this requires a dens existing Wi-Fi/network in the urban area to be successful. The third issue is that the GPS data needs to be uploaded to a server from the smartphone via the internet and this could have cost implications for participants depending on their data plans. The fourth issue is that travel surveys are normally done per household and not per individual. Not all household members will probably have a smartphone, which results in the need for normal GPS devices again. The last issue is that a smartphone sample will probably be biased (income group, age and gender) (NCHRP, 2014a).

2.6.6 Determining Trip Ends in Large Passive GPS Data Sets

Once GPS data are collected, an algorithm is established to identify possible trip ends. The primary variable used to identify possible trip ends for vehicular travel is the time gap, also referred to as the stopped time or dwell time (Aultman-Hall, 2007). The stopped time occurs in the GPS data when a vehicle is stationary for a period of time. The challenge with this variable is to select a time period that will have the most accurate identifications of trip ends. If the period selected is too short, stops at traffic signals or other traffic conditions could incorrectly be identified as trip ends. If the period selected is too long, short stops at destinations could be missed (Aultman-Hall, 2007). Stop times that have been used in a range of studies are shown in Table 2-10.

Table 2-10: Stopped Times Recommended by Studies or Documents

No	Stopped Time or Time Gap or Dwell Time (seconds)	Title of Study / Document	Author	Year
1.	120	Exploring the potential of automatically collected GPS data for travel behaviour analysis	Schonfelder et al.	2002
2.	45	GPS and travel surveys: Results from 1997 Austin household survey	Pearson	2001
3.	180	Trip reporting in household travel diaries: A comparison to GPS-collected data	Casas & Arce	1999
4.	120	Using GPS Data Loggers to Replace Travel Diaries in the Collection of Travel Data	Wolf	2000
5.	120	GPS, Location, and Household Travel	Stopher	2004
6.	More than 300 - confident Between 300 and 120 - suspicious delays Between 120 and 20 - probable	Eighty Weeks of GPS Traces, Approaches to Enriching Trip Information	Axhausen et al.	2004
7.	120	Post-processing procedures for passive GPS based travel survey	Liu et al.	2013
8.	300	Trip destination prediction based on past GPS log using a Hidden Markov Model	Alverez - Garcia	2010
9.	120	Applying GPS Data to Understand Travel Behaviour	NCHRP Volume 1	2014
10.	60	Should we change the rules for trip identification for GPS travel records	Shen et al.	2013
11	12	Automatic segmentation and classification of movement trajectories for transportation modes	Biljecki	2010

According to Bhat most studies recommend 120 seconds as the stopped time value (Bhat, 2009). The literature reviewed confirmed this statement, as shown in Table 2-10.

There are some patterns in the GPS data that could assist in identifying possible trip ends. These patterns are:

- Heading changes of approximately 180˚;
- Parking patterns;
- Repeated use of road links;
- Distance from the road network, and
- Circuitous routes.

Heading changes of approximately 180˚ can occur in the proximity of the trip end as shown in Figure 2-8 schematically. The point where the heading change occurs is not necessarily the exact trip end location. The trip end can occur before or after the heading change as well (Aultman-Hall, 2007).

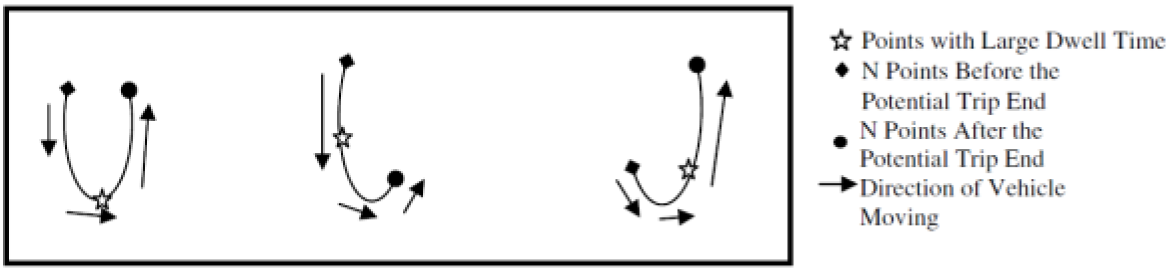


Figure 2-8: Three Schematic Scenarios of Heading Change Around Trip Ends (Aultman-Hall, 2007).

A parallel parking track is shown in Figure 2-9. The movements performed by a vehicle during a parking manoeuver are usually distinctly different from normal driving. There are a lot of direction changes at low speed to enter and exit a parking space. The trip end should be located on the parking bay (Aultman-Hall, 2007).

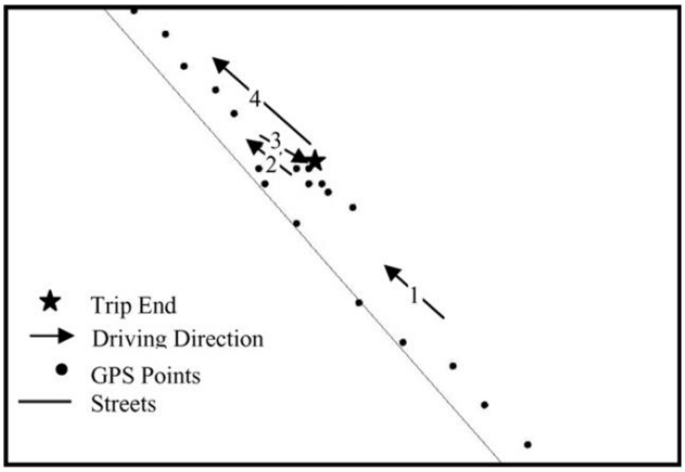


Figure 2-9: Parallel Parking Tracks (Aultman-Hall, 2007).

Figure 2-10 shows a GPS track that travels to a point, turns around and travels back on the same route. This track is usually present near a trip end. The trip end is usually located near the turnaround point. The driver reaches his or her destination and then returns the same way (Aultman-Hall, 2007).

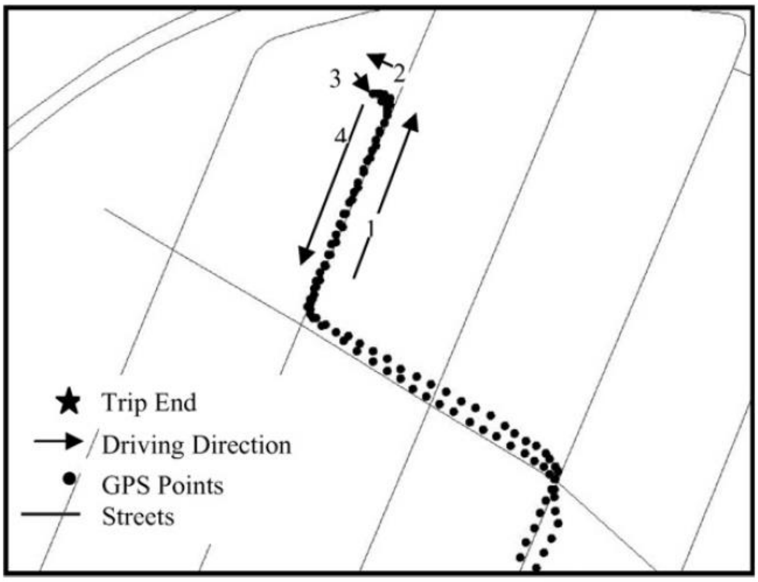


Figure 2-10: Repeated Use of Road Links (Aultman-Hall, 2007).

Distance away from the road network as can also indicate a possible trip end, shown in Figure 2-11. Accurate GIS road network data and GPS data are required for this method. This variable could identify many false trip ends, if the GIS road network and the GPS logs don't match well (Aultman-Hall, 2007).

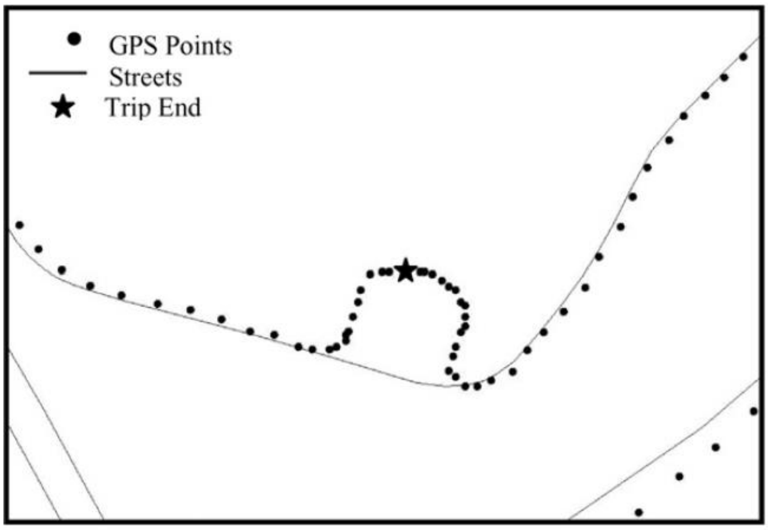


Figure 2-11: Points off the Road Network (Aultman-Hall, 2007).

Circuitous routes are an indication that a trip end was probably missed, as shown in Figure 2-12. If there is clearly a shorter and faster route between origin and destination, it could be that two trips are shown as one (Aultman-Hall, 2007).

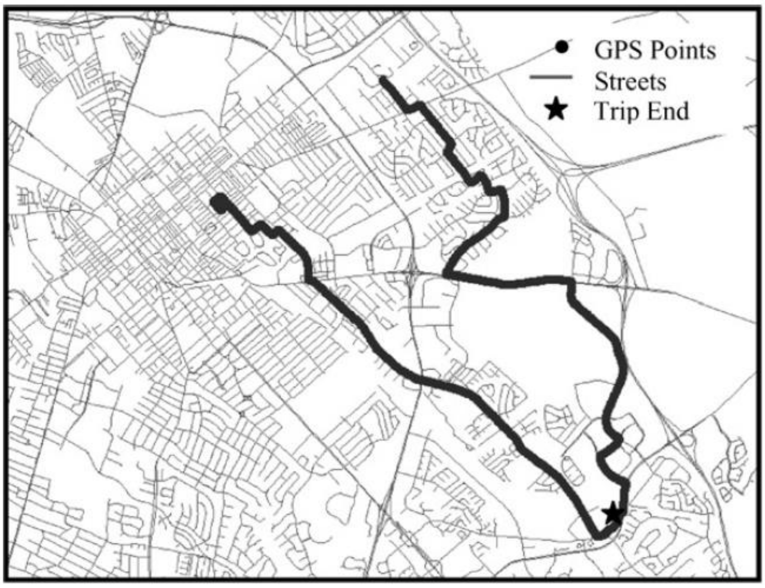


Figure 2-12: Circuitous Route (Aultman-Hall, 2007).

Due to the complexity of the travel behaviour there is no stop time interval that will identify all trip ends accurately. Aultman recommends that the following criteria be used to estimate trip ends:

- Maximum and minimum stopped time;
- Heading changes, and
- Distance from the road network.

2.6.7 Trip Purpose Based on Trip Ends

The trip purpose can be estimated by identifying the land uses around the trip end. The land uses that are used to identify the purpose of the trip are limited to a 200 m to 300 m radius around the trip end (Clifford, 2008). Table 2-11 shows possible trip purposes for certain land uses.

Table 2-11: Land Uses Associated with Possible Trip Purpose

Land use	Trip Purpose
Residential	Return home
Offices	Go to work
Shopping centre	Shopping and social
Educational	Go to school

To assist with the identification of trip purpose, respondents to the GPS survey can be requested to provide some addresses. The addresses requested should include: workplace, home, schools and shopping centres which are often visited by them. If an end point is located near these addresses, then the trip purpose is easily identifiable (Clifford, 2008).

Trips made to mixed land use developments or areas present a challenge to identifying the trip purpose (Chen et al., 2010). Chen used a probabilistic model to evaluate various factors that could provide some indication of what the trip purpose could be. The factors considered were: time of day, history dependence and land use characteristics. With these three factors two multinomial logit models were developed (one for home based trips and the other one for non-home based trips) to calculate the probability of a trip having a specific purpose (Chen et al., 2010).

Axhausen had a different approach which based probability of the purpose on the distance from the end point to different destinations. The land use located within 50 m of the trip end was given a weight of 1.5, between 50 m and 100 m a weight of 1, between 100 m and 200 m a weight of 0.7 and between 200 m and 300 m a weight of 0.4 (Axhausen et al., 2003).

2.6.8 Parking Types Influencing the Estimation of Destinations

In the case of vehicular trips, the trip end is also where the driver of the vehicle decides to park. Drivers usually park as close as possible to their destinations (Reimers, 2013). However, the type of parking available will have an influence on where drivers decide to park. The different types of parking are: on street parking; garage parking and parking lots. Parking lots are constructed to serve a particular development or shopping centre and are theoretically only used by people coming to visit this particular development. For shopping centre developments, it is recommended that parking should not be provided further than 120 m from the shopping centre entrances (Cloete, 2002).

On street parking and garage parking serves multiple developments. It is thus difficult to estimate what development was the destination of the drivers. Street parking is preferred over garage parking because on average the driver can park closer to his or her destination. However, the availability of open on street parking and the tariffs of street parking compared to garage parking could influence this preference (Kobus et al., 2013).

2.6.9 Calculating Trip Lengths from GPS Data

After identifying trip ends, there are two main methods for determining the trip lengths from the GPS data. These methods are:

- The Point-to-Point method - sum of distances between points over the entire trip (PP), and
- The Link-to-Link method - sum link lengths over the entire trip after matching the GPS points to road network links (LL).

Each of these methods has advantages and disadvantages. The LL method requires accurate road network data with GPS data linked to the network. The PP method could calculate the trip length inaccurately, due to the GPS that might have lost signal during the specific trip (Bhat, 2009). Thus, the LL method could be more accurate if an accurate GIS database is available with the required road links data, while the PP method is much faster and easier to calculate. Caution should be taken for any errors or missing GPS points when using the PP method.

2.6.10 The V Model for Designing GPS Data Collection

The V model for designing a GPS survey is shown in Figure 2-13. This model can be used to design a GPS survey. Each of these blocks is discussed briefly below.

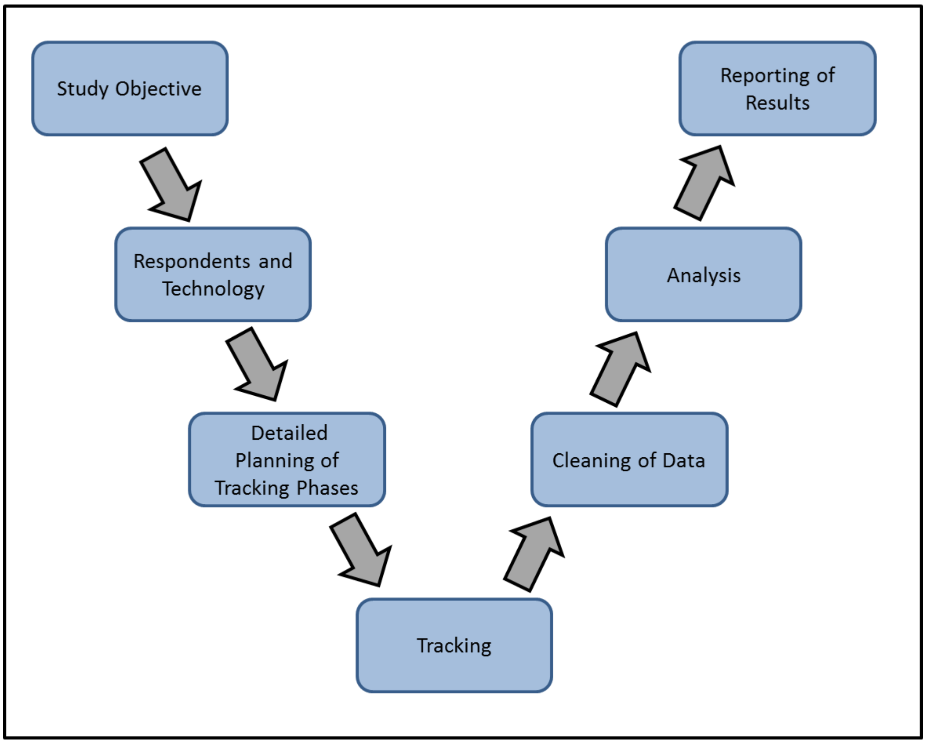


Figure 2-13: The V Model for GPS Tracking (Rasouli, 2014)

2.6.10.1 Study Objective

The study objective should be evaluated against the following three questions (Rasouli, 2014):

- Does the data already exist?
- If not, is GPS tracking actually the right technology to choose?
- Can the study objective be reached in a theoretical grounded way?

If the answers on these three questions are, no the data does not exist, GPS technology is the most appropriate technology and the objectives can be reached in a theoretical grounded way with the GPS data, then the GPS survey may be conducted (Rasouli, 2014).

2.6.10.2 Respondents and Technology

This block focuses on the interplay between the technology and the respondents. With all the different GPS data gathering technologies available the most effective technology should be chosen considering the respondents, cost and study objective (Rasouli, 2014). For example if the decision is taken to use smartphones as the GPS gathering technology, only people with smartphones can participate in the survey. If this bias does not impact on the objective of the study the smartphone technology can be used. However, if the bias impacts the study objectives negatively, other technology should be used or the respondents without smartphones should be provided with another technology. It is recommended that a pilot study be conducted to test the technology with a small group of respondents and to identify any problems with the data gathering methodology (Rasouli, 2014).

2.6.10.3 Detailed Planning of Tracking Phases

This phases involves the following steps (Rasouli, 2014):

1. Recruitment of respondents;
2. Information provision to respondents about details of the project and their tasks;
3. Collecting legal contracts regarding participation from respondents;
4. Distribution of GPS device or smartphone apps;
5. Tracking;
6. Collecting of GPS devices, and
7. Feedback to and from respondents.

2.6.10.4 Tracking

During this phase the devices are tracking (gathering GPS data) the movements of respondents. It is recommended to set up a hotline for respondents to use if they have questions or something goes wrong the equipment. If the tracking technology allows real time monitoring, devices should be monitored to identify any devices malfunctioning and participants that are not participating as requested (Rasouli, 2014).

2.6.10.5 Cleaning of Data, Analysis and Reporting

The cleaning of the GPS data are a time and resource consuming task. This task should not be underestimated. Proper data cleaning is required before the data can be used for analyses. The cleaned GPS data are analysed to reach the study objectives. The result of the study can then be reported on (Rasouli, 2014).

2.7 Summary of Literature Review

The bulk contributions calculation for roads contains a variable for the halved adjusted trip length. This variable is the average distance travelled to or from a development which is then halved and adjusted to only reflect the distance travelled on roads under the jurisdiction of the municipality, excluding distance travelled on Class 4 and Class 5 roads. It was found that limited data are available in South Africa on average trip length per land use type including shopping centres. Currently shopping centre trip lengths are estimated using the TMH17 document for calculating bulk contributions. Trip length data can be gathered during traditional household travel surveys and trip dairies. However, research has shown that the accuracy of the lengths recalled by participants is usually questionable.

GPS technology can assist in gathering more accurate trip data. GPS technologies consist of passive and active devices. The passive GPS is the most basic device which only records position at a time interval for the duration of the survey. The active GPS requires the participant to enter information on each trip, in addition to recording the position at a set time interval. Algorithms are used to identify trip ends within the GPS data. The main criteria used to determine trip ends is the stop time; most studies use a stop time of 120 seconds. After the trip ends are identified the following can be determined or estimated: trip purpose, trip length, origin destination pairs, average speed and travel time.

3. Methodology

3.1 Introduction

This chapter discusses the methodology followed to estimate the average trip lengths for shopping centre trips based on GPS data. The methodology can be summarised in the following steps:

- Obtaining the GPS data;
- Cleaning the GPS data;
- Determining home coordinates of participants;
- Determining trip ends;
- Determining which trip ends are at shopping centres, and
- Calculating average trip lengths.

3.2 Obtaining the GPS Data

3.2.1 Background of the GPS data

The GPS data used for this study were obtained from a previous study undertaken at the University of Pretoria. The purpose of the original study was to evaluate the impact that the E-toll system (open-road tolling on freeways in Gauteng Province) had on driving patterns in the province. The study was designed as a panel study; data were initially gathered before the implementation of the E-toll system to obtain baseline driving patterns, and the exercise will be repeated in the post e-toll implementation phase. The baseline data were analysed with a view to establishing freeway use patterns for various user groups, including calculating vehicle kilometres of travel and speed of travel per road type. With this information a method was developed to estimate the fuel consumption of private vehicles from GPS data (see Venter and Joubert, 2013). A second analysis was undertaken to determine the freeway vehicle kilometres of travel for private vehicle users of different income groups. This was then used to determine the cost and equity implications to the private vehicle freeway user of electronic freeway tolling compared to an additional fuel tax (Venter and Joubert, 2014).

The GPS dataset was considered a suitable data source for the present study, for the following reasons:

- The size of the sample;
- The data was already collected and available;
- No funding was available for additional data collection;
- Some initial data cleaning was already conducted, and

- The GPS data was coupled with a user characteristics survey which included home suburb, income and gender of participants. Thus this data presented the opportunity to compare trip length data across socio-economic characteristics.

The detailed procedure followed during collection of this set of GPS data are described below.

3.2.2 Participant Selection

The participants were recruited through face-to-face home visits after recruiting through telephonic methods were not successful. The participants were chosen based on a stratified random approach. The size of the sample in every area was pre-selected to be proportional to the percentage of freeway users residing in the area. This was also done to ensure large enough sub-samples across income groups were surveyed. The intention was to have a representative sample of car owners, spatially and demographically, in the Gauteng Province (Venter, 2013). The sample is described in Section 4.2.

3.2.3 GPS Survey Process

The GPS survey process is summarised as follows (Venter, 2013):

- Field workers delivered GPS logger unit to participants;
- Participants were instructed to place the logger unit within their vehicles and to continue with their normal travel patterns within the Gauteng Province until the logger was collected;
- Weekends were included in the study but public holidays and out of province trips were excluded;
- Field workers revisited each participant after three to five days to collect the GPS loggers and to complete a short questionnaire on household, demographical information, work location and vehicle details;
- Participants were offered a R 200 gift voucher for participating;
- Participants signed an agreement in terms of which the data would be treated confidentially and anonymously, and
- The GPS data was downloaded and processed.

3.2.4 The GPS Device

The GPS device used for this survey was a Tracking Key 3100-INT model manufactured by LandAirSea Systems, Inc. A picture of the device is shown in Figure 3-1. This GPS is a passive device which records the position at every second that the vehicle is moving. It can record up to 300 hours of data and records the position with a horizontal accuracy of 2.5 m. The battery life is between

5 and 8 days (Venter, 2013). The device has no buttons and the fact that it requires no interaction with the user makes it extremely robust and simple to use.

Figure 3-1: Photo of the 3100-INT GPS Device

3.2.5 The GPS Data

The GPS devices were given to 726 participants to place in their vehicle's glove compartments. The GPS devices were configured to record the vehicle's position every second until the devices were retrieved. Each participant had the GPS in their vehicles on average for approximately four days. The travel patterns were recorded between November 2011 and March 2012 (Venter, 2013).

Once the GPS devices were retrieved from each participant the data were downloaded. A total of 608 (approximately 84%) participants' data were successfully retrieved from the 726 participants. The remaining 118 (approximately 16%) participants data were not used for this study for a variety of reasons, including device malfunctions and batteries that died during the survey period and travel outside of the study area.

The GPS device provides its data in a "las" file format. These files were converted to Microsoft Excel csv (Comma delimited). An extract of the GPS data in the csv file is shown in Table 3-1. The GPS id number is given in the first column. Columns 2 to 6 are used for the date and time. Columns 7 to 10 are the Latitudes, Longitudes, X coordinates and Y coordinates. Columns 11 and 12 give the street name and type of road based on a high level GIS road network. Column 14 shows the time elapsed between records and Column 15 provides the speed of travel between records.

Table 3-1: Example of GPS data csv file

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
id	month	day	hour	minute	second	latitude	longitude	x	y	name	type	distance (m)	time (s)	speed (km/h)
81	0	0	14	31	57	-2540.51	2809.898	28.16497	-25.6751	DAAN DE WET NEL STREET	MAIN ROADS	0	0	0
81	0	0	14	31	58	-2540.51	2809.898	28.16496	-25.6752	DAAN DE WET NEL STREET	MAIN ROADS	10.329354	1	37.185674
81	0	0	14	31	59	-2540.52	2809.898	28.16496	-25.6753	DAAN DE WET NEL STREET	MAIN ROADS	8.813749	1	31.729495
81	0	0	14	32	0	-2540.52	2809.897	28.16494	-25.6753	DAAN DE WET NEL STREET	MAIN ROADS	7.907853	1	28.468272
81	0	0	14	32	1	-2540.52	2809.897	28.16494	-25.6754	DAAN DE WET NEL STREET	MAIN ROADS	5.501722	1	19.8062
81	0	0	14	32	2	-2540.53	2809.897	28.16494	-25.6754	LINEATA STREET	STREET	5.501722	1	19.8062
81	0	0	14	32	3	-2540.53	2809.896	28.16494	-25.6755	LINEATA STREET	STREET	3.330509	1	11.989833
81	0	0	14	32	4	-2540.53	2809.896	28.16494	-25.6755	LINEATA STREET	STREET	2.43424	1	8.763263
81	0	0	14	32	5	-2540.53	2809.896	28.16494	-25.6755	LINEATA STREET	STREET	0.906634	1	3.263883
81	0	0	14	32	36	-2540.52	2809.919	28.16532	-25.6753	DAAN DE WET NEL	MAIN ROADS	43.766605	31	5.082574
81	0	0	14	32	37	-2540.53	2809.915	28.16526	-25.6755	DAAN DE WET NEL	MAIN ROADS	16.600118	1	59.760427
81	0	0	14	32	38	-2540.53	2809.911	28.16519	-25.6756	DAAN DE WET NEL STREET	MAIN ROADS	14.904189	1	53.655082
81	0	0	14	32	39	-2540.54	2809.906	28.1651	-25.6757	DAAN DE WET NEL STREET	MAIN ROADS	13.982674	1	50.337626
81	0	0	14	32	40	-2540.55	2809.899	28.16499	-25.6758	DAAN DE WET NEL STREET	MAIN ROADS	14.286545	1	51.431563
81	0	0	14	32	41	-2540.55	2809.892	28.16487	-25.6758	LINEATA STREET	STREET	13.87014	1	49.932504
81	0	0	14	32	42	-2540.55	2809.885	28.16475	-25.6759	BOSCHBERG	STREET	14.699758	1	52.919128
81	0	0	14	32	43	-2540.56	2809.878	28.16463	-25.676	BOSCHBERG	STREET	14.195844	1	51.105039
81	0	0	14	32	44	-2540.56	2809.87	28.1645	-25.676	BOSCHBERG	STREET	15.517915	1	55.864493
81	0	0	14	32	45	-2540.57	2809.861	28.16436	-25.6761	BOSCHBERG	STREET	16.773609	1	60.384993
81	0	0	14	32	46	-2540.57	2809.853	28.16422	-25.6762	BOSCHBERG	STREET	15.883241	1	57.179668
81	0	0	14	32	47	-2540.58	2809.844	28.16406	-25.6763	BOSCHBERG	STREET	18.06831	1	65.045915
81	0	0	14	32	48	-2540.58	2809.835	28.16391	-25.6764	BOSCHBERG	STREET	17.867726	1	64.323814

3.3 Cleaning the GPS Data

Each csv file had to be checked for the following:

- Whether the GPS was activated before it was given to the participant;
- Whether the GPS was deactivated immediately after retrieving it from the participant, and
- Whether the GPS operated correctly during the survey period?

A manual process was worked through where all the GPS data not relevant to the participants were removed from data sets. Some GPS logs where removed completely due to the GPS malfunctioning and only 502 (approximately 70%) participants' data could be used in this study.

Furthermore, GPS data usually contains anomalies. These anomalies occur when the GPS device loses signal or has limited signal. The criteria used to identify these anomalies were as follows:

- Speed between two points more than 140 km/h;
- The X coordinate not between 27.000000 and 29.000000, and
- The Y coordinate not between -25.000000 and -27.000000.

The speed of 140 km/h was selected based on the highest speed limit being 120 km/h in South Africa. The 140 km/h allows for some drivers who drive above the legal speed limit and removes the impossible speeds that can be recorded by GPS devices. The X and Y coordinate limits results in a block around Gauteng as shown in Figure 3-2.

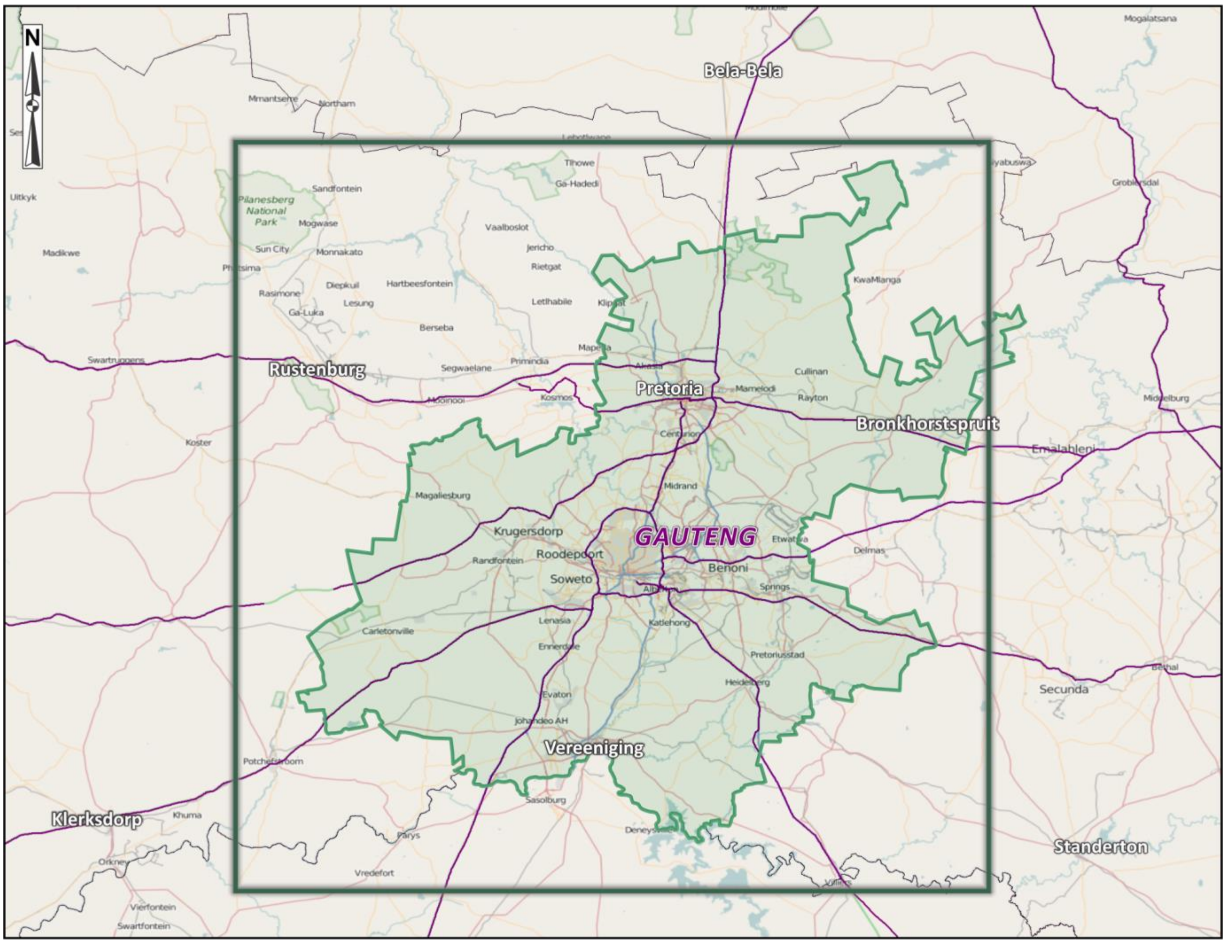


Figure 3-2: Block around Gauteng

If the X and Y coordinates were not between the specified values, those points were removed from the data set. If the speed between two points was more than the specified speed, a default speed of 40 km/h was assumed and used further in the cumulative distance calculations of the entire track log. In some instances Microsoft Excel could not perform the X and Y coordinate criteria, as the coordinate data from the GPS was not recognised as numbers by Microsoft Excel. This resulted in #Value and #Num messages when data was analysed. The data logs giving these error messages were identified and the anomalies in the X and Y coordinates were removed manually.

3.4 Determining the Home Coordinates of Participants

Before the Home to Shop and Shop to Home trips could be identified, the participant's home coordinates had to be estimated. This was done through a manual process of determining where the vehicle spent most of the evening and early morning hours (night hours). The locations were identified by viewing the GPS track in Google Earth satellite imagery. The Home coordinates were then visually estimate based on the GPS track. The identified Home locations were tested with the provided suburb of residence in the questionnaire. A list of home coordinates was then compiled.

With this manual and visual method the GPS tracks was interpreted to identify the correct home coordinates. This method has a high probability of providing correct Home coordinates. The down side to this method is that it is extremely time consuming.

3.5 Determination of Trip Ends

The criteria used to identify the probable trip ends in the GPS data sets are Stopped time and Repeated Use of Road Links. Both of these criteria are discussed in more detail below.

3.5.1 Stopped Time

When a vehicle is moving, the GPS device records a position every second. However, when the vehicle is stationary the GPS records the location and the number of seconds the vehicle is stationary at this position. The stop time criterion makes use of the number of seconds for which the vehicle is stationary to determine whether the stop is a trip end or not. The challenge with this variable is to select a time period that will have the most accurate identification of trip ends. The literature shows that the stop time variable used in studies can be between 45 to 300 seconds (refer to Section 2.6.6), with 120 seconds being the most popular.

For this study the appropriate minimum stopped time was analysed by randomly selecting 50 trip logs (approximately 10% of total trip logs) and identifying the trip ends based only on different stop time values. The stop time value was varied between 45 and 600 seconds. The trip ends were evaluated based on location as shown in Table 3-2. The approach essentially acknowledged that the likelihood of a stopped signal of a certain length corresponding to an actual trip end depends on the location of the stop. For instance, a stopped time of 60 seconds is less likely to indicate a true trip end if it occurred at a congested intersection than if it occurred outside the road reserve. In this manner identified trip ends are classified as incorrect, uncertain, or correct.

Table 3-2: Criteria Used to Evaluate Trip Ends

Stop Time (seconds/Minutes)	Location of Trip End					
	On Freeways, On-Ramp or Off-Ramp (Class 1)	**On Major Roads (Class 2)**	**On Road Near Intersection (Class 3)**	**On Road Far From Intersection (Class 3)**	**On Minor Road (Class 4 & 5)**	**Within Property**
45 / 0:45	Incorrect	Incorrect	Incorrect	Incorrect	Uncertain	Correct
60 / 1:00	Incorrect	Incorrect	Incorrect	Incorrect	Uncertain	Correct
80 / 1:20	Incorrect	Incorrect	Incorrect	Incorrect	Uncertain	Correct
100 / 1:40	Incorrect	Incorrect	Incorrect	Uncertain	Correct	Correct
120 / 2:00	Incorrect	Incorrect	Incorrect	Uncertain	Correct	Correct
150 / 2:30	Incorrect	Incorrect	Incorrect	Uncertain	Correct	Correct
180 / 3:00	Incorrect	Incorrect	Uncertain	Uncertain	Correct	Correct
240 / 4:00	Incorrect	Incorrect	Uncertain	Uncertain	Correct	Correct
300 / 5:00	Incorrect	Uncertain	Correct	Correct	Correct	Correct
600 / 10:00	Incorrect	Uncertain	Correct	Correct	Correct	Correct

The reasoning behind the criteria is as follows:

- On Freeways, On-Ramp or Off-Ramp
  - Under normal driving conditions trips don't end on freeways or on ramps.
  - Thus all trips ends identified on freeways were marked as incorrect and are believed to be caused by traffic congestion.
- On Major Roads
  - It is considered unlikely that a trip end would be located on a major road. Major roads usually have some sort of access management and no on street parking is provided. However, in some cases (e.g. people parking within road reserve as not sufficient parking is provided within a property) it is possible, thus for longer stopped times they will be marked as uncertain.
  - To determine this longer stopped time it is observed that due to saturated traffic conditions, traffic may be stopped at signalised intersections for up to two cycles. At a typical length of 150 seconds, this would equate to a stopped time of 300 seconds. Thus 300 seconds was selected as the cut-off point between incorrect and uncertain.
- On Road Near Intersection
  - A trip end less than 100m from an intersection on a Class 3 road was categorised as a near intersection trip end.
  - Both signal, stop and priority control are used on these roads. Traffic congestion occurs mainly at intersections which increase the probability that short stop times

could identify traffic congestion as trip ends. Thus stop times up to one cycle length (150 seconds) was identified as incorrect trip ends.

- Cycle failure less likely on Class 3 roads. Thus long stops (300 seconds plus) not likely to be caused by oversaturated signals, so assumed to be correctly identified trip ends.
- Time in between is considered as uncertain.

- On Road Far From Intersection
  - Very short stops not likely to be trip end as there is not enough time to leave vehicle, do activity and return to the vehicle. Given likely walking distance from Class 3 road to property, it was estimated that at least 100 seconds would be required to complete even a short time activity. Thus trip ends shorter than 100 seconds are labelled incorrect.
  - Between 100 and 300 seconds stops are less likely but not impossible. On street parking is not common but occurs sometimes. Thus these are labelled as uncertain.
  - Long stops (300seconds plus) not likely to be caused by oversaturated traffic conditions, so assumed to be correctly identified trip ends.
- On Minor Road
  - Minimum stop time of 100 seconds still applies. However, due to shorter walking distances between on street parking space and property, there is a larger possibility that shorter stops could be actual stops. Thus shorter stops were labelled uncertain.
  - Low traffic congestion levels and low control delay at junctions, thus medium and longer stop times more likely to indicate true trip ends (as compared to higher order roads). Thus anything longer than 100 seconds was labelled as correct.
- Within Property
  - All trip ends within a property was marked as correct.

The results of the analyses are shown in Table 3-3 and Figure 3-3.

Table 3-3: Stop Time Experiment Results

Stop Time (Seconds)	Stop Time (Minutes)	Total Number of Trip Ends Identified	Correct Trip Ends Identified	Uncertain Trip Ends Identified	Incorrect Trip Ends Identified	Correct Trip Ends Missed	Sum of Uncertain, Incorrect and Correct Trip Ends Missed	% of Total Uncertain, Incorrect and Correct Trip Ends Missed
45	0:45	1050	760	104	186	0	290	28%
60	1:00	884	735	83	66	25	174	20%
80	1:20	804	699	72	33	61	166	21%
100	1:40	777	689	64	24	71	159	20%
120	2:00	740	667	54	19	93	166	22%
150	2:30	710	641	51	18	119	188	26%
180	3:00	680	615	48	17	145	210	31%
240	4:00	637	573	47	17	187	251	39%
300	5:00	608	545	46	17	215	278	46%
600	10:00	507	453	39	15	307	361	71%

It was assumed that with the 45 seconds stop time, no correct stop ends was missed. With this assumption the correct trip ends missed by greater stop times were calculated. The uncertain and incorrect trip ends were also added together for the purpose of the analysis.

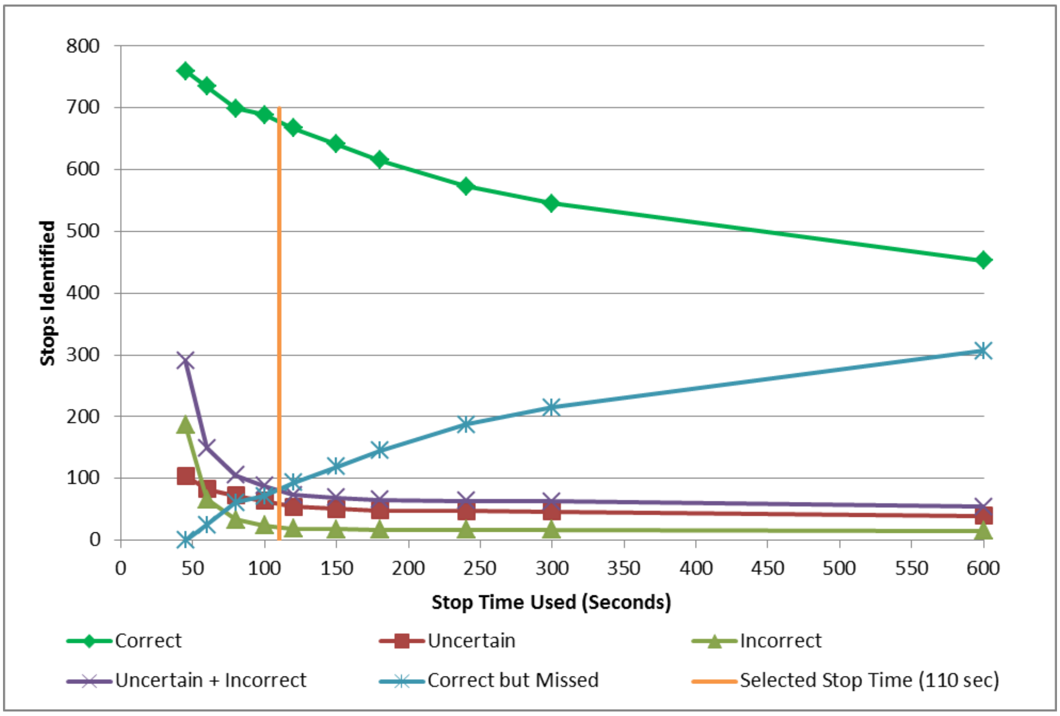


Figure 3-3: Stop Time Analyses

The optimal stop time will be the stop time value which minimises both the correct trip ends missed (blue line) and the uncertain + incorrect trip ends (purple line) as shown in Figure 3-3. The stop time analysis shows that 110 seconds should be used for the stop time value. The 110 seconds stop time value compares well with the literature (refer to Section 2.6.6) which shows most studies uses 120 second for their stop time value. The 110 seconds was subsequently used as the stop time value for this study.

3.5.2 Repeated Use of Road Links

Repeated use of road links occur when a vehicle travels along a road and then turns around, returning along the same road, as shown in Figure 2-10. The method used to detect these movements was as follows and is shown in Figure 3-4:

- Calculate the distance between the 50th preceding and 50th next point;
- Calculate the distance between the 40th preceding and 40th next point;
- Calculate the distance between the 30th preceding and 30th next point;
- Calculate the average of the three distances, and
- If the average distance is less than 20m, the point is marked as a trip end.

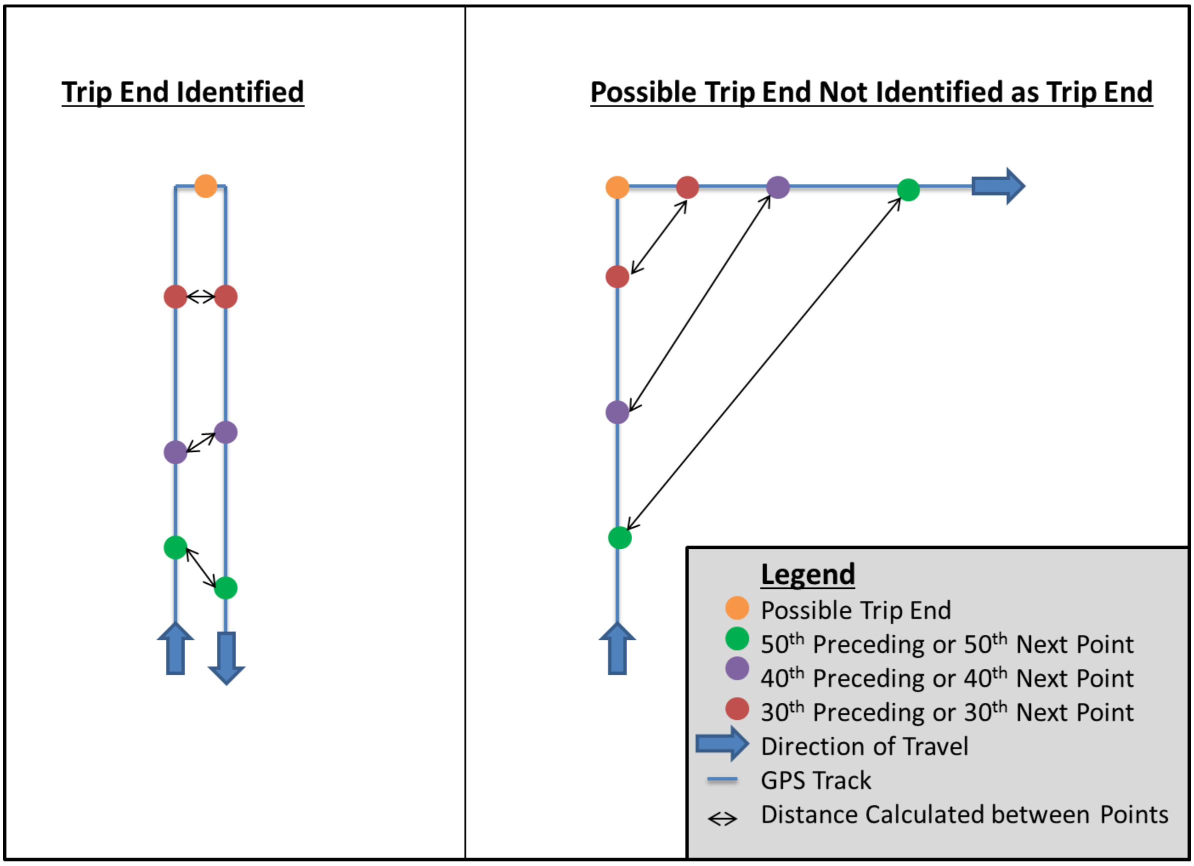


Figure 3-4: Illustration of Identifying Repeated Use of Road Link in GPS Data

The 20 m distance is based on two lanes per direction with a 6m wide median cross section as shown in Figure 3-5. The width of the two centre lanes plus the median equals 13 m. The accuracy level of the GPS points of 2.5 m was also taken in to account by adding 2.5 m for both points to the 13 m, which equals 18 m. The 18 m was then rounded to 20 m.

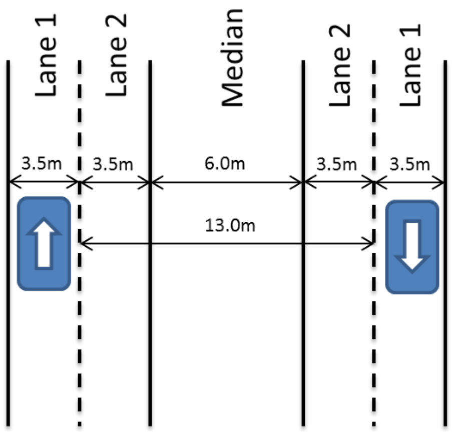


Figure 3-5: Cross Section

3.5.3 Cleaning Trip Ends

With these criteria the trip ends were identified. It was found that in some cases the same trip end was identified by both of the criteria, leading to consecutive trip ends located in close vicinity to each other representing one actual trip end. An additional measure was added to combine all of these trip ends into one trip end. The following measure was implemented: The straight line distance between sequential stop points was calculated and if a stop point was closer than 300 m, it was ignored and removed from the trip end results. The 300 m limit was selected as it is a comfortable walking distance and it was reasoned that a trip with a trip length of 300 m would not be conducted with a vehicle but rather by walking.

3.5.4 Heavy Traffic Conditions

Trip ends were also identified on freeways due to heavy traffic congestion. This was overcome by checking whether vehicles were stuck in heavy traffic conditions. The following measure was implemented:

- The average speed from the previous 30th point to the next 30th point was calculated;
- If the average speed at the specific point is less than 10 km/h, this point was deemed to be in heavy traffic conditions, and

- The trip ends identified within heavy traffic conditions were removed from the trip end results except if the stop time was more than 3600 seconds (1 hour).

3.6 Determining Which Trip Ends were at Shopping Centres

The list of trip ends was compared to a database for shopping centres in Gauteng Province. Using GIS all trip ends within a specified radius of the shopping centre point was identified as trip ends at shopping centres.

3.6.1 Shopping Centre Database

The shopping centre database contains the following information: the location (X Y coordinate), name of shopping centre and the GLA of the shopping centre. This database was compiled by sourcing information from the South African Shopping Centre Directory (SACSA, 2012) and measuring shopping centre buildings on Google Earth where required. The database contains 725 shopping centres, with an average centre size of 13 337m^2. These shopping centres were categorised according to the shopping centre types as described in Section 2.3.4 – refer to Table 3-4.

Table 3-4: Shopping Centres per Type

Shopping Centre Type	GLA (m^2)	Number of Shopping Centres	Percentage of Total Shopping Centres	Cumulative Percentage
Convenience Centres	500 to 5 000	235	32%	32%
Neighbourhood Centres	5 001 to 12 000	282	39%	71%
Community Centres	12 001 to 25 000	118	16%	88%
Small Regional Centres	25 001 to 50 000	56	8%	95%
Regional Centres	50 001 to 100 000	29	4%	99%
Super Regional Centres	More than 100 000	5	1%	100%
	Total	**725**	**100%**	

Source: Based on SACSA, 2012

The majority of shopping centres in Gauteng are Neighbourhood Centres. As shown in Figure 3-6 approximately only 10% of shopping centres in Gauteng have a GLA of more than 30 000m^2.

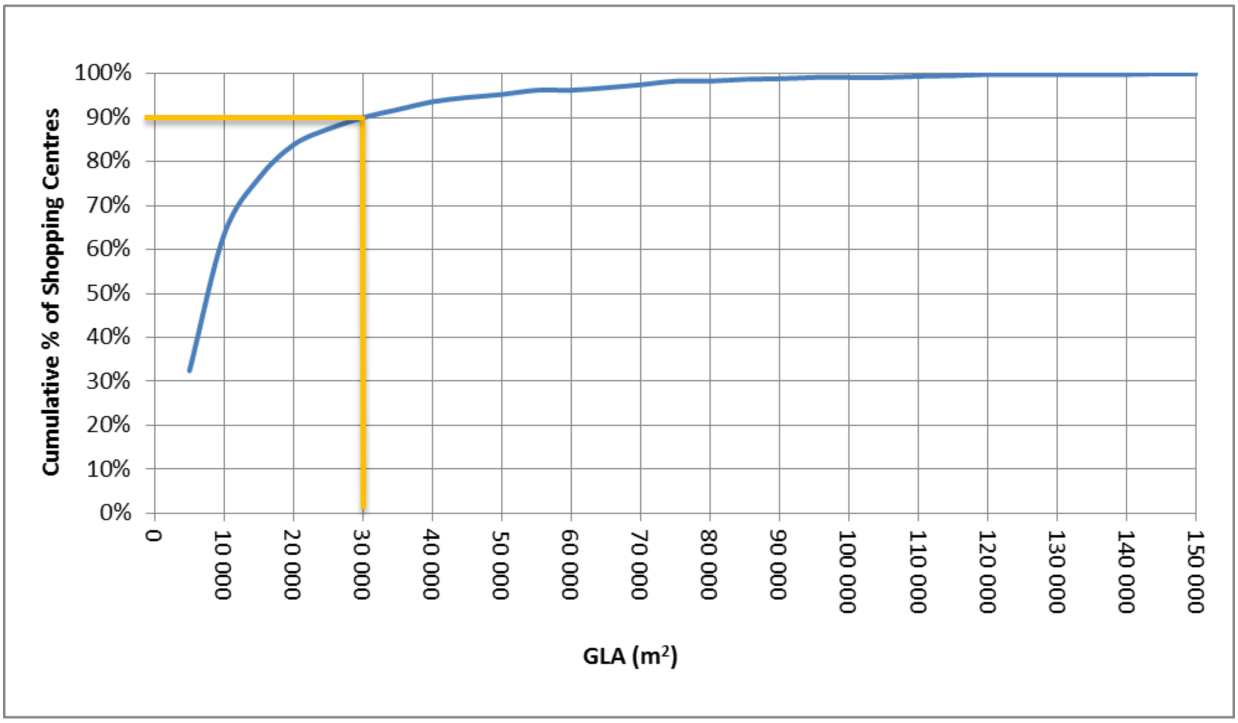


Figure 3-6: Cumulative Percentage of Shopping Centres Based on GLA (Source: Based on SACSA, 2012)

3.6.2 Radius Used per Shopping Centre Type

Each shopping centre in the shopping centre database has an XY coordinate to define its location. All trip ends within a specified radius of the shopping centre point are identified as trip ends at shopping centres. An analysis was done to determine the required radius length per shopping centre type. The radius length should be long enough to cover the parking area of the shopping centre but not too long to pick up other trip ends on properties adjacent to the shopping centres.

Ten shopping centres where randomly selected from each of the shopping centre types, except for the super regional centres for which all five was selected. Each selected shopping centre was viewed in Google Earth satellite imagery with its XY coordinate. The distance between the XY coordinate and the furthest boundaries/parking spaces of the shopping centres was measured. This length was then recorded as the required radius length. The average, 75th percentile and 25th percentile values were calculated per shopping centre type. The results are shown in Table 3-5 as well as the recommended radius length per shopping centre type.

Table 3-5: Radius Analysis

Shopping Centre Type	GLA (m^2)	25th Percentile Radius Required (m)	Average Radius Required (m)	75th Percentile Radius Required (m)	Recommended Radius (m)
Convenience Centres	500 to 5 000	56	67	75	70
Neighbourhood Centres	5 001 to 12 000	83	100	120	100
Community Centres	12 001 to 25 000	120	149	180	150
Small Regional Centres	25 001 to 50 000	180	212	238	220
Regional Centres	50 001 to 100 000	265	278	300	280
Super Regional Centres	More than 100 000	290	301	320	300

As expected the required radii increased with the size of the shopping centres. The recommended radii are based on the average radii. These recommended radii were used to identify the shopping centre trip ends.

3.7 Calculating Trip Lengths for Shopping Centres

For each trip log, a cumulative distance was calculated by adding the distances between the points recorded every second (PP method as discussed in Section 2.6.9). The trip ends at shopping centres are already identified at this point. The trip length to a shopping centre is calculated by subtracting the shopping centre trip end's cumulative distance from the previous trip end's cumulative distance value. For the trip length from a shopping centre, the same method is followed except the next trip end after the shopping centre trip end's cumulative distance is used. This procedure gives the trip length to and from the shopping centres.

3.8 Calculating Distance Travelled per Road Class

During the conversion process of the GPS files to Microsoft Excel csv, an attribute was added to each data point named road type. The data points were matched to a high level GIS road network to identify on which type of road the point was recorded. The GIS road network had six different road categories. These categories and the corresponding road Class for each are shown in Table 3-6. The length of travel on each road Class group was totalled for each trip.

Table 3-6: GIS Road Categories and Corresponding Road Classes

GIS Road Categories	Corresponding Road Classes	Road Class Group
Highway	Class 1	Class 1
Main Roads	Class 2 + 3	Class 2 +3
Secondary	Class 2 + 3	
Streets	Class 4 + 5	Class 4 + 5
Other	Class 4 + 5	
Unknown	Class 4 + 5	

4. Analysis and Results

4.1 Introduction

This chapter discusses the sample and examine the representativeness of the sample with respect to the driving population of the Gauteng province. After this the chapter focuses on the results obtained from analysing the GPS shopping centre trip length data.

4.2 Sample Description

The sample of 502 participants was investigated to determine whether the sample is representative of the driving population of the Gauteng province both spatially and demographically. According to the Census 2011 the population of the Gauteng province was estimated to be approximately 12.27 million people. The sample equals 0.004% of the estimated population of Gauteng.

4.2.1 Spatial Representation

The Gauteng Province is divided into three metropolitan and two district municipalities. These municipalities are:

- City of Tshwane (CoT) Metropolitan Municipality;
- City of Johannesburg (CoJ) Metropolitan Municipality;
- Ekurhuleni Metropolitan Municipality (EMM);
- West Rand District Municipality, and
- Sedibeng District Municipality.

These major municipalities are shown in Figure 4-1.

Figure 4-1: The Major Municipalities of Gauteng Province

The residential addresses of the 502 participants were plotted on a Gauteng Province map as shown in Figure 4-2. The red dots indicate the residential addresses of the participants. The dots are located mainly in three of the five municipalities namely: CoT, CoJ and EMM. The dots appear to be denser around the national freeways, in line with the sampling strategy for the original study which aimed at obtaining a representative sample of freeway users in Gauteng. This introduces a potential sampling bias, as freeway users may be a biased sub-set of all drivers. The question of bias is examined below in relation to spatial and socio-economic distributions.

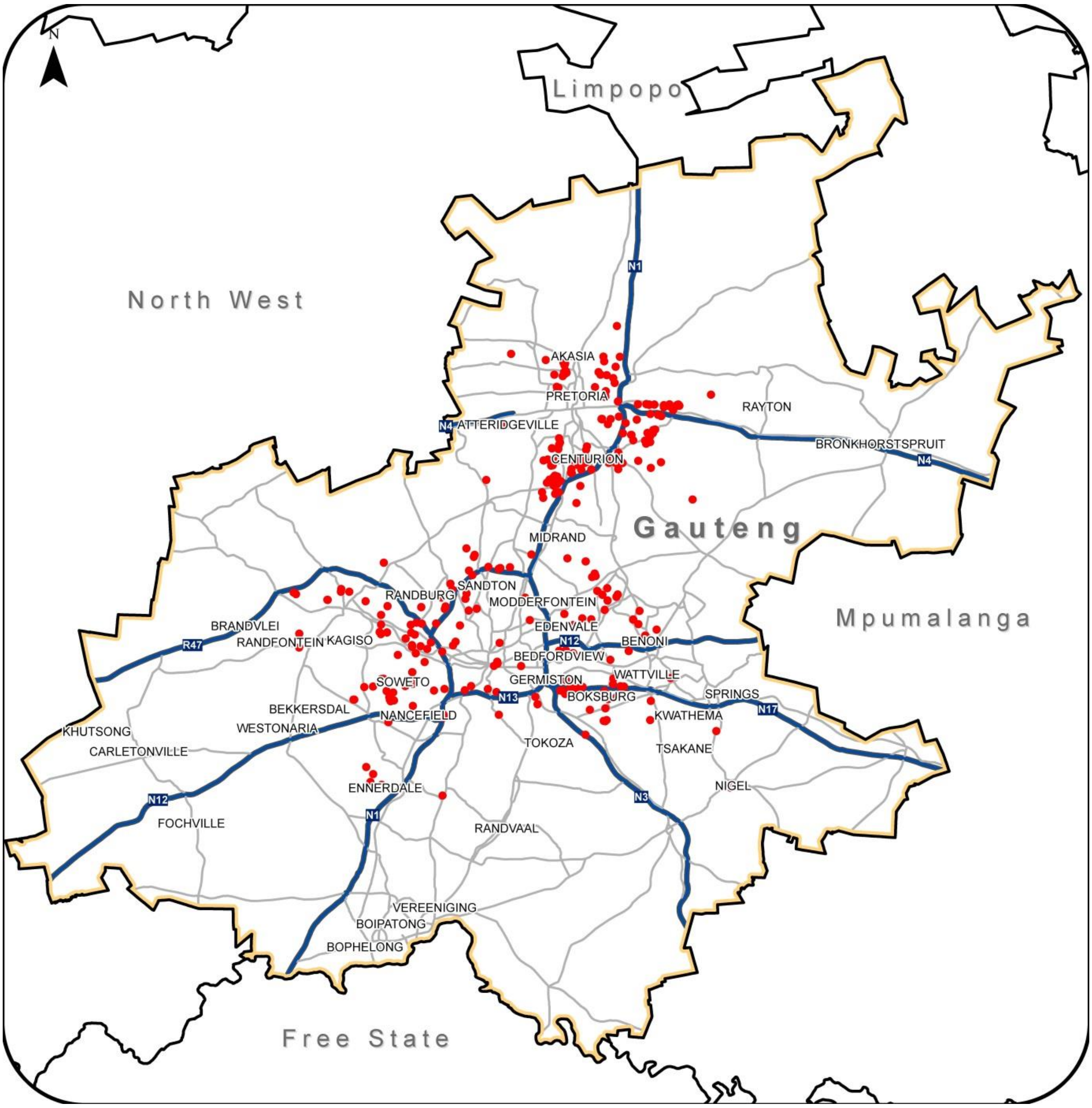


Figure 4-2: Location of Participants Residential Address

According to Statistics South Africa (Stats SA) CoT, CoJ and EMM have the highest populations among the five municipalities, containing 86% of the Gauteng population. The percentage residential addresses of the participants per municipality are compared to the Gauteng population in Figure 4-3. The sample compares well with the Census 2011 population distribution, thus spatially the sample is deemed to be representative of the population in Gauteng.

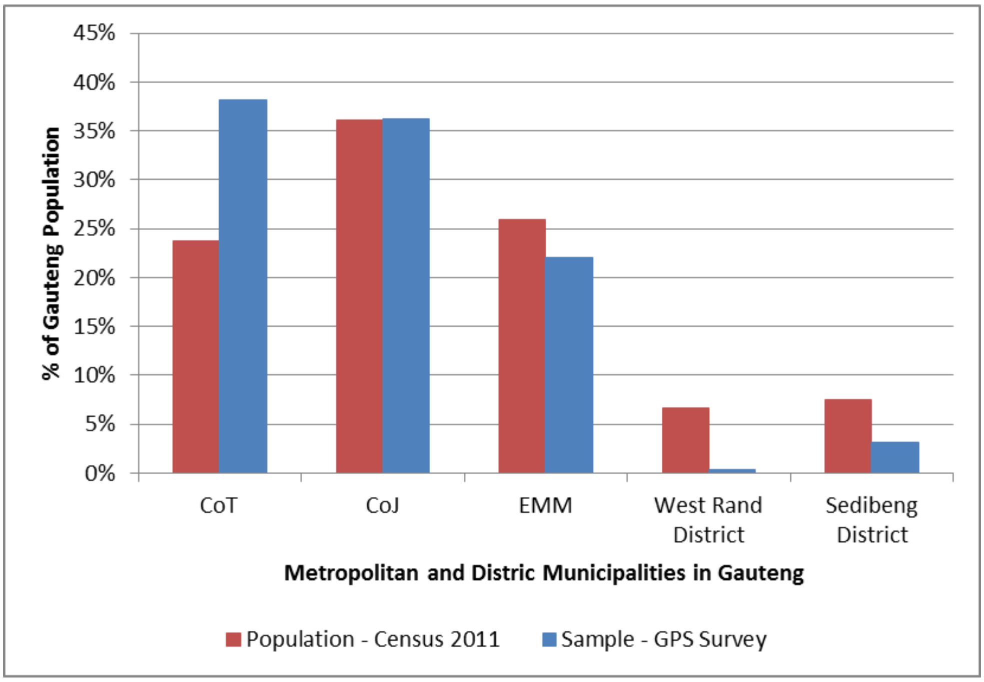


Figure 4-3: Gauteng and Sample Population Distribution

4.2.2 Demographic Representation

4.2.2.1 Age and Gender Distribution

Stats SA released statistics from the National Household Travel survey conducted in 2013. The number of people with a driving licence per age group and gender in South Africa is shown in Table 4-1. These statistics could unfortunately not be found for only the Gauteng Province. However, the distribution should be similar for the country as a whole.

Table 4-1: Persons aged 18 years and older by age group, type of driver's licence and Gender

Age Group	Motorcycle ('000)			Light motor vehicle ('000)			Heavy motor vehicle ('000)		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
18-25	42	31	11	621	363	259	294	218	76
26-39	98	66	32	1999	1135	864	1263	941	322
40-49	89	65	24	1383	804	579	668	553	115
50-59	97	77	21	1063	606	457	461	396	65
60+	71	51	20	1028	514	514	284	248	37
Total	**389**	**290**	**108**	**6095**	**3422**	**2673**	**2970**	**2355**	**616**

Source: Stats SA, National Household Travel Survey, 2014

The distributions are shown in Figure 4-4 and Figure 4-5 for males and females with driving licences compared to the sample's distribution. The general shapes of the distributions are similar. The age group of 26 to 39 for the male sample has a significantly higher percentage than the national distribution while the age groups of 50 to 59 and 60 plus are significantly lower. The female sample distribution also has a significantly lower percentage for the age group of 60 plus compared to the national distribution.

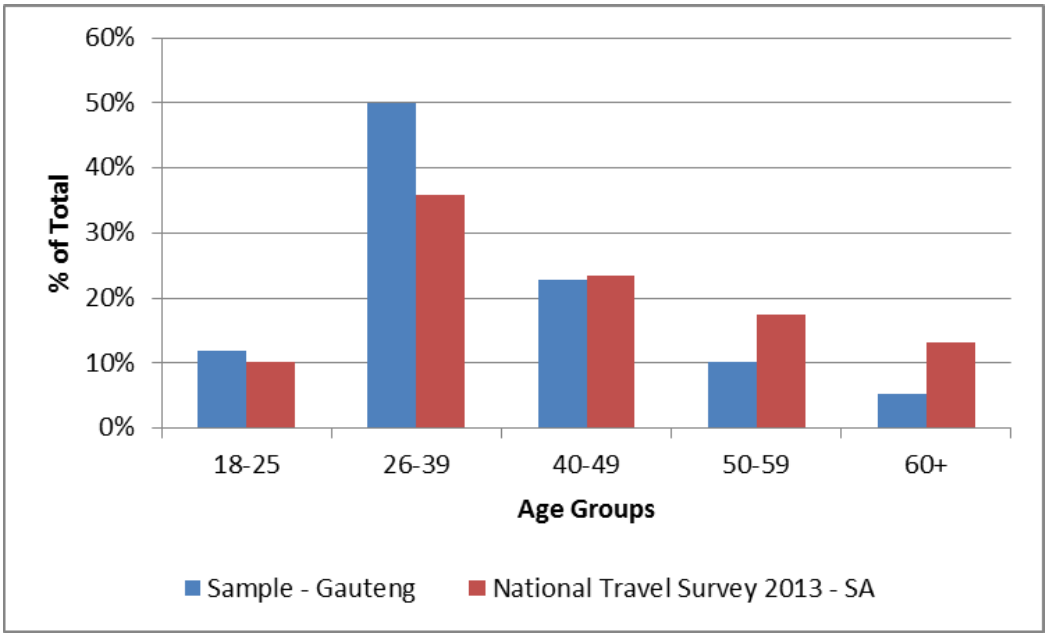


Figure 4-4: Male Driving Licence Age Distribution South Africa

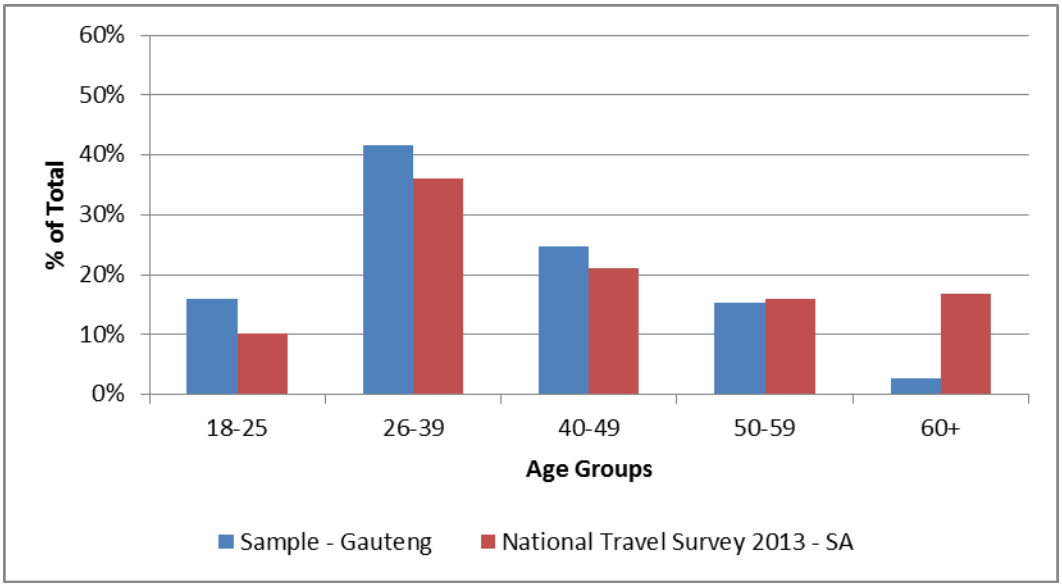


Figure 4-5: Female Driving Licence Age Distribution South Africa

Figure 4-6 shows the gender split for South Africa (purple) compared to Gauteng Province (green). The gender splits are approximately 50% for both. Figure 4-6 also shows the gender split between people with driving licenses in South Africa (red) and the Sample (blue). Since the general gender split for the country and Gauteng province is similar (purple and green), it was assumed that the people with driving licences in Gauteng province have a similar gender split as for the entire South Africa (red). Comparing the Sample (blue) gender split with country's driving licence (red) population gender split the gender splits are similar. Both the sample and the people with driving licenses in South Africa have a gender split of approximately 60% male to 40% female. It is clear from this that males are more prone to obtaining a driving licence in South Africa.

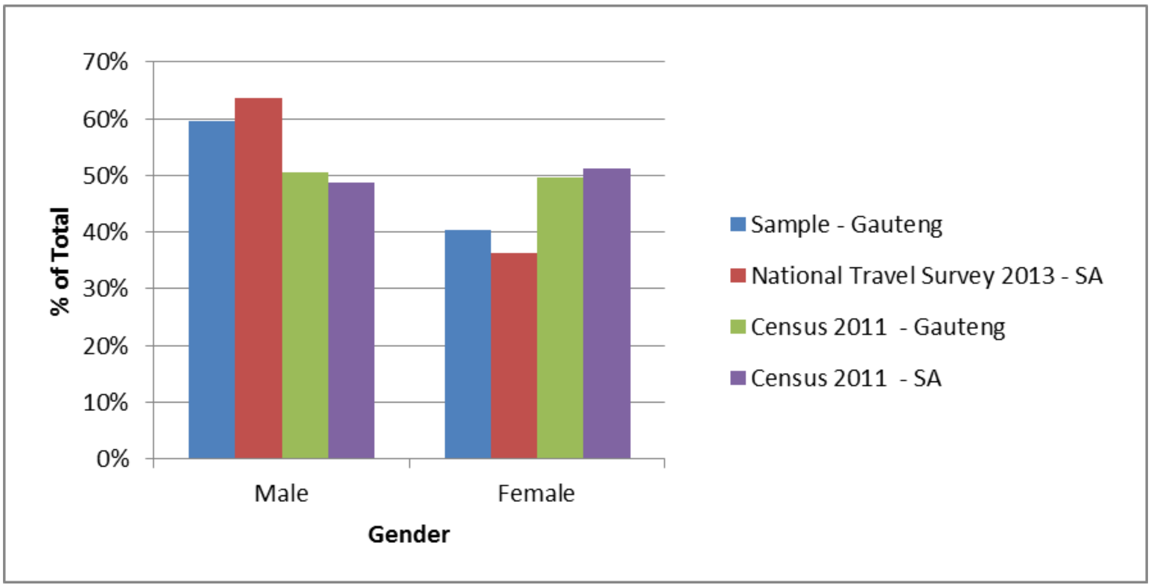


Figure 4-6: Gender Split

Since the Gauteng distribution is not known for both genders and age groups with driving licences, and the sample's distribution is similar to the national driving licence distribution, it is concluded that the sample is sufficiently representative in terms of gender and age groups to support robust analyses for the Gauteng Province.

4.2.2.2 Income Groups

The income information gathered from the survey participants is shown in Figure 4-7. The income distribution of the GPS survey was compared to data from a licence plate survey of freeway users in Gauteng conducted by SANRAL in 2009, in which freeway users were photographed and asked during a telephone interview for personal information including their income. Figure 4-8 shows the comparative results. The GPS sample has an over-representation of the two lower income groups (up to R6 000 and R6 001 to R11 000) and an under-representation of the higher income groups (R11 001 to 20 000 and R20 001 and more) compared to the SANRAL survey.

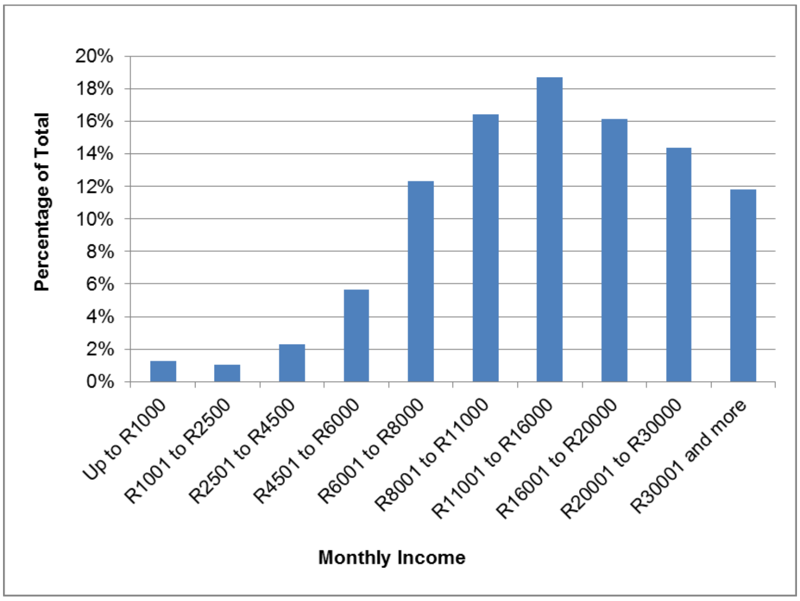


Figure 4-7: Income Distribution within the Sample

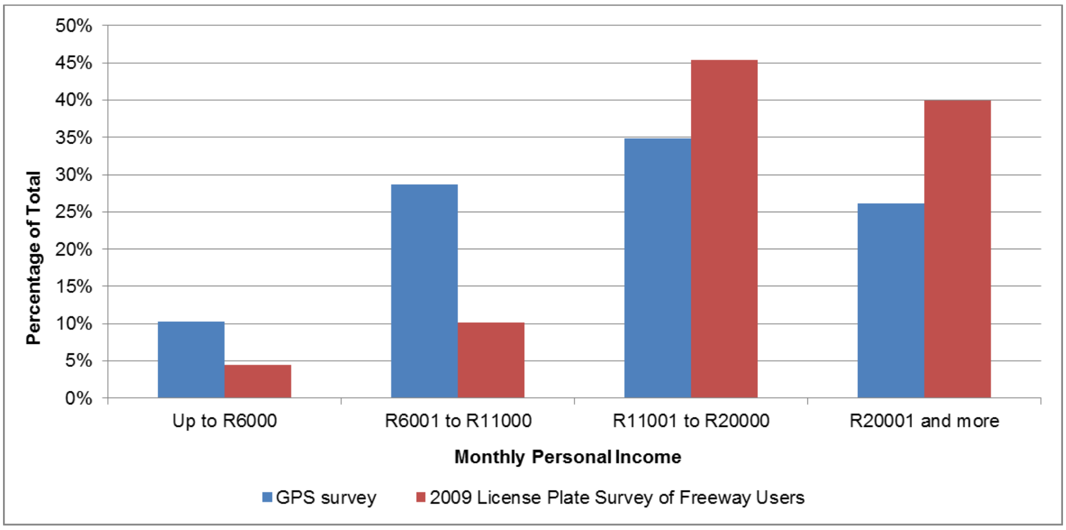


Figure 4-8: GPS survey income groups compared to Licence Plate Survey 2009

4.2.3 Sample Representation

It is concluded that the sample contains some bias with respect to income groups. However, this had limited impact on the results of the trip length investigation as the trip lengths between income groups did not differ significantly, as discussed in Section 4.7. The trip lengths could thus be investigated and the results would representative of the Gauteng province.

4.3 Shopping Centre Trip Lengths

From the sample trip lengths for the following trip types were calculated:

- Home to Shopping Centre;
- Non-home to Shopping Centre;
- Shopping Centre to Home, and
- Shopping Centre to Non-home.

The number of trips, average trip lengths, and standard deviations of the trip types are shown in Table 4-2. It was found that the average and standard deviation of all the trip types were similar. The average trip lengths range from 7.0 to 7.8 and the standard deviations range from 9.9 to 11.4. To investigate the similarities between the trip types, the trip length distributions for each trip type was calculated and is shown in Figure 4-9. The trip length distributions are similar as well.

An ANOVA statistical test was done to determine whether there was a statistical difference between the trip types. The F value from the ANOVA test was 0.43 which was below the F critical value of 2.6, thus according to the ANOVA test there is no significant difference between the average trip lengths of the trip types. The trip types were combined into the following three trip types:

- To Shopping Centres;
- From Shopping Centre, and
- To and From Shopping Centres (combination of first two).

Table 4-2: Average Trip Lengths per Trip Type

Trip Type	Trips Identified in Data Set	Average Trip Length (km)	Standard Deviation (km)
Home to Shopping Centre	134	7.3	11.3
Non-home to Shopping Centre	725	7.8	10.9
To Shopping Centre	**859**	**7.8**	**11.0**
Shopping Centre to Non-home	722	7.8	9.9
Shopping Centre to Home	186	7.0	11.4
From Shopping Centre	**908**	**7.6**	**10.3**
To & From Shopping Centre	**1767**	**7.7**	**10.6**

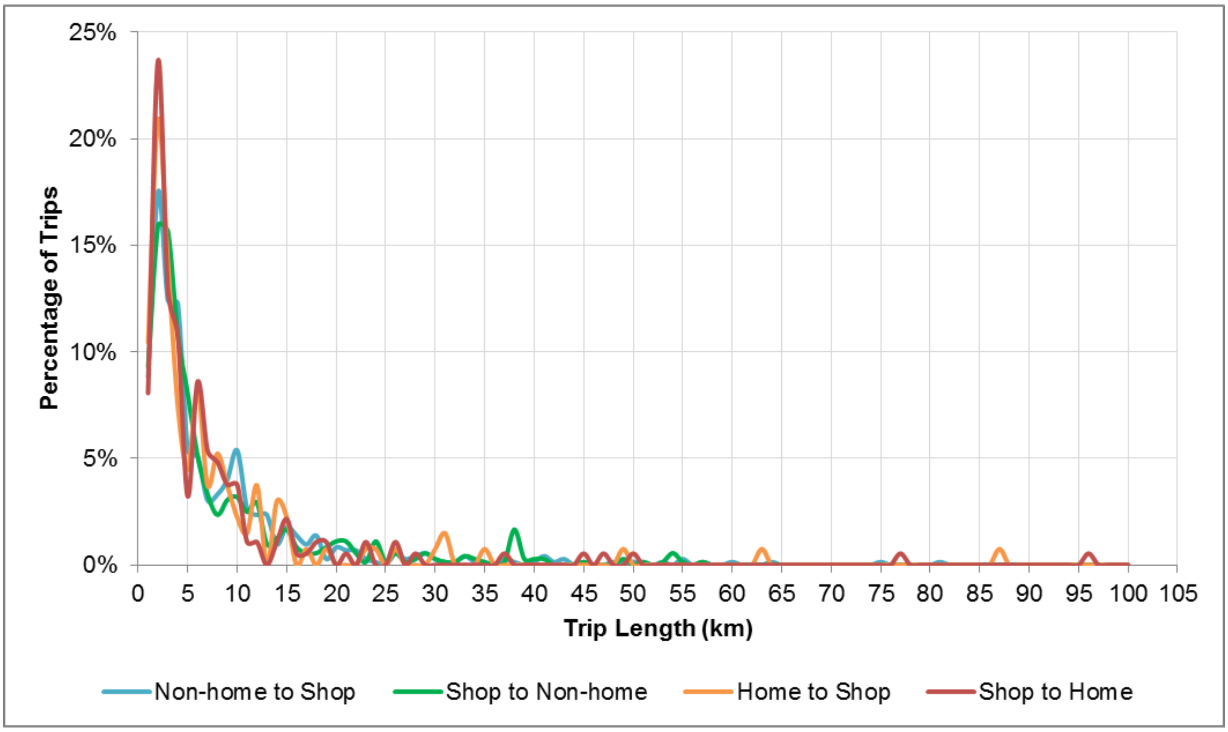


Figure 4-9: Trip Length Distribution per Trip Type

The trip lengths to shopping centres were plotted against the GLA of the shopping centres, as shown in Figure 4-10. This was done for each trip type and is shown in Appendix A. Each blue dot represents a trip length. The linear correlation between trip lengths and GLA was tested with the Pearson Product-Moment Correlation Coefficient (PPMCC) and the R^2 as shown in Table 4-3. The R^2 values range from 0.000 to 0.021 and the PPMCC values range from 0.024 to 0.145. The interpretation of these variables is as follows:

- R^2 1.0 is a perfect linear correlation while 0.0 shows no linear correlation
- PPMCC 1.0 or -1.0 is a perfect linear correlation while 0.0 shows no linear correlation

Both of these statistical variables indicate that there is almost no linear correlation between the trip length and shopping centre size. This is an important finding as it suggests that GLA alone is not a good predictor of shopping trip lengths, and is not appropriate as a base for bulk service contribution calculations. A better indicator might be shopping centre type, as examined in the next section.

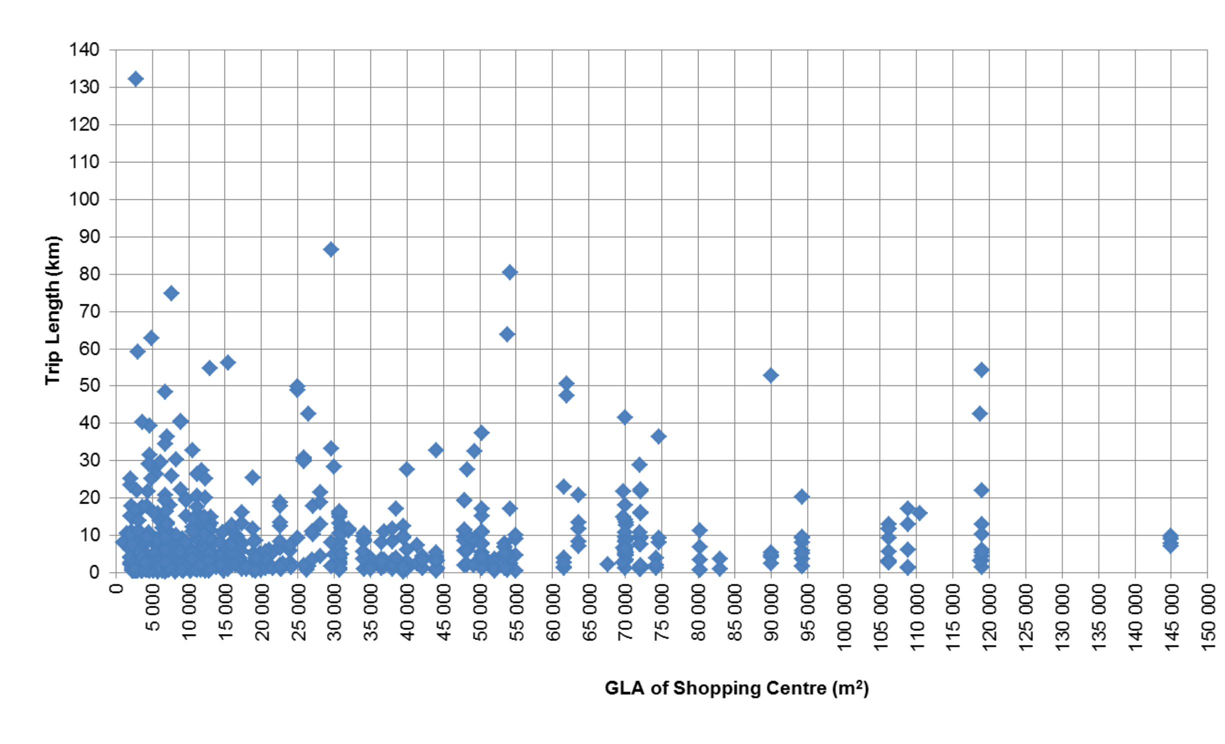


Figure 4-10: Trip Lengths to Shopping Centres shown against Shopping Centre Size

Table 4-3: Linear Relationship between Trip Length and GLA

Trip Type	PPMCC	R^2
Home to Shopping Centre	0.024	0.001
Non-home to Shopping Centre	0.092	0.008
To Shopping Centre	**0.082**	**0.007**
Shopping Centre to Non-home	0.145	0.021
Shopping Centre to Home	0.139	0.019
From Shopping Centre	**0.144**	**0.021**
To & From Shopping Centre	**0.113**	**0.013**

4.4 Average Trip Lengths for Different Shopping Centre Types

The trips were categorised per shopping centre type. The number of trips and percentage of total trips per shopping centre type are shown in Table 4-4 and Figure 4-11. The number of trips relating to the super regional shopping centres is significantly less than those to the other shopping centre types. This is the result of people not visiting the Super Regional Centres as regularly as the other shopping centre types.

Table 4-4: Number of Trips per Shopping Centre Type

Shopping Centre Type	Shopping centre Distribution	Home to Shop		Shop to Home		Non-home to Shop		Shop to Non-home	
		Trips	% of Total	Trips	% of Total	Trips	% of Total	Trips	% of Total
Convenience Centres	32%	22	16%	35	19%	111	15%	116	16%
Neighbourhood Centres	39%	44	33%	67	36%	199	27%	197	27%
Community Centres	16%	21	16%	22	12%	143	20%	146	20%
Small Regional Centres	8%	20	15%	23	12%	137	19%	131	18%
Regional Centres	4%	23	17%	34	18%	107	15%	104	14%
Super Regional Centres	1%	4	3%	5	3%	28	4%	28	4%
Total	100%	134	100%	186	100%	725	100%	722	100%

The Neighbourhood Centres have the largest percentages of trips in this data set while the Super Regional Centres have less than 5% of the total trips. The rest of the trips are spread approximately in similar percentages from 10% to 15%. This follows the distribution of shopping centres per type in Gauteng.

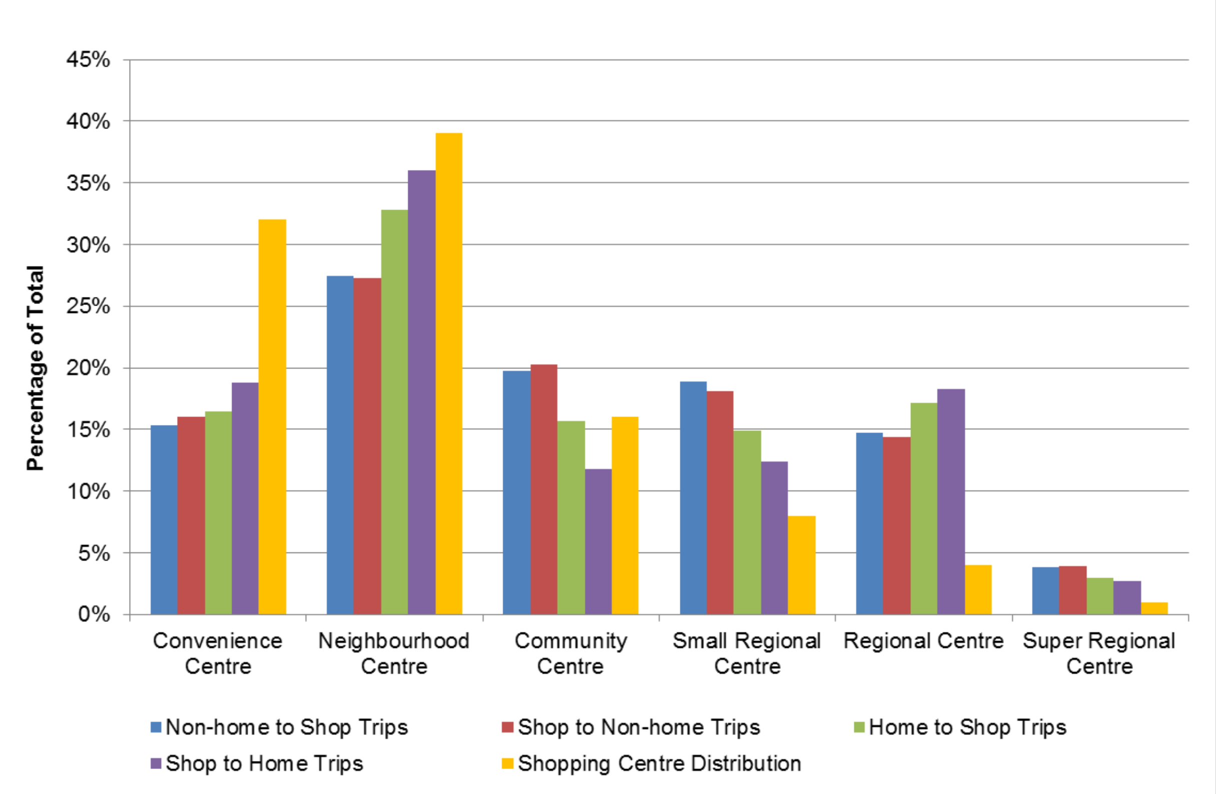


Figure 4-11: Percentage of Trips per Shopping Centre Type

The average trip lengths per shopping centre type are shown in Table 4-5 and Figure 4-12. Generally the average trip lengths increase as the shopping centre sizes increase. Average trip lengths range from 6.0 km to 13.1 km. Regional and Super Regional Centres have the longest average trip length as expected. As mentioned in Section 2.3.5, the attraction of a sufficient number of customers is critical for a shopping centre's feasibility. Figure 4-12 shows that the larger the shopping centre, the larger the shopping centres attraction area must be on average to be feasible.

Convenience Centres have an unexpectedly long average trip length. The reason for this could be that people regularly visit Convenience Centres before or after work. This finding might be related to the size and nature of the study area: as a large metropolitan conurbation, the Gauteng City Region probably has longer trip lengths (especially work trips) than smaller cities. To the extent that Convenience Centres draw more secondary trips (e.g. shops visited as a part of a trip chain between home and work or vice versa), it is possible that these types of centres might be associated with longer average trip lengths. Further research is required to confirm these results and clarify causes.

Table 4-5: Average Trip Lengths per Shopping Centre Type

Shopping Centre Type	GLA (m^2)	Trip Type		
		To Shop (km)	From Shop (km)	To & From Shop (km)
Convenience Centres	500 to 5 000	8.5	7.8	8.2
Neighbourhood Centres	5 001 to 12 000	6.5	6.1	6.3
Community Centres	12 001 to 25 000	6.6	7.6	7.1
Small Regional Centres	25 001 to 50 000	8.2	6.0	7.1
Regional Centres	50 001 to 100 000	9.7	10.7	10.2
Super Regional Centres	More than 100 000	10.3	13.1	11.8
All Shopping Centres	500 to More than 100 000	7.8	7.6	7.7

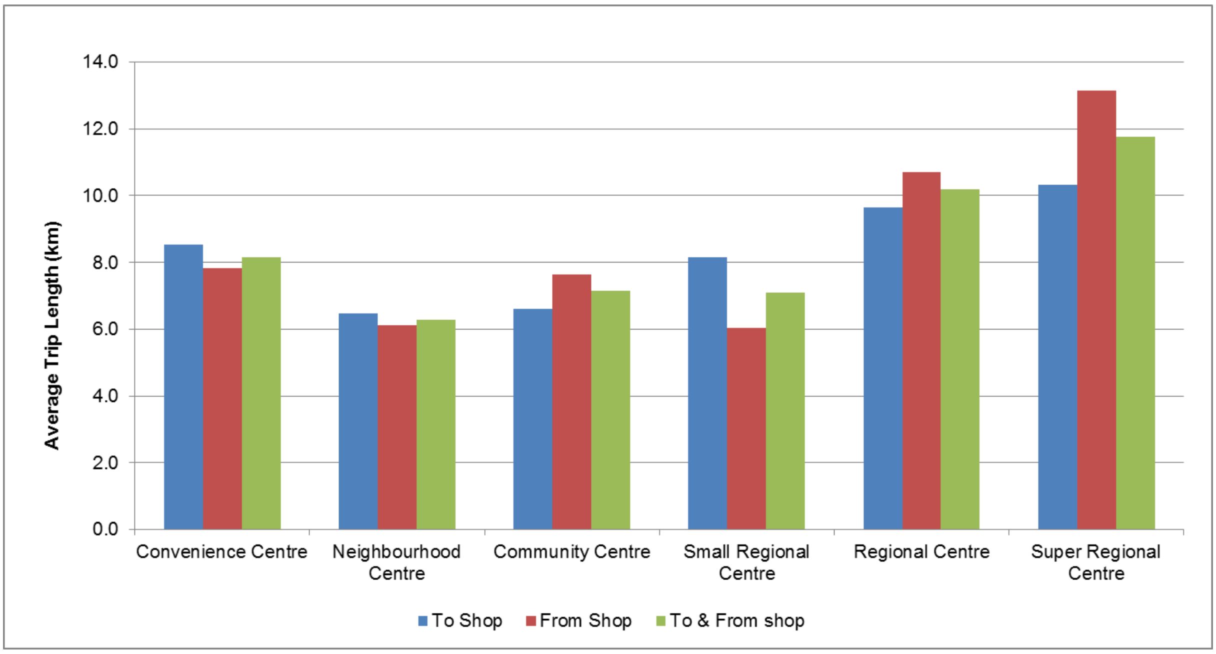


Figure 4-12: Average Trip Length per Shopping Centre Type

The statistical level of confidence of the average trip length results was then investigated.

4.5 Sample Size Compared to Confidence Levels

The number of trips per shopping centre was used to calculate the level of confidence based on sample size, standard deviation and error value. The error value was set as 1 km and the population standard deviation was assumed to be equal to the samples standard deviation for all level of confidence calculations. As shown in Table 4-6 the level of confidence ranges from 99% to 47%. The required sample size for a 95% confidence level was also calculated. The highest number of trips required is for the Regional shopping centres for which 650 trips are required for a 95% confidence level with an error in average trip length of 1 km. The error for the current sample size 95% confidence level is given as well.

Table 4-6: Confidence Levels

Shopping Centre Type	GLA (m^2)	To & From Shop Trip Length (km)	Standard Deviation	Number of Shopping Centre Trips in Sample	Trips Required for 95% Confidence Level and Error of 1 km	Confidence Level with this Sample and Error of 1 km	Error value for 95% Confidence Level with this Sample (km)
Convenience	500 to 5 000	8.2	13.0	284	646	81%	1.51
Neighbourhood	5 001 to 12 000	6.3	9.1	507	318	99%	0.79
Community	12 001 to 25 000	7.1	8.7	332	289	96%	0.93
Small Regional	25 001 to 50 000	7.1	9.0	311	312	95%	1.00
Regional	50 001 to 100 000	10.2	13.0	268	650	79%	1.56
Super Regional	More than 100 000	11.8	12.7	65	622	47%	3.09

It is recommended for future studies that the aim should be to retrieve 650 trips for Regional shopping centres. If this aim is accomplished, all the other shopping centre types, except Super Regional Centres, should have enough trips to reach the 95% confidence level with an error of 1 km. The challenge is to determine how many participants and what duration is required to collect these 650 trips for Regional shopping centres.

The number of participants and the duration required to obtain 650 trips for regional shopping centres is estimated as follows. The first step is to estimate the percentage regional shopping centre trips of the total shopping centre trips. From the GPS data the percentage was calculated to be 15% of total shopping centre trips. Thus 4 334 shopping centre trips are estimated to be required to obtain 650 trips for regional shopping centres. A shopping centre trip rate of 0.84 shopping centre trips per person per day was calculated from the GPS data. With this shopping centre trip rate and the assumption of losing 15% of participant's data, Table 4-7 and Figure 4-13 were derived to give guidance on the duration and number of participants required to collect a sample of shopping centre trips. Using Table 4-7, if the duration is 7 days, then 900 participants are required to collect an estimate of 5 022 shopping centre trips. However, if one decides to reduce the duration to 4 days the participants increase to 1 600 participants for the same number of shopping centre trips.

Table 4-7: Estimated Shopping Centre Trips in Sample with 15% of Participants Data Lost

	Duration of Survey (Days)						
Participants	**1**	**2**	**3**	**4**	**5**	**6**	**7**
200	143	287	430	574	717	861	1004
400	287	574	861	1148	1435	1722	2009
600	430	861	1291	1722	2152	2583	3013
800	574	1148	1722	2296	2870	3444	4018
1000	717	1435	2152	2870	3587	4305	5022
1200	861	1722	2583	3444	4305	5166	6027
1400	1004	2009	3013	4018	5022	6027	7031
1600	1148	2296	3444	4592	5740	6888	8036
1800	1291	2583	3874	5166	6457	7749	9040
2000	1435	2870	4305	5740	7175	8610	10045
2200	1578	3157	4735	6314	7892	9471	11049
2400	1722	3444	5166	6888	8610	10332	12054
2600	1865	3731	5596	7462	9327	11193	13058
2800	2009	4018	6027	8036	10045	12054	14063
3000	2152	4305	6457	8610	10762	12915	15067

It is recommended that the maximum duration should be limited to 7 days. This will allow a survey to record a week's shopping centre travel behaviour. When the survey duration exceeds 7, the same trips made by the same person could get recorded several times. This creates a risk of potentially gathering a biased data set.

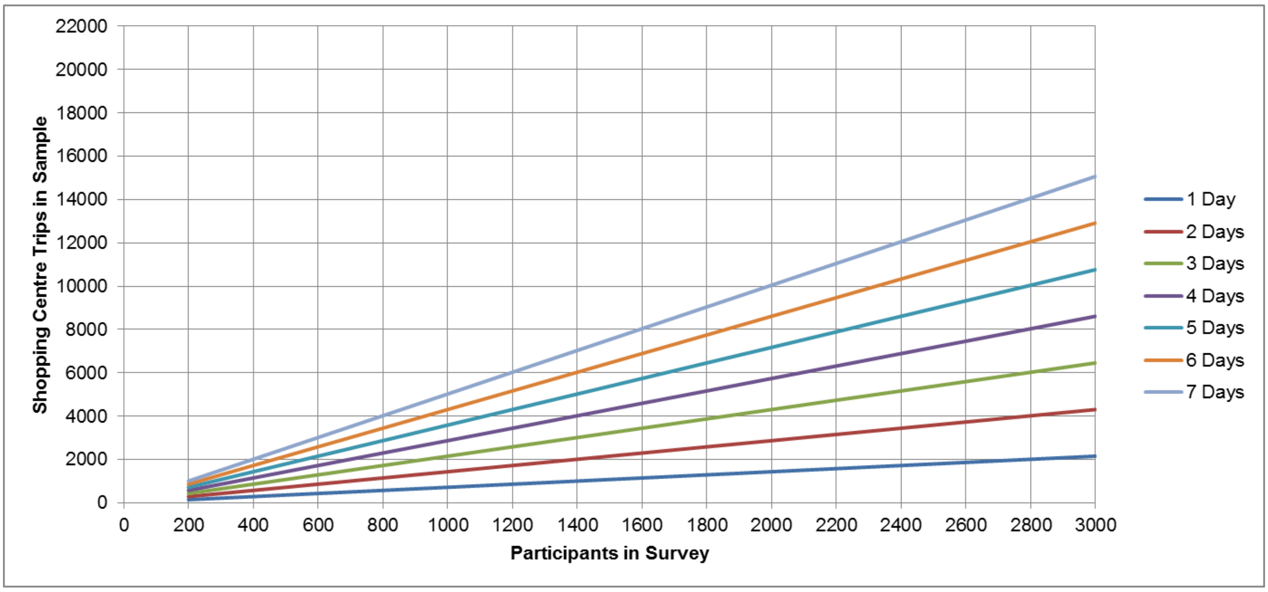


Figure 4-13: Estimated Shopping Centre Trips in Sample with 15% of Participants Data Lost

Both the number of participants and duration can be decreased if the percentage of participants' data lost is decreased through tighter quality control during the data collection process. Appendix E contains tables for sample sizes corresponding to 0% to 30% participant data loss.

4.6 Trip Length Frequency Distributions

The purpose of this section is to estimate best-fitting distributions to the trip lengths for each type of shopping centre. This would be valuable for future updating of the findings for the purposes of transferring the results to other areas, as smaller trip samples obtained locally could be used to estimate the parameters under the assumption that the distributional form remains constant. From the literature (Section 2.5) it was recommended the Gamma or Weibull distributions should be used for the trip length frequency distributions. The exponential distribution was added as a third distribution type.

Trip length frequency distributions were calculated with intervals of 1 km. A frequency distribution function was estimated by fitting Gamma, Weibull and Exponential distributions through the trip length frequency points for each trip type (To Shop, From Shop and To & From Shop). The least squares method was used to estimate the variables of the functions. The trip length frequency Gamma distributions for trips to shopping centres are shown in Figure 4-14. Appendix B contains the distributions for the other trip types and distributions.

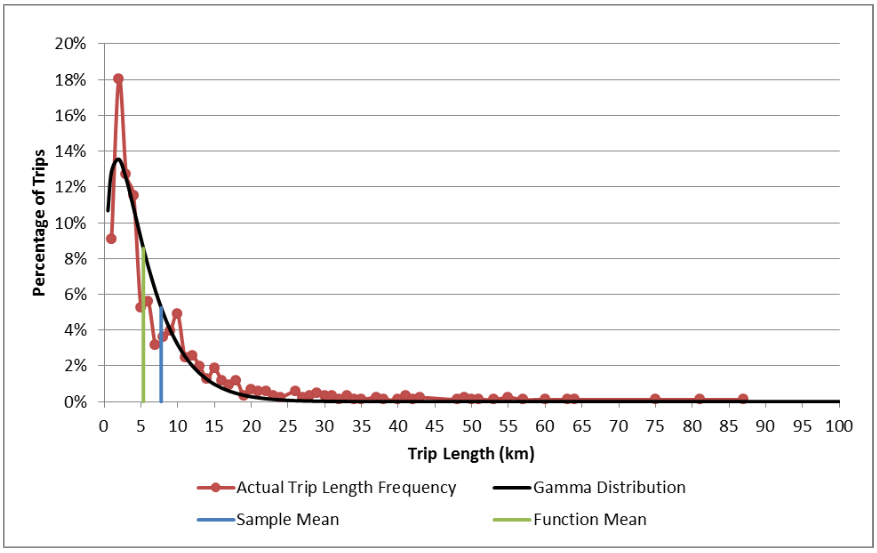


Figure 4-14: TLFD - To Shopping Centre – Gamma Function Mean Not Equal to Sample Mean

The least squares method returns a function with a distribution mean not equal to the sample's mean, as shown in Figure 4-14. This occurs for all three distribution functions. The difference between the sample mean and the Gamma function mean is approximately 2.4 km. The least squares method can be adjusted to minimise the sum of the squared errors while still providing a function that has a mean equal to the sample's mean, as shown in Figure 4-15. However, this increases the sum of the squared errors and decreases the R^2 values slightly.

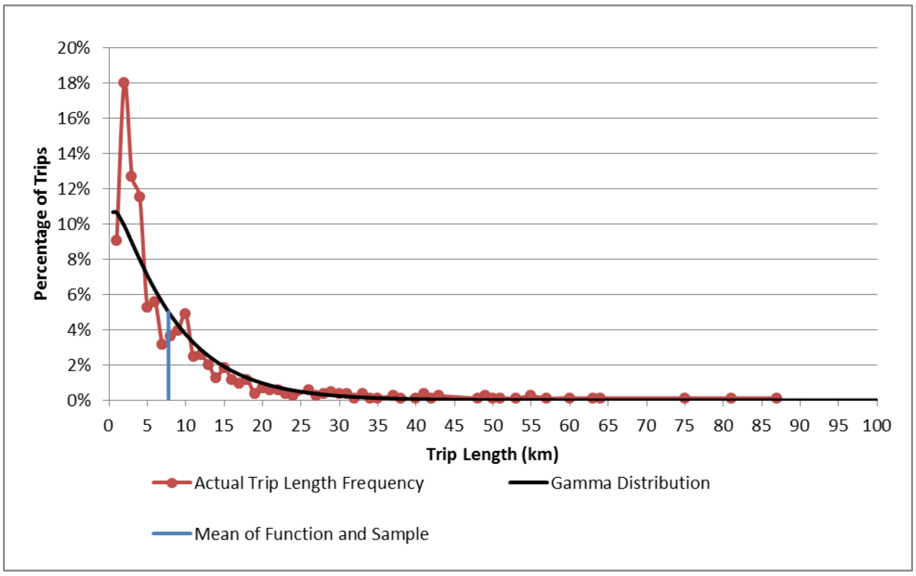


Figure 4-15: TLFD – To Shopping Centre – Gamma Function Mean Equal to Sample Mean

The comparison between distribution functions is shown in Figure 4-16 and Figure 4-17. Statistical values used to evaluate the goodness of the functions fitted are shown in Table 4-8. Refer to Appendix C for similar tables and graphs for each of the trip types.

Table 4-8: Goodness of Fit Evaluation for Trips to Shopping Centres

Trip Type: To Shopping Centre						
Variable	Function Mean Not Equal to Sample Mean			Function Mean Equal to Sample Mean		
	Gamma	Weibull	Exponential	Gamma	Weibull	Exponential
R^2	0.891	0.908	0.861	0.832	0.867	0.829
*RMS Error	0.012	0.012	0.013	0.015	0.015	0.015
Error Squared	0.007	0.008	0.009	0.011	0.011	0.011
**Mean Diff (km)	-2.4	-2.5	-1.9	0.0	0.0	0.0

* RMS: Root Mean Square

** Mean Diff: The difference between function mean and sample mean

*** Green shading indicates best fit value

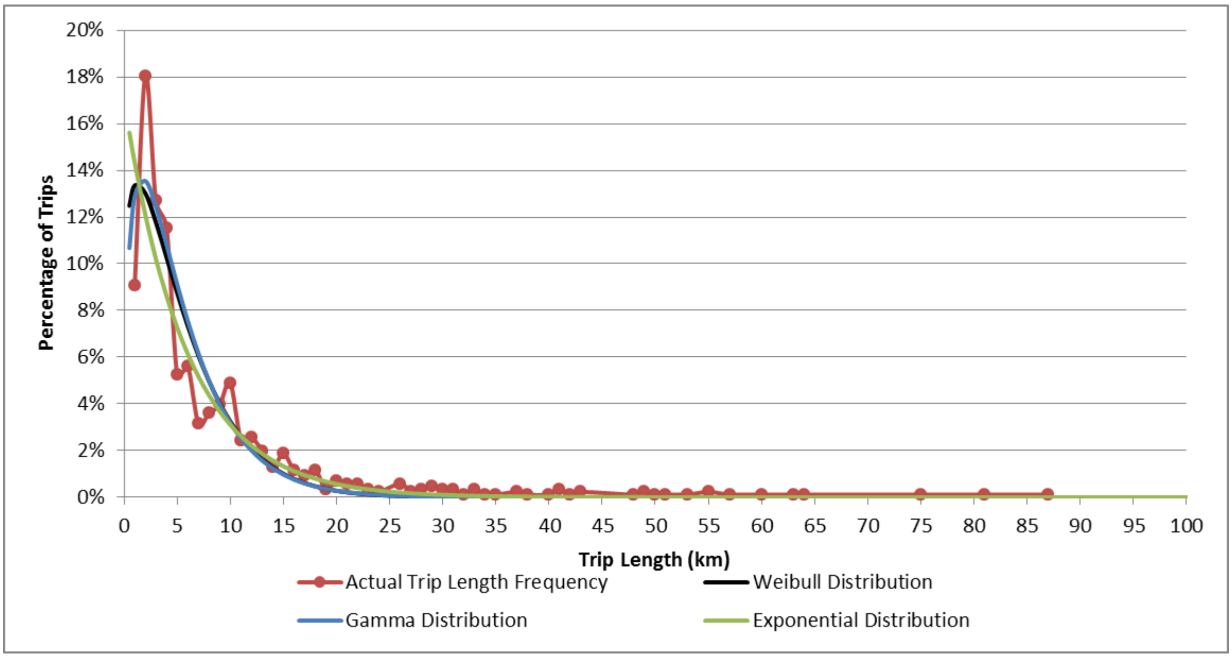


Figure 4-16: To Shop Trip Length Comparison, Sample Mean Not Equal to Function Mean

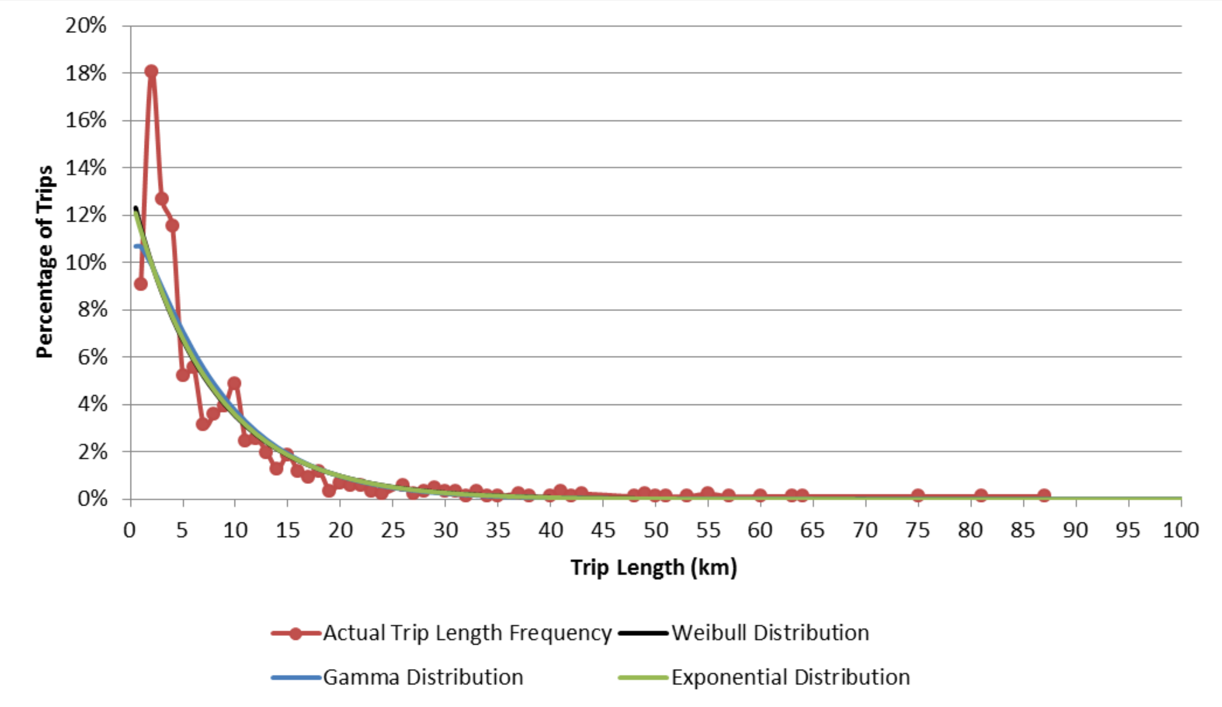


Figure 4-17: To Shop Trip Length Comparison, Sample Mean Equal to Function Mean

As shown in Table 4-8, Figure 4-16 and Figure 4-17 the three distribution functions produce similar distribution functions. The closer the R^2 value is to 1.0 the better the fit, the closer to zero the RMS error and error squared values are, the better the fit. The least difference between the sample mean and the function mean was also used to evaluate the goodness of fit. The distribution functions were compared and the best fit statistical value was marked green in Table 4-8. This was also done for the other trip types as shown in Appendix C. It was found that the Weibull distribution fits the best for all three trip types for both scenarios. The bulk contribution calculations uses the average trip length value, thus it was decided that the Weibull distribution function's average trip length needs to be equal to the sample's mean. The best fit Weibull trip length distribution function for each trip type is shown in Figure 4-18. The distributions are virtually the same.

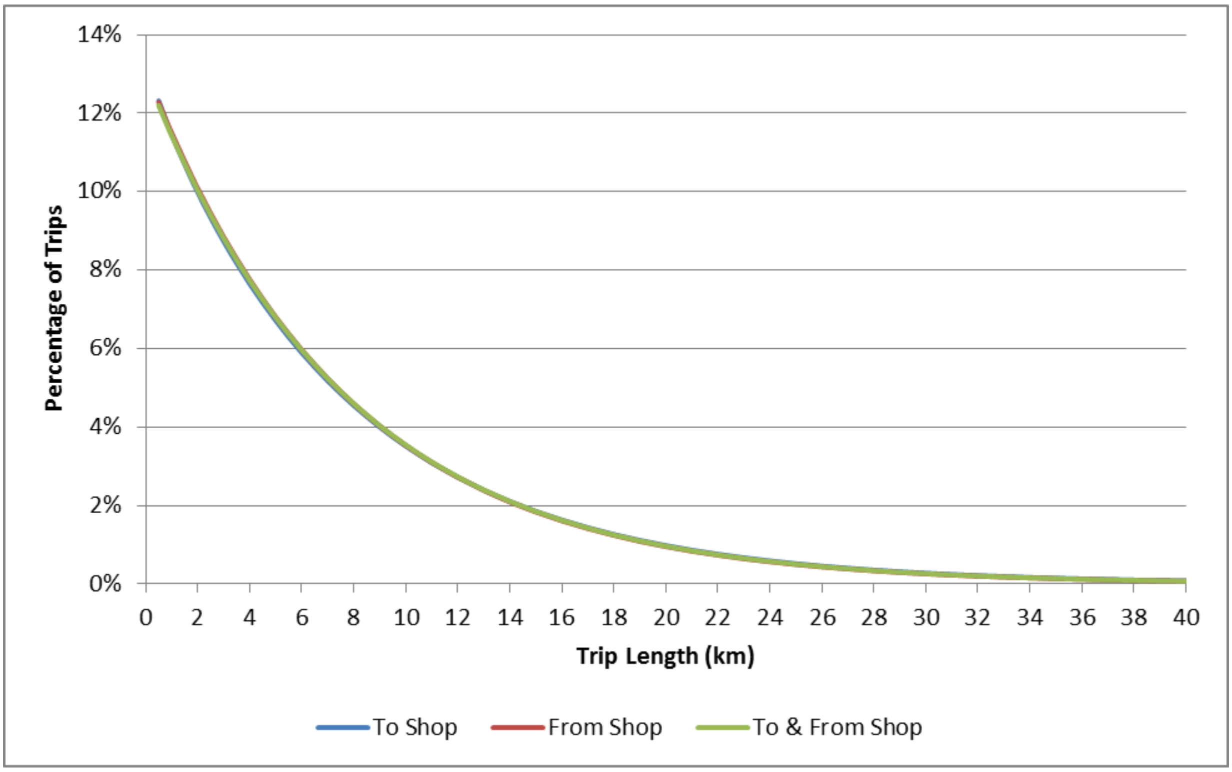


Figure 4-18: Weibull TLFD for All Shopping Centres

Since the average trip length variable in bulk contributions is based on both trips to and from the shopping centres, only the trip type To & From Shop was used for the Weibull distribution functions for each shopping centre type. The best fit Weibull distribution with the sample mean equal to the function mean for each shopping centre type is shown in Figure 4-19 and the distribution parameters are given in Table 4-9. The distribution for each shopping centre type is given separately in Appendix D.

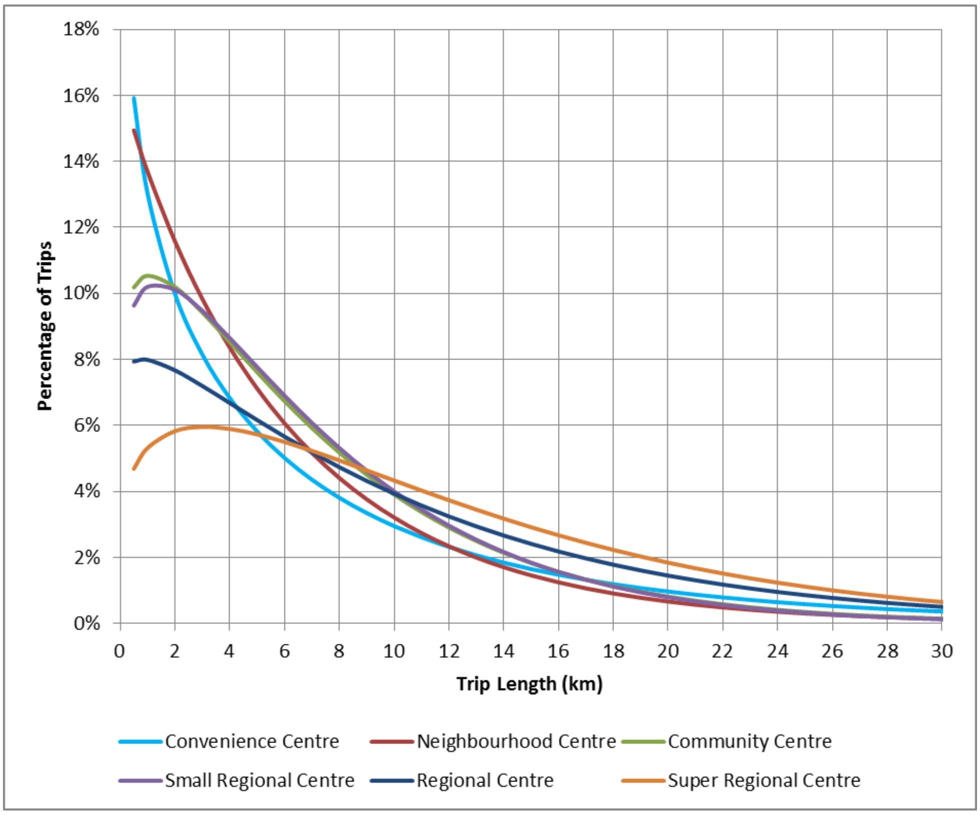


Figure 4-19: Weibull Distributions for Shopping Centre Types

Table 4-9: Weibull Distributions Parameters for Shopping Centre Types

Shopping Centre Type	Alpha (α)	Beta (β)
Convenience Centre	0.83	7.41
Neighbourhood Centre	0.99	6.26
Community Centre	1.13	7.45
Small Regional Centre	1.16	7.48
Regional Centre	1.07	10.48
Super Regional Centre	1.22	12.56

In Table 4-10 the goodness of fit is evaluated for each shopping centre type. The R^2 value is higher than 0.7 for all shopping centre types, except for the Super Regional Centres which has a value of 0.369. Generally in transportation engineering field a R^2 value of 0.7 is acceptable for function to fit fairly well through data points. The low R^2 value for Super Regional Centres could be a result of the small sample size and the function mean being forced to be equal to the sample mean. The Super Regional Centre distribution function gives similar results to the other statistical parameters. The RMS Error and Error Squared values are all close to zero. Thus the distribution functions fitted are deemed to be acceptable, although the results for Super Regional Centres should be used with caution.

Table 4-10: Weibull Distribution for Shopping Centre Types Statistical Values

Shopping Centre Type	R^2	RMS Error	Error Squared
Convenience Centre	0.702	0.025	0.022
Neighbourhood Centre	0.795	0.021	0.018
Community Centre	0.823	0.017	0.010
Small Regional Centre	0.724	0.023	0.018
Regional Centre	0.789	0.014	0.008
Super Regional Centre	0.396	0.021	0.012

4.7 Average Trip Length per Income Group

During the respondent surveys, participants were given ten categories to select from to indicate their monthly income:

- R 0 to R 1 000;
- R 1 001 to R 2 500;
- R 2 501 to R 4 500;
- R 4 501 to R 6 000;
- R 6 001 to R 8 000;
- R 8 001 to R 11 000;
- R 11 001 to R 16 000;
- R 16 001 to R 20 000;
- R 20 001 to R 30 000, and
- R 30 001 or more.

The income ranges were categorised into High, Middle and Low income group categories as shown in Table 4-11. A total of 390 (78% approximately) of the 502 participants provided an income response. The lower income group (up to R8 000) constitutes 23% of all participants, while the medium income group (R8 001 to R30 000) contains 66% and the high income group (R30 0001 and more) contains 12% of the sample.

Table 4-11: Income Groups

Income Group	Monthly Income (After Income Tax)	% of Total Participants Providing Income Data
Low	R0 to R8 000	23%
Middle	R8 001 to R30 000	66%
High	More than R30 001	12%

It was reasoned that the low income group could have the shortest average trip length due to financial constraints while the medium income group could have the longest average trip length. The high income group was expected to have an average trip length between the low and medium income groups. It was reasoned that the high income group was located in well-developed areas with shopping centres in close vicinity. The medium income group was assumed to be located in areas with fewer shopping centres, while still having the financial means to travel to preferred shopping centres.

Income ranges for Low, Medium and High Income respondents are given in Table 4-11. The trip lengths from the participants who did not provide an income range were excluded from this analysis. A total of 390 (78% approximately) of the 502 participants provided income data. The percentage of the total trips per income group is shown in Figure 4-20. The middle income group has more than 50% of all the shopping centre related trips in this data set, while the low and high income groups have approximately 20% of the trips each.

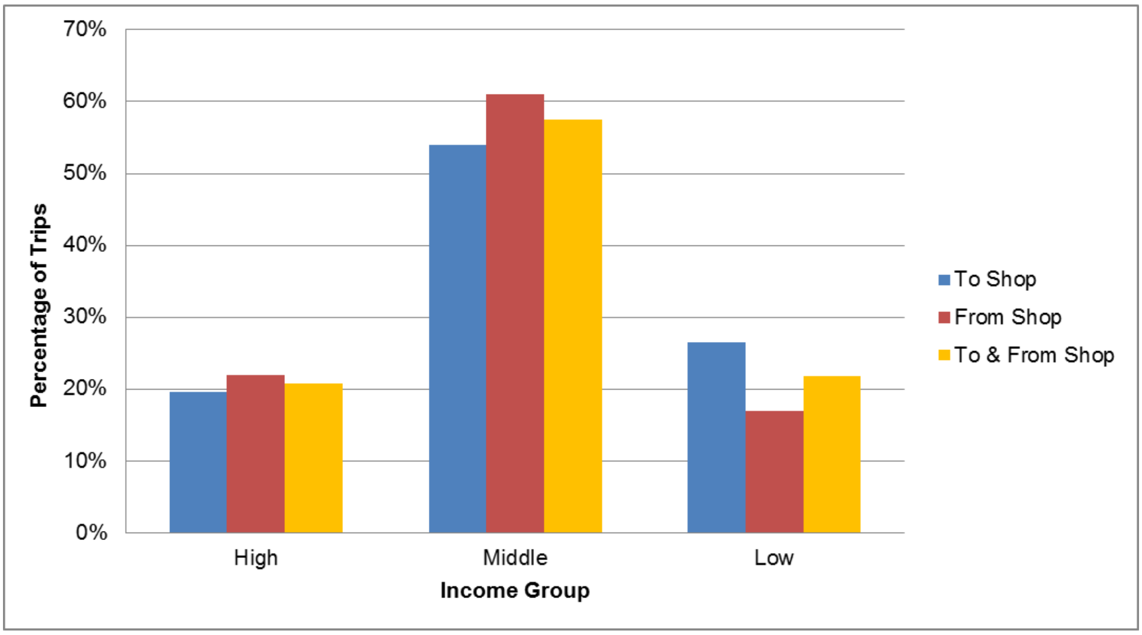


Figure 4-20: Percentage of Total Trips per Income Group

The shopping centre trip rate per income group per day was calculated with the data as shown in Table 4-12. As expected the trip rate for the low income group is significantly lower than for the medium and high income groups and the high income group has the largest trip rate.

Table 4-12: Shopping Centre Trip Rates per Income Group

Income Group	Shopping Centre Trips per Day per Person
High	0.51
Middle	0.42
Low	0.34

The average trip lengths per income group and trip type are shown in Table 4-13 and Figure 4-21.

Table 4-13: Trip Lengths per Income Group

Trip Type	Income Group	Number of Trips	Average Trip Length	Standard Deviation (km)
To Shopping Centre	High	91	7.86	12.11
	Medium	251	7.45	9.71
	Low	123	7.07	8.50
From Shopping Centre	High	98	9.34	14.74
	Medium	273	6.97	8.40
	Low	76	6.44	7.02
To & From Shopping Centre	High	189	8.63	13.52
	Medium	524	7.20	9.06
	Low	199	6.83	7.96

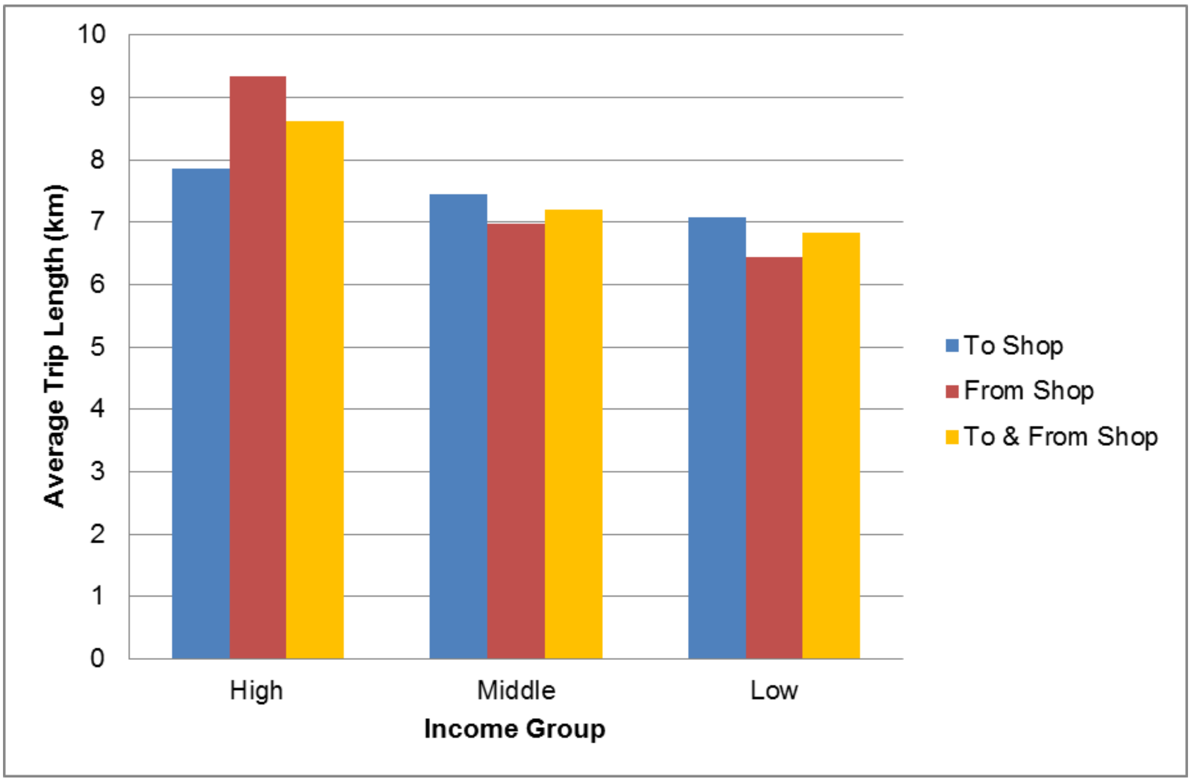


Figure 4-21: Average Trip Length per Income Group

The average trip length per income group has a downward trend with the high income group having the longest average trip length and the low income group the shortest average trip length. The average trip length per income group and per trip type has a range of 6.44 km to 9.34 km. This is a difference of 2.90 km between maximum and minimum average trip length. An ANOVA test was done to determine if the differences were statistically significant. The F value (1.24) was less than the F critical value (1.94) thus the differences in average trip length are statistically insignificant. The reason for this is evidently the high degree of variation, indicating that there is high variability in individual shopping trip behaviour.

4.8 Average Trip Length per Gender

The shopping centre related trips were categorised per gender as shown in Table 4-14. In general the split between the number of female trips to male trips was approximately 40:60.

Table 4-14: Gender Split of Trips

Trip Type	Male		Female		Total	
	Trips	% of Total	Trips	% of Total	Trips	% of Total
To Shop	517	60%	342	40%	859	100%
From Shop	543	60%	365	40%	908	100%
To & From Shop	1060	60%	707	40%	1767	100%

The average trip length per gender is shown in Table 4-15 and Figure 4-22 for all trip types. The average trip lengths for the male participants were more than for the female participants.

Table 4-15: Average Trip Length per Gender

Trip Type	Male		Female	
	Average Trip Length (km)	Standard Deviation (km)	Average Trip Length (km)	Standard Deviation (km)
To Shop	7.9	10.5	7.5	11.7
From Shop	7.9	11.0	7.2	9.0
To & From Shop	7.9	10.8	7.3	10.4

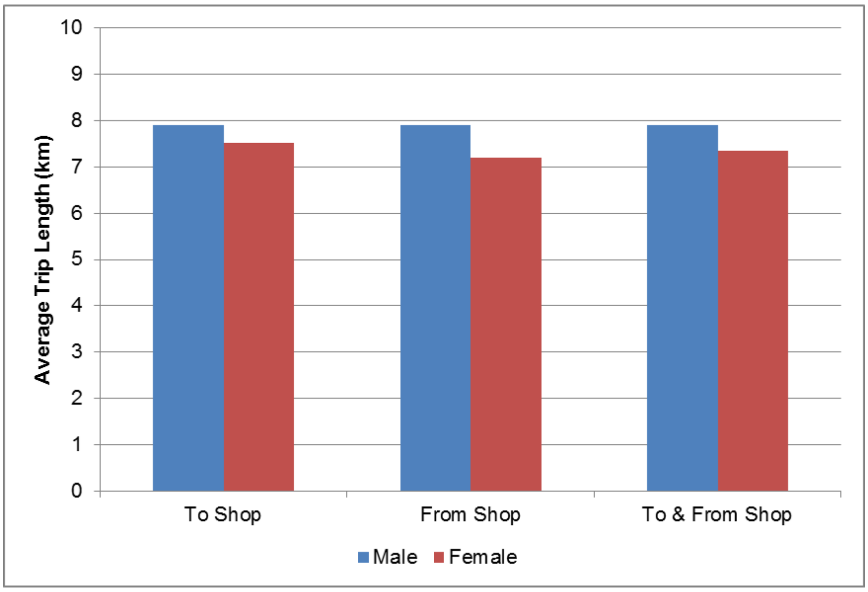


Figure 4-22: Average Trip Length per Gender

An ANOVA test was done to determine if the differences were statistically significant. The F value (0.51) was less than the F critical value (2.22), thus the differences in average trip length are statistically insignificant.

4.9 Shopping Centre Trip Lengths per Class of Road

The bulk contribution calculation excludes the distance travelled on Class 4 and Class 5 roads as well as roads under jurisdictions of other authorities like SANRAL (Class 1 roads). This creates additional challenges for calculating the average trip length for a development. From the trip length data the distance travelled on Class 1, Class 2-3 and Class 4-5 was calculated for the shopping centre trips. The percentage per Class of road was then calculated as shown in Figure 4-23.

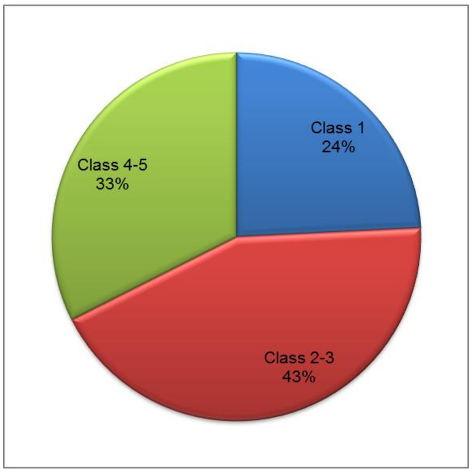


Figure 4-23: Average Trip Length Percentage per Class of Road

If one assumes that Class 2-3 roads are under the jurisdictions of the municipality, only 43% of the average shopping centre trip length should be used in bulk contributions calculations according to this sample. The impact of this and other findings on the accuracy of bulk service contribution calculations is examined in Chapter 5.

5. Impact on Bulk Contribution Calculations

5.1 Average Trip Lengths for Shopping Centres

The trip length value currently recommended by COTO (2012) is given in the TMH17. The average trip lengths given in the TMH17 is based on data from South Africa and the United States of America (USA). Owing to the insufficient amount of local data on average trip lengths, the data was supplemented with studies done in Florida in the USA. The TMH17 average trip length per shopping centre size is compared to the average trip length per shopping centre type calculated from the GPS data set. The comparison is shown in Figure 5-1 together with the upper and lower 95% confidence levels.

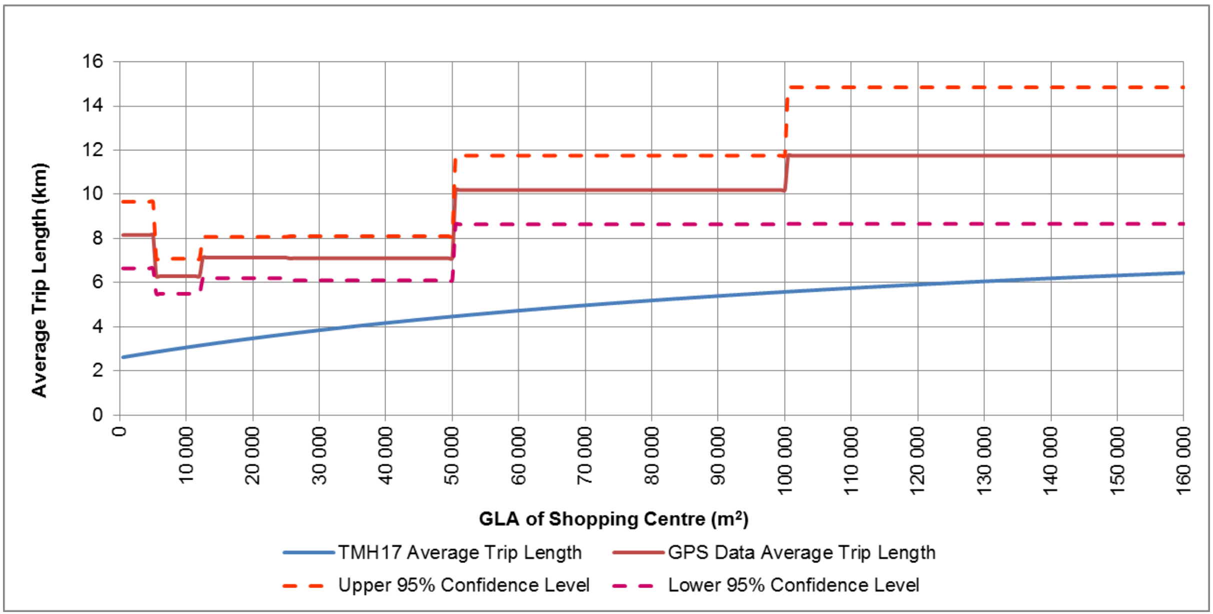


Figure 5-1: TMH17 Compared to GPS Data

The average trip lengths calculated from the GPS data are longer than the recommended TMH17 trip lengths for all GLA values. Owing to the trip lengths being calculated per shopping centre type, the average trip length is constant for the range of the specific shopping centre type and then increases or decreases in a step function for the next shopping centre size. The TMH17 has a curved shape (refer to Equation 2-10) which increases gradually with the size of the shopping centres. The differences in average trip lengths were calculated based on the midpoint GLA value for each shopping centre size, as the latter was used in the formula (Equation 2-10) provided to calculate the TMH17 recommended average trip length. This was then compared to the average trip length of the specific shopping centre type as shown in Table 5-1. Significant differences are observed; differences range between 3.0 km and 5.4 km.

Table 5-1: TMH17 Compared to GPS Data

Shopping Centre Type	GLA (m^2)	Middle GLA / GLA used in TMH17 Calculation (m^2)	TMH17 Average Trip Length (km)	GPS Data Average Trip Length (km)	Difference (km)
Convenience	500 to 5 000	2 750	2.7	8.2	5.4
Neighbourhood	5 001 to 12 000	8 500	3.0	6.3	3.3
Community	12 001 to 25 000	18 500	3.4	7.1	3.7
Small Regional	25 001 to 50 000	37 500	4.1	7.1	3.0
Regional	50 001 to 100 000	75 000	5.1	10.2	5.1
Super Regional	More than 100 000	150 000	6.3	11.8	5.4

Even though further research is required to confirm the average trip lengths results, three options were evaluated to possibly revise the recommended TMH17 average trip length values. These options are:

- Propose a new function based on the observed relationship between GLA and trip lengths;
- Revise current function based on the observed relationship between GLA and trip lengths, and
- Propose a new step function based on the observed relationship shopping centre type and trip lengths.

For the first option a straight line was fitted through the average trip length corresponding to the midpoint GLA of each shopping centre type as shown in Figure 5-2. The straight line function has the best fit compared to higher order functions and has an acceptable R^2 value of 0.82 with a positive gradient. This line function can be used to estimate average trip lengths for shopping centres based on the GLA.

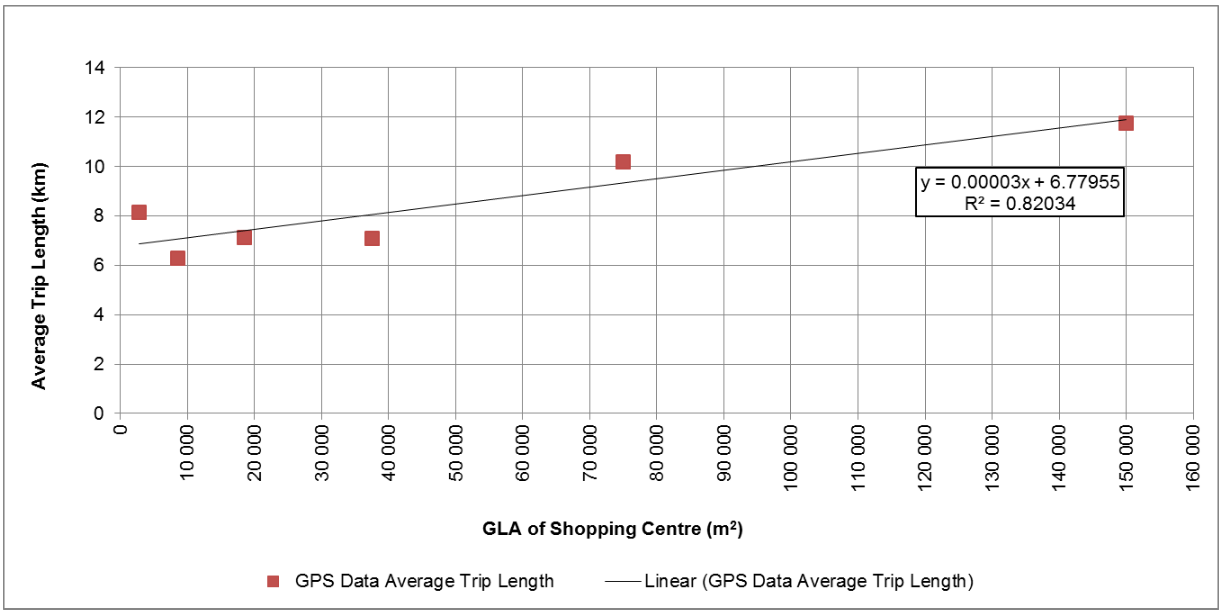


Figure 5-2: Proposed Line Function for Estimating Average Trip Lengths

For the second option a constant was added to the current function to adjust the average trip length to be similar to the results obtained from the GPS data. The least squares method was used to calculate the constant value of 4.8 km (rounded to one decimal) that has to be added to the original TMH17 function to have the estimated average trip lengths similar to the GPS data, as shown in Figure 5-3.

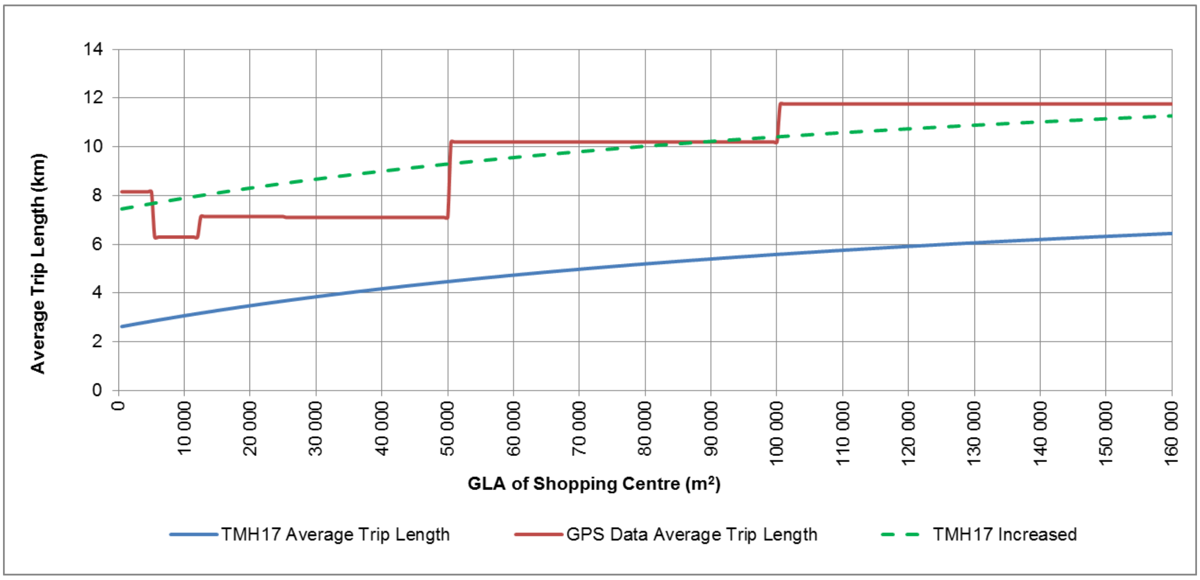


Figure 5-3: Proposed Revised TMH17 Function

The third option is to classify shopping centres based on their GLA (refer to Section 2.3.4) and to estimate average trip lengths based on the shopping centre type.

All three options are shown in Figure 5-4. Both the new function and the adjusted TMH17 function under- or over-estimate the average trip lengths as compared to the average trip length per shopping centre type obtained from the GPS data. It is proposed that the average trip lengths for shopping centres should be estimated based on shopping centre type, as this is most accurate and based on empirical evidence.

Figure 5-4: Proposed Average Trip Length

5.2 Shopping Centre Average Trip Length on Road Class 2 and 3

In the bulk contribution calculations the total shopping centre trip length is first halved and then has to be decreased to remove distance travelled on Class 4-5 roads and travel on roads under jurisdiction other than the municipality. The TMH17 recommends the following steps:

- Calculate total average trip length with adjustment factor;
- Halve the trip length;
- Multiply the halved trip length with 0.6 (40% of trip length distance is on roads under other jurisdiction), and
- Subtract 1 km for travel on Class 4-5 roads.

With this method if the total trip length was 10 km, only 2 km would be used further in the bulk contributions calculations.

The average trip length per road Class was calculated from the GPS data set as discussed in Section 4.9. According to the GPS data 43% of the travelled distance occurs on Class 2-3 roads for shopping centres. It should be noted that within the urban areas of Gauteng many Class 2 roads are under the jurisdiction of the municipalities. Thus it was assumed that the distance travelled on the Class 2 roads should be included in the bulk contribution calculation.

The 43% calculated from the GPS data was then compared to the TMH17 method with an assumed total trip length of 10 km. Firstly the 10 km is halved, resulting in 5 km. Multiplying the 5 km with the 43% results in an adjusted halved trip length of 2.15 km. The TMH17 method multiplies the 5 km with 60% and then subtracts 1 km as discussed above. This results in an adjusted halved trip length of 2.0 km. Thus there is no significant difference in the result between the TMH17 method or the GPS calculated factor. However, using one factor to decrease the halved trip length is far less complicated.

5.3 Impact on Bulk Contributions

The difference in results for the average adjusted half trip length ($L_D/2$ in Equation 2-4) for shopping centres between the proposed shopping centre type and the current THM17 method is shown in Figure 5-5. There is a significant difference in the results.

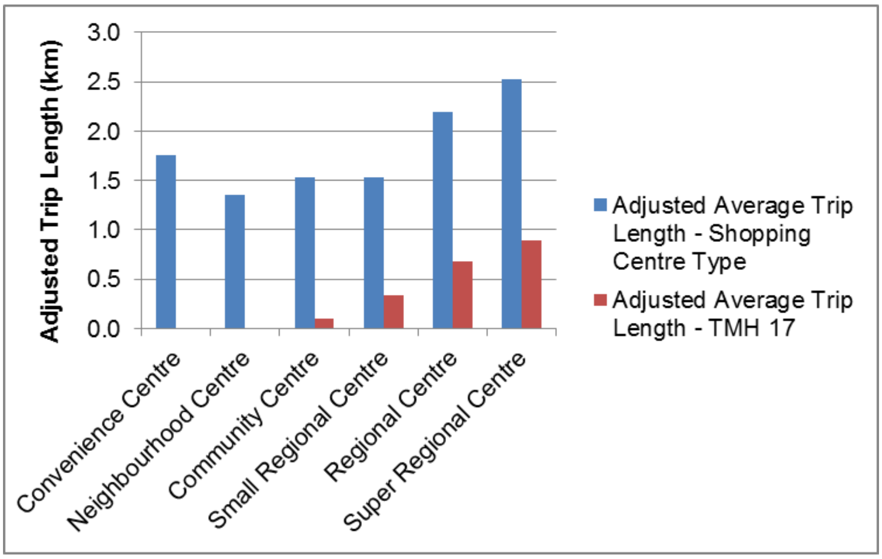


Figure 5-5: Comparison of Adjusted Average Trip Length

Substituting Equation 2-5 into Equation 2-4, Equation 5-1 is derived.

Equation 5-1: Capacity Component for a Particular Land Use (C_{AQD})

$$C_{AQD} = A_D * F_{QD} * AADT_D * (L_D/2) * R_Q$$

Where:

C_{AQD} = Capacity component for a particular land use

A_D = Size of the land use rights in appropriate size units

T_D = Impact trip rate (trips per hour per size unit)

F_{QD} = Traffic factor to convert AADT to an impact trip rate

$AADT_D$ = AADT trip generation rate (trips per day per size unit)

$L_D/2$ = Half adjusted average trip length (km)

R_Q = Cost rate per veh-km/hour for the capacity component

For the Convenience and Neighbourhood centres the C_{AQD} contribution will result in zero when using the TMH17 half adjusted length of 0 km for Convenience and Neighbourhood centres in Equation 5-1. The TMH17 estimates that the entire halve average trip lengths of the Convenience and Neighbourhood centres are travelled on Class 4 and 5 roads. According to the proposed method the halve adjusted trip length values for Convenience and Neighbourhood centres should be 1.8 km and 1.4 respectively. If the proposed method is implemented, this will have a significant impact on the C_{AQD} contribution amount for these shopping centre types. For the Community, Small Regional, Regional and Super Regional on average the proposed adjusted halve trip length is 2.86 times longer than the TMH17's adjusted halve trip lengths. Since the $L_D/2$ is directly proportional to the C_{AQD}, the C_{AQD} contribution amount will increase on average by 286%.

If it is assumed that the delivery vehicles will have the same adjusted halve trip lengths as the light vehicles, the strength component contribution (Equation 2-7) will increase similarly to the capacity component contribution. The assumption that delivery vehicles will have the same adjusted halve trip lengths as the light vehicles is made as a result of not having any data on delivery vehicle average trip lengths.

The increase in the adjusted halve trip lengths will significantly increase the road contributions for shopping centres.

6. Conclusions and Recommendations

6.1 Conclusions

The bulk contributions calculation for roads contains a variable for the halved adjusted trip length. This variable is the average distance travelled to or from a development which is then halved and adjusted to only reflect the distance travelled on roads under the jurisdiction of the municipality excluding distance travelled on Class 4 and Class 5 roads. It was found that limited data are available for average trip length per land use type including shopping centres. Currently shopping centre trip lengths are estimated using the TMH17 document for calculating bulk contributions. The average trip lengths given in the TMH17 are based on South African and supplemented with data from the USA due to the scarceness of trip length data in South Africa. This dissertation investigated the possibility of using GPS data to estimate trip lengths, specifically for shopping centres.

The GPS data used for this study were obtained from a previous study undertaken at the University of Pretoria. The purpose of the original study was to evaluate the impact that the E-toll system had on driving patterns in the province. After receiving the raw GPS data, the data was cleaned. An investigation was done on what methods are available to identify trip ends in large GPS data sets. From this investigation it was found that the criteria of the stop time (110 seconds) and repeated use of road links would be the most suitable for this data set. The trip ends were then successfully determined. Trip ends were then compared with the database of shopping centres to identify shopping centre trips. With the trip ends identified and the trips ending at shopping centres, trip lengths could be calculated for shopping centre trips. This indicates the potential of future study approaches using GPS technology to collect more accurate trip length data for South Africa.

The average trip lengths between gender and income groups did not differ significantly. ANOVA tests were done to determine if the differences were statistically significant and the ANOVA tests showed the differences were not statistically significant. The reason for this is evidently the high degree of variation, indicating that there is high variability in individual shopping trip behaviour. The Weibull distribution statistically has the best fit for the trip length frequency distributions and Weibull distribution functions were calculated for each shopping centre type. The Weibull distribution functions can be used as the basis for transferring these results to other areas in South Africa.

The sample (GPS data) contains some bias with respect to income groups. However, this had limited impact on the results of the trip length investigation as the trip lengths between income groups do not differ significantly. As expected the average trip length increases generally as the shopping centre size increases. However, statistical variables indicate that there is almost no linear correlation between the trip length and shopping centre size. This is an important finding as it suggests that GLA alone is not a good predictor of shopping trip lengths, and is not appropriate as a base for bulk service contribution calculations. A better indicator namely shopping centre type was proposed.

Comparing the GPS data to the TMH17, It was found that the average trip length per shopping centre size is longer by approximately 4.8 km compared to the prescribed TMH17 average trip lengths. These results need to be confirmed with further research. Even though further research is required to confirm the average trip lengths results, three options were evaluated to possibly revise the recommended TMH17 average trip length values. The method of estimating average trip length based on shopping centre type was deemed to be the preferred method.

If the proposed method is implemented and the average trip length values are confirmed, the impact on C_{AQD} contribution for Convenience and Neighbourhood centres will be significant. The C_{AQD} contribution would have been zero, but with the proposed method there will be an amount for the C_{AQD} contribution. For the Community, Small Regional, Regional and Super Regional centres on average the proposed adjusted halve trip length is 2.86 times longer than the TMH17's adjusted halve trip lengths. Since the $L_D/2$ is directly proportional to the C_{AQD}, the C_{AQD} contribution amount will increase on average by 286%. If it is assumed that the delivery vehicles will have the same adjusted halve trip lengths as the light vehicles, the strength component contribution C_{AHD} will increase similarly to the capacity component contribution C_{AQD}. The increase in the adjusted halve trip lengths will significantly increase the road contributions for shopping centres.

6.2 Recommendation

The following recommendations are made:

- A similar study is required to estimate the average trip lengths for the shopping centre delivery vehicles. This variable is required in Equation 2-7 to calculate the strength component for the road bulk contributions;
- The V Model for designing GPS data collection should be used for future GPS studies;
- Further research is required to improve trip end and trip purpose identification algorithms;
- The difference results between the point to point method compared to the link to link method in travel distance should be researched;
- During the GPS survey, the GPS loggers should only start recording after the loggers were given to the participant and should stop recording immediately after the loggers were retrieved. This will ease the data processing and cleaning processes;

- The tables and graphs in this document should be used to determine the required survey duration and participants required to collect a sample that can yield a 95% confidence level for shopping centre trips;
- It should be considered to conduct further research to confirm the average trip length results, as this study showed the TMH17 might underestimate the average shopping centre trip lengths significantly;
- Trip lengths for other land uses should be investigated with GPS data as well, and
- The high variability in shopping centre trip lengths across individuals and weak correlation by income group suggests further work on shopping centre trip behaviour is required.

7. References

Aultman-Hall, L & Du, J 2007, 'Increasing the accuracy of trip rate information from passive multi-day GPS travel datasets: Automatic trip end identification', *Transportation Research Part A*, 41 (2007) 220-232.

Axhausen, KW, Schonfelder, S, Wolf, J, Oliveira, M & Samaga, U 2003,'80 weeks of GPS traces: Approaches to enrich the trip information', *Arbeitsbericht Verkehrs- und Raumplanung, 178, Institut für Verkehrsplanung und Transportsysteme*, ETH Zürich, Zürich.

Biljecki, F 2010, *Automatic segmentation and classification of movement trajectories for transport modes* (Dissertation, MSc in Geomatrics). Delft University of Technology, TU Delft.

Banks, HB 2002, *Introduction to Transportation Engineering*, 2nd edn, McGraw-Hill Companies Inc, United States of America, New York.

Bhat, C, Brick, S & Srinivasan, S 2009, 'Methodology for converting GPS navigational streams to the travel diary data format'.

Borgers, A & Voster, C 2011, 'Assessing preference for mega shopping centres: A conjoint measurement approach', *Journal of Retailing and Consumer Services* 18 (2011) 322-332.

Bricka, SG, Sen, S, Paleti, R & Bhat, CR 2012, 'An analysis of the factors influencing differences in surveys-reported and GPS-recorded trip', *Transportation Research Part C* 21 (2012) 67-88.

City of Johannesburg 2010, *Consolidated Johannesburg Town Planning Scheme*, Draft.

Clifford, E, Zhang, J & Stopher, P 2008, 'Determining trip information using GPS data', *Institute of transport and logistics studies*, The University of Sydney, ITS-WP-08-01.

Cloete, CE 2002, *Management of commercial and industrial properties*, South African Property Education Trust, South Africa, Sandton.

Chen, C, Gong, H, Lawson, C & Bailostozky, E 2010, 'Evaluating the feasibility of a passive travel survey collection in a complex urban environment: Lessons learned from the New York City case study', *Transportation Research Part A* 44 (2010) 830-840.

Committee of Transport Officials (COTO) 2012, *South Africa Engineering Contributions Manual for Municipal Road Infrastructure*, Version 0.00, Draft.

Department of Transport (DOT) 2005a August, *The First South African National Household Travel Survey 2003 – Technical report*, Final Report.

Department of Transport (DOT) 2005b August, *The First South African National Household Travel Survey 2003 – Key results of the National Household Travel Survey*, Final Report.

Department of Transport (DOT) 1995 June, *South African Trip Generation Rates*, 2nd edn.

Johnson, RA 2005, *Miller & Freud's Probability and Statistics for Engineering*, 7th edn, Pearson Education Ltd, England, London.

Kobus, MBW, Gutierrez-i-Puigarnau, E, Rietveld, P & Van Ommeren, JN 2013, 'The on-street parking premium and car drivers' choice between street and garage parking', *Regional Science and Urban Economics* 43 (2013) 395-403.

Liu, J, Zheng, H, Feng, T, Yuan, S & Lu, H 2013, 'Post-processing procedures for passive GPS based travel survey', *Procedia - Social and Behavioural Sciences* 96 (2013) 310 – 319.

Murakami, E & Wagner, DP 1999, 'Can using global positioning systems (GPS) improve trip reporting?', *Transportation Research Part C* 7 (1999) 149-165.

Murat, CH 2010, 'Sample size needed for calibrating trip distribution and behaviour of the gravity model', *Journal of Transport Geography* 18 (2010) 183–190.

National Cooperative Highway Research Program (NCHRP) 2014a, *Applying GPS Data to Understand Travel Behaviour*, Volume 1: Background, Methods and Testing, Report 775.

National Cooperative Highway Research Program (NCHRP) 2014b, *Applying GPS Data to Understand Travel Behaviour*, Volume 2: Guidelines, Report 775.

Pearson, DF, Stover, VG & Benson, JD 1974, 'A Procedure for Estimation of Trip Length Frequency Distributions', *Texas Transport Institute*, Report No.: TTI-2-10-74-17-1.

Prinsloo, DA 2010, *Classification and hierarchy of retail facilities in South Africa*, Urban Studies.

Rasouli, S & Timmermans, H 2014, *Mobile Technologies for Activity-Travel Data Collection and Analysis*. IGI Global. United States of America.

Reimers, V 2013, 'Convenience for the car born shopper: Are malls and shopping strips driving customers away?', *Transportation Research Part A* 49 (2013) 35-47.

Shen, L & Stopher, PR 2013, '*Should we change the rules for trip identification for GPS travel records*', Proceedings of the 36th Australasian Transport Research Forum ATRF 2013 – "Transport and the New World City", Brisbane Australia.

Shen, L & Stopher, PR 2014, 'Using SenseCam to pesue "gound truth" for global positioning systems travel surveys' , *Transportation Research Part C* 42 (2014) 76-81.

South African Council of Shopping Centres (SACSC) 2012, *South African Shopping Centre Directory*.

Statistics South Africa (Stats SA) 2012, *Census 2011 Municipal Report Gauteng*, Report No. 03-01-55.

Statistics South Africa (Stats SA) 2014, *National Household Travel Survey, February to March 2013*, Embargoed until 17 July 2014, Statistical Release P0320.

Stopher, PR, FitzGerald, C & Zhang, J 2008, 'Search for a global positioning system device to measure person travel', *Transportation Research Part C*, 16 (2008) 350-369.

Stopher, PR & Greaves, SP 2007, 'Household travel surveys: Where are we going?', *Transportation Research Part A* 41 (2007) 367-381.

Stover, VG & Koepke, FJ 2002, *Transportation and Land Development*, 2nd edn, Institute of Transport Engineers. United States of America.

Timmermans, H & Rasouli, S 2013, 'Assessment of model uncertainty in destinations and travel forecasts of models complex spatial shopping behaviour', *Journal of Retailing and Consumer Services*, 20 (2013) 139-146.

TMH17 Committee of Transport Officials (COTO) September 2013, *South African Trip Data Manual*, Version 1.01.

TTT Africa 2008 November, *Policy for Engineering Service Contributions for Roads and Stormwater,* Johannesburg Road Agency (JRA), Final Report.

Van As, SC 2008, *Applied Statistics for Civil Engineering*.

Venter, CJ & Joubert, JW 2013, 'Use of Multisource GPS Data to Characterized Multiday Driving Patterns and Fuel Usage in a Large Urban Region', *Journal of the Transportation Research Record*, No 2338, page 1-10.

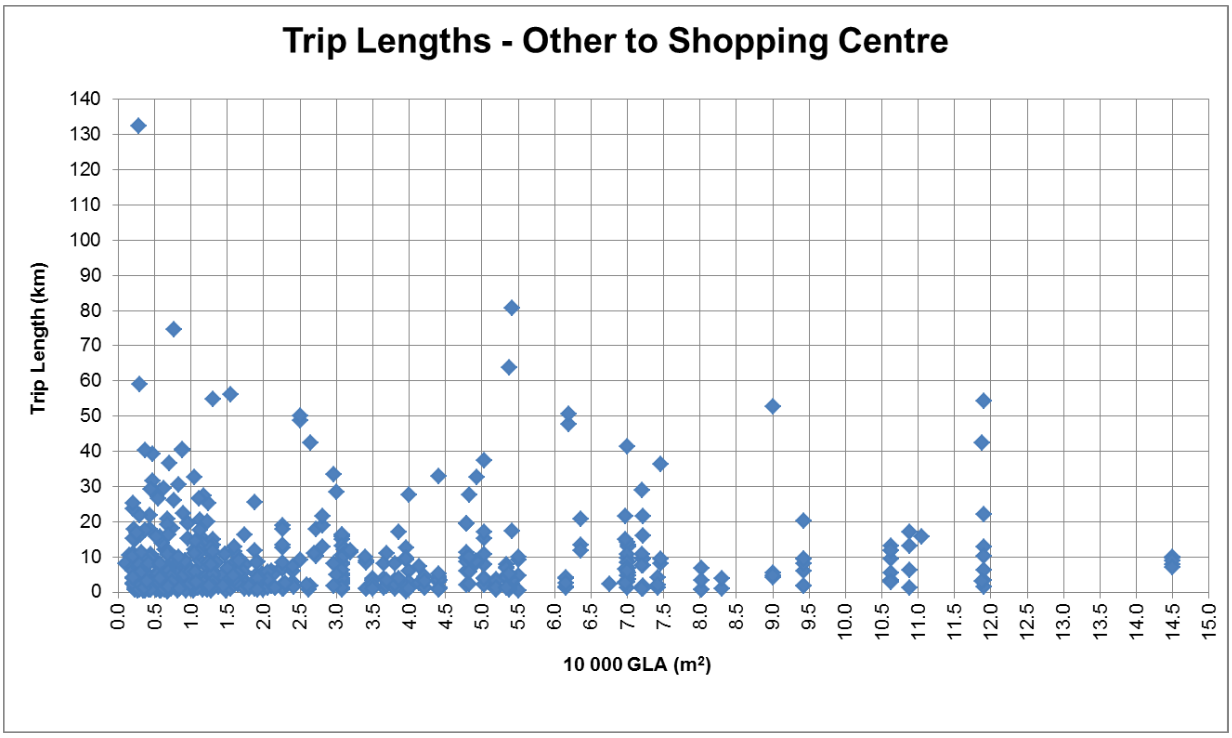
Venter, CJ & Joubert JW 2014, 'Tax or toll? GPS-based assessment of equity impacts of large scale electronic freeway tolling in Gauteng, South Africa', *Transportation Research Record: Journal of the Transportation Research Board*, No. 2450, D.C. 2014 pp 62-70.

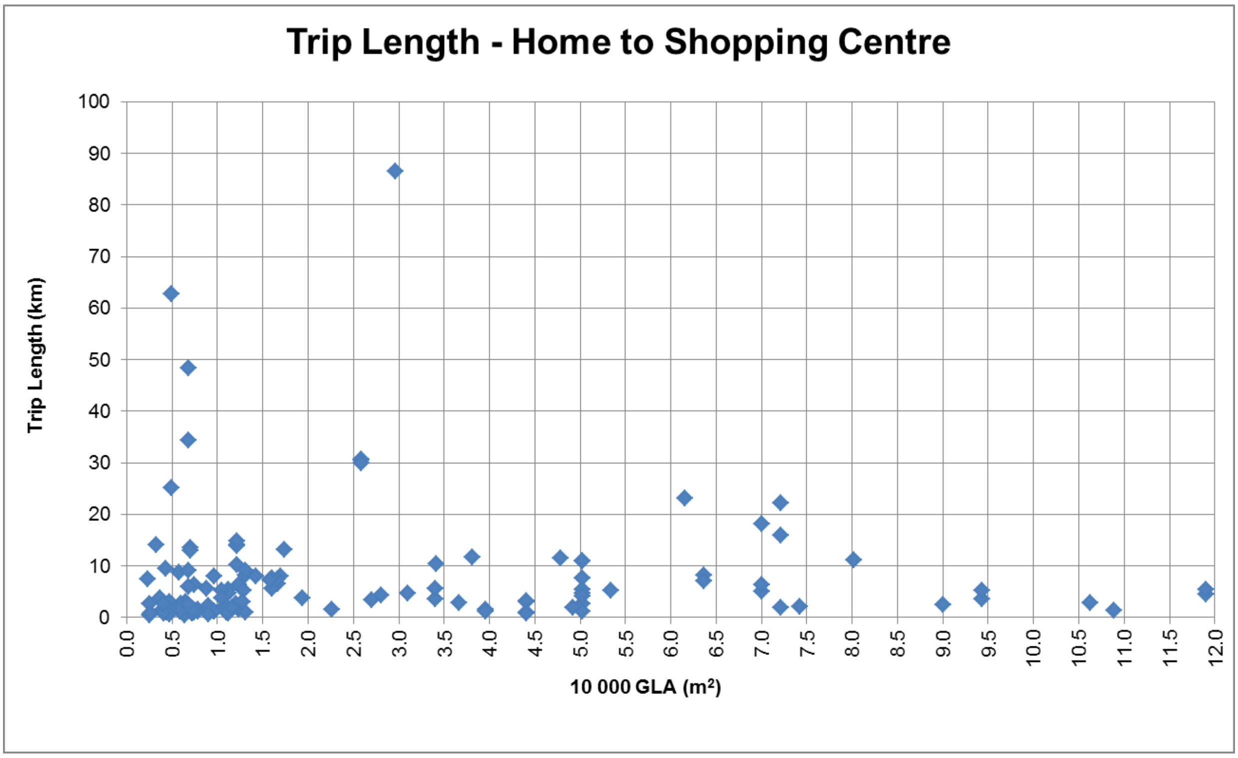
Woehrle, M & Brouwers, N 2013, 'Dwelling in the canyons: Dwelling detection in urban environments using GPS, Wi-Fi, and geolocation', *Pervasive and Mobile Computing 9*, (2013) 665-680.

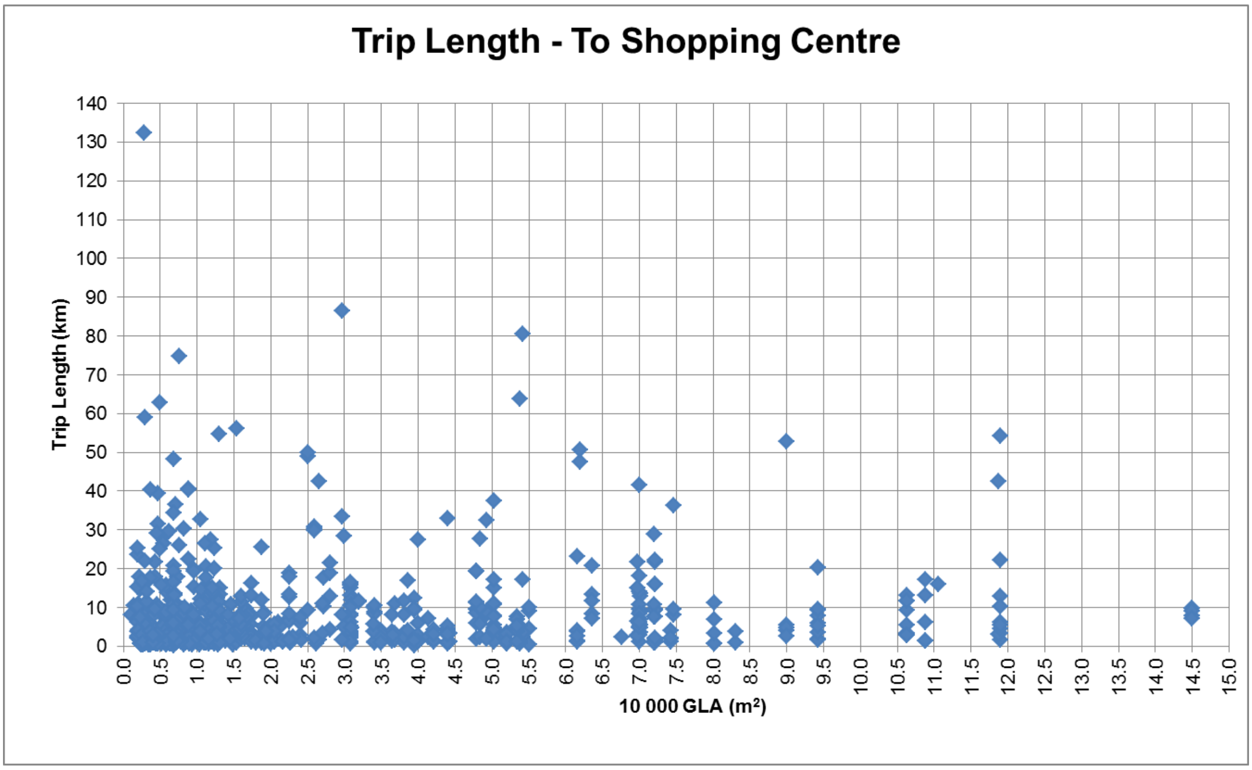
Wolf, J 2000, *Using GPS Data Loggers to Replace Travel Diaries in the Collection of Travel Data*, Dissertation, Georgia Institute of Technology, School of Civil and Environmental Engineering, Atlanta, Georgia, July 2000.

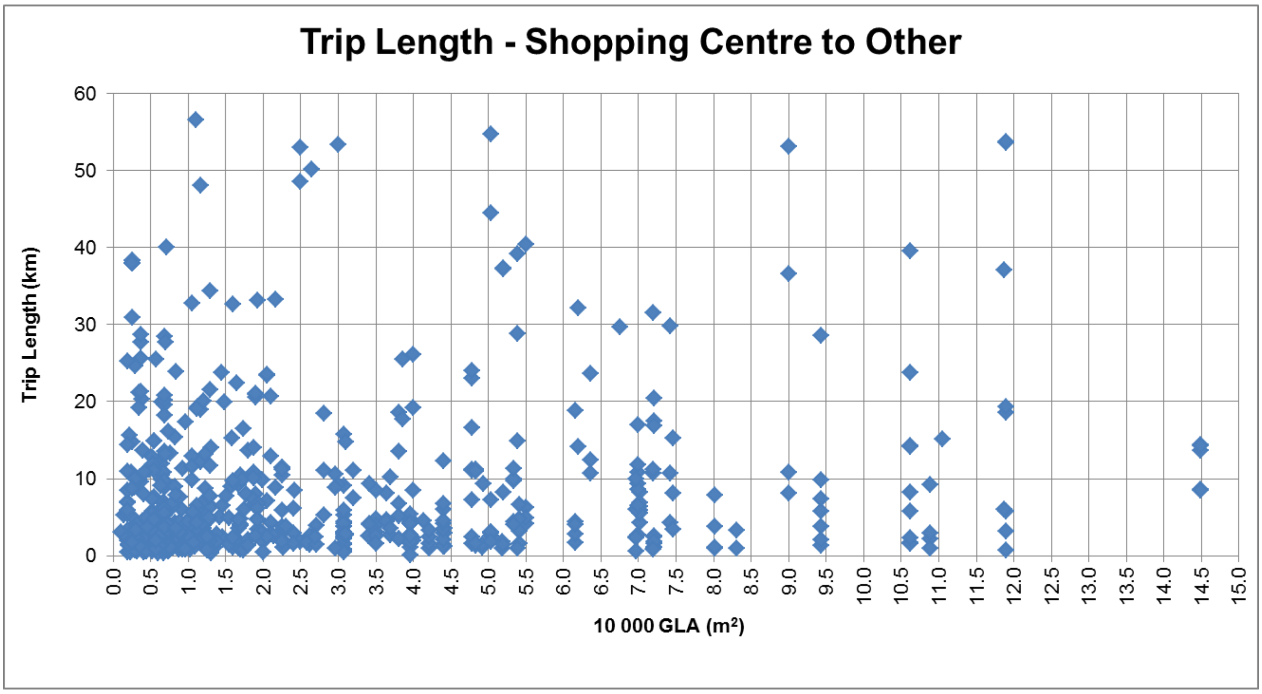
Appendix A

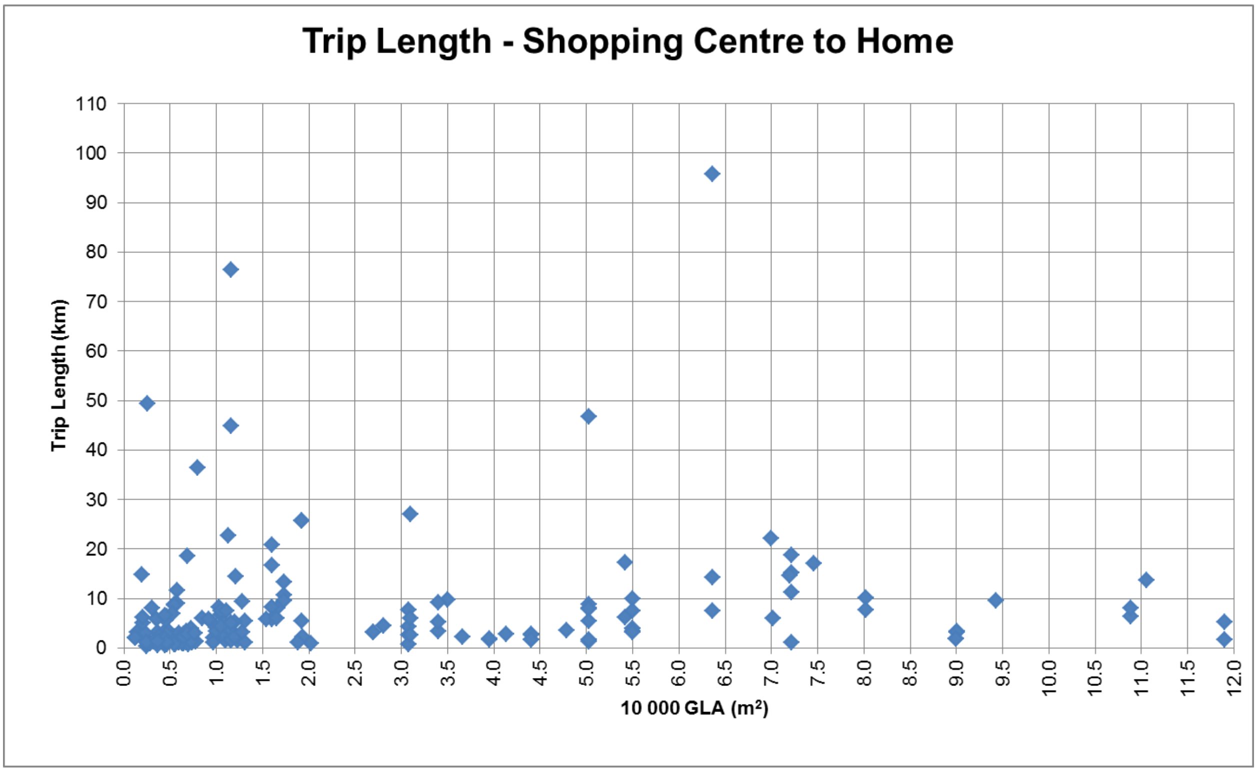
Trip Length Graphs

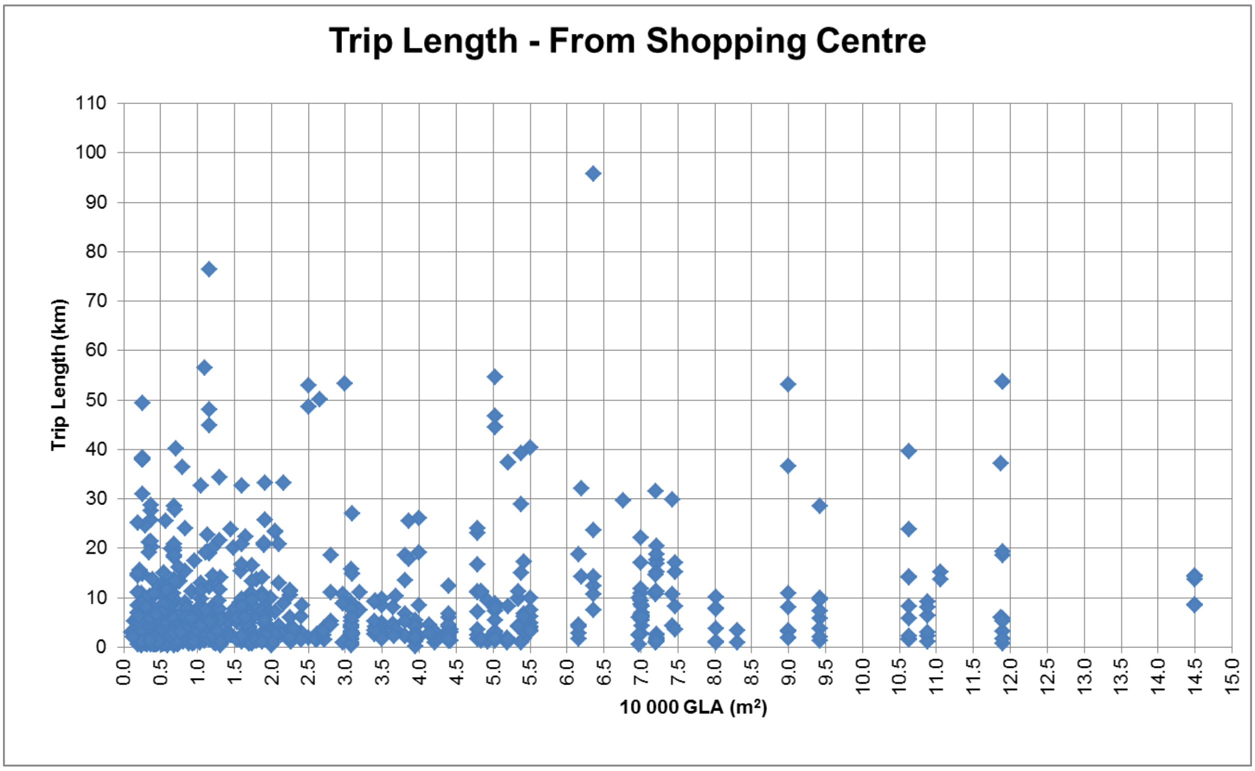

Trip Lengths - Other to Shopping Centre
Trip Length (km)
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0
12.5
13.0
13.5
14.0
14.5
15.0
10 000 GLA (m²)

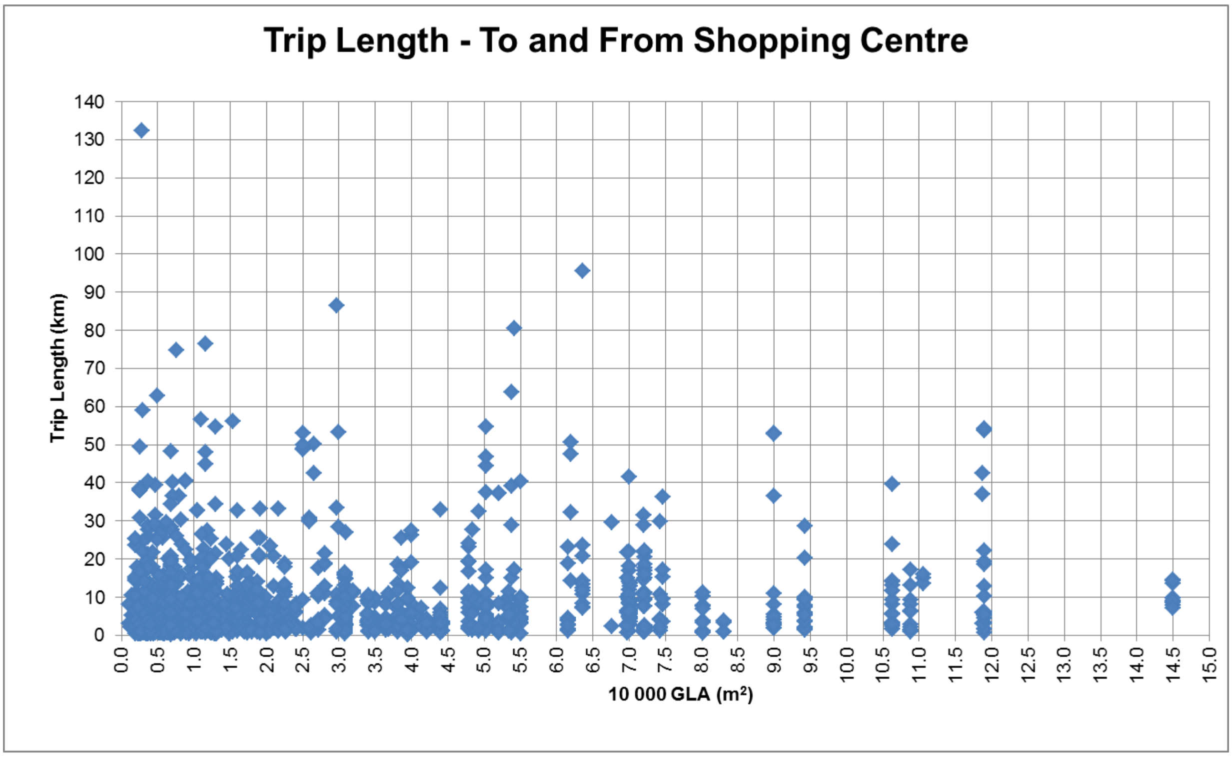

Trip Length - Home to Shopping Centre
Trip Length (km)
100
90
80
70
60
50
40
30
20
10
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0
10 000 GLA (m²)


Trip Length - To Shopping Centre
Trip Length (km)
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 10.0 10.5 11.0 11.5 12.0 12.5 13.0 13.5 14.0 14.5 15.0
10 000 GLA (m²)


Trip Length - Shopping Centre to Other
Trip Length (km)
60
50
40
30
20
10
0
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 10.0 10.5 11.0 11.5 12.0 12.5 13.0 13.5 14.0 14.5 15.0
10 000 GLA (m²)


Trip Length - Shopping Centre to Home
Trip Length (km)
10 000 GLA (m²)

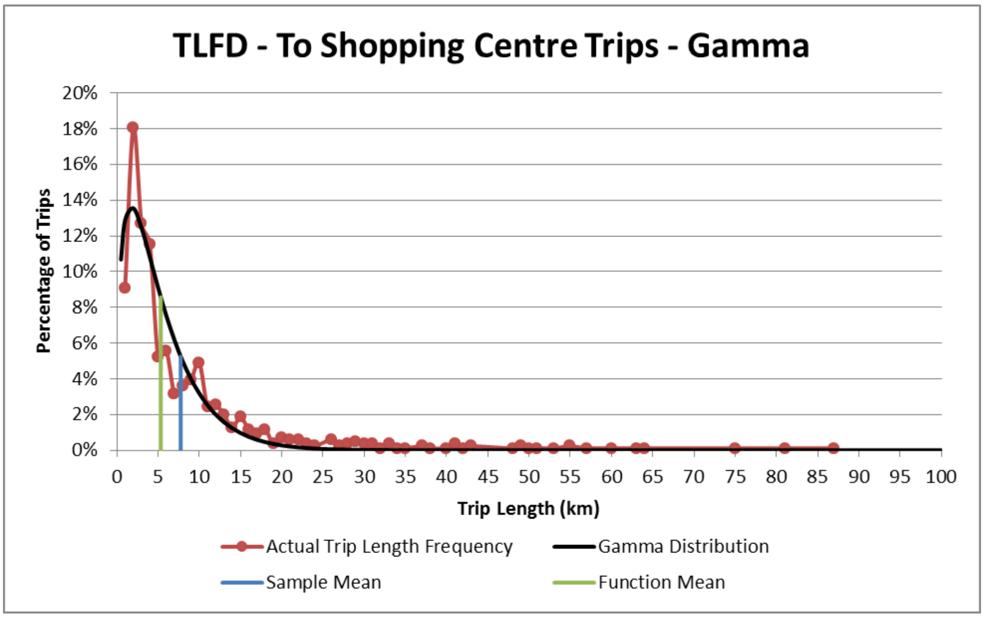

Trip Length - From Shopping Centre
Trip Length (km)
10 000 GLA (m²)

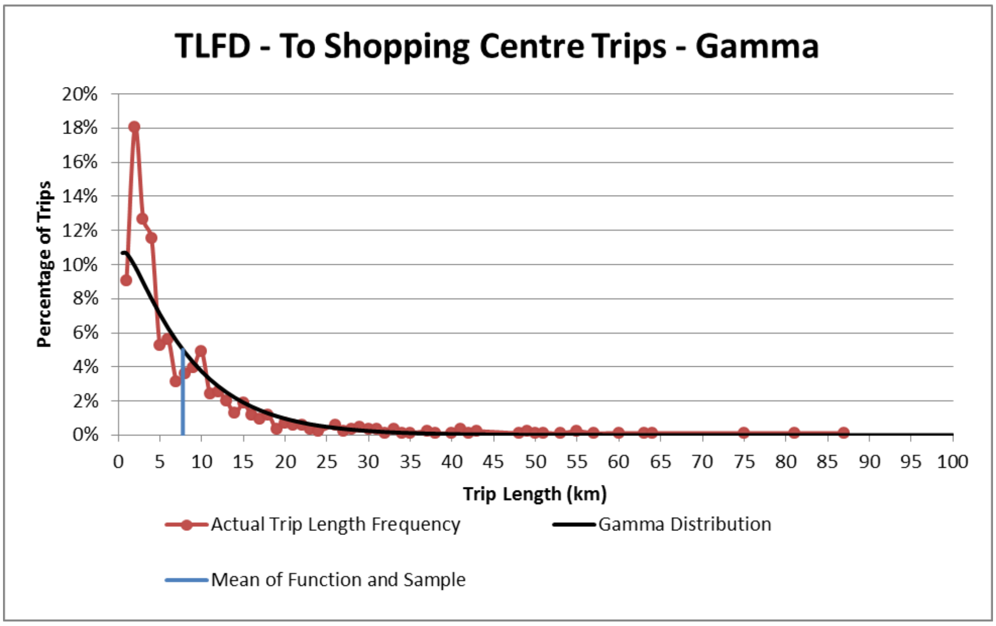

Trip Length - To and From Shopping Centre
Trip Length (km)
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0
12.5
13.0
13.5
14.0
14.5
15.0
10 000 GLA (m²)

Appendix B

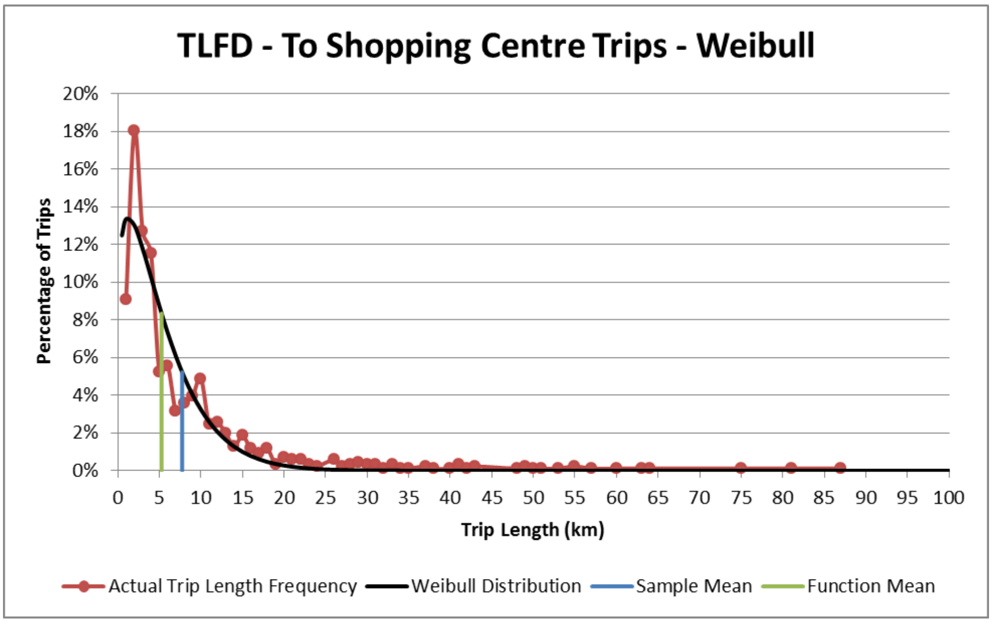
Shopping Centre Trip Length Distributions

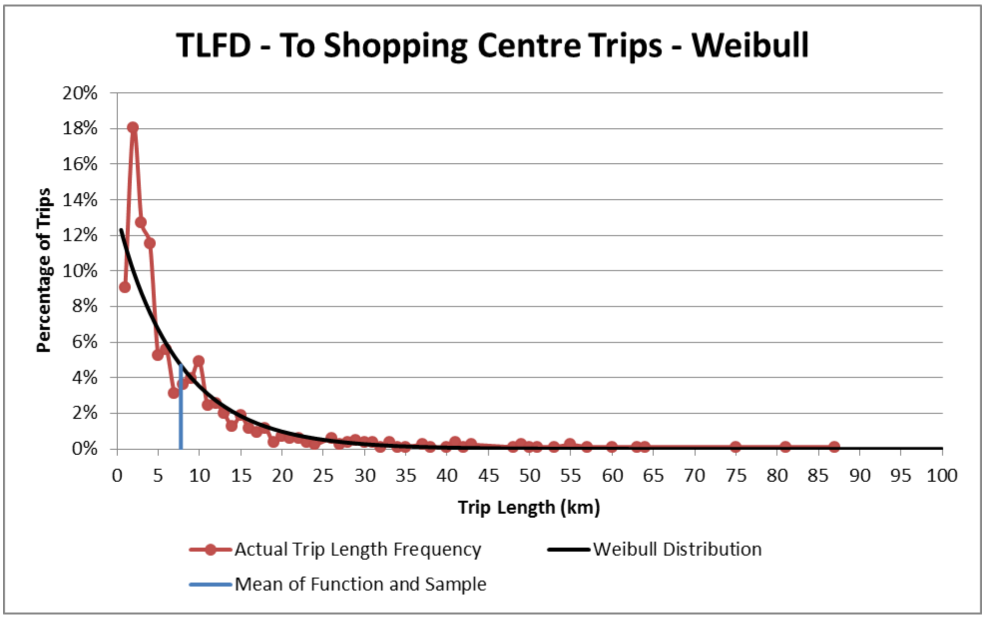
TLFD – To Shopping Centre Trips – Gamma Distribution		
Variable	**Function Mean Not Equal to Sample Mean**	**Function Mean Equal to Sample Mean**
Alpha	1.47	1.099
Beta	3.63	7.048
R^2	0.891	0.832
RMS Error	0.012	0.015
Error Squared	0.007	0.011
Mean Diff (km)	-2.4	0.0



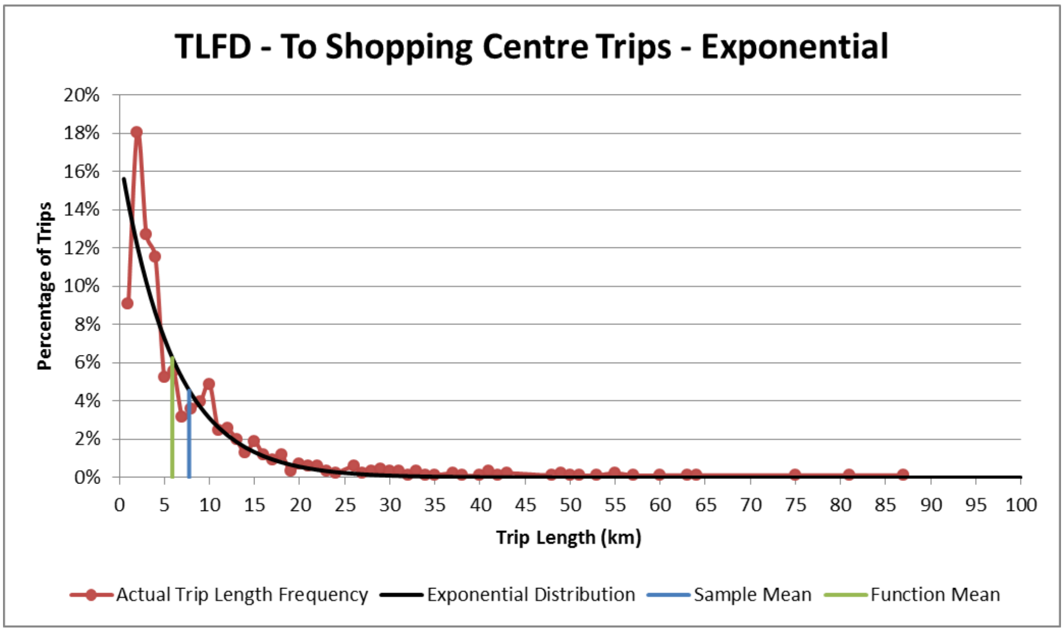


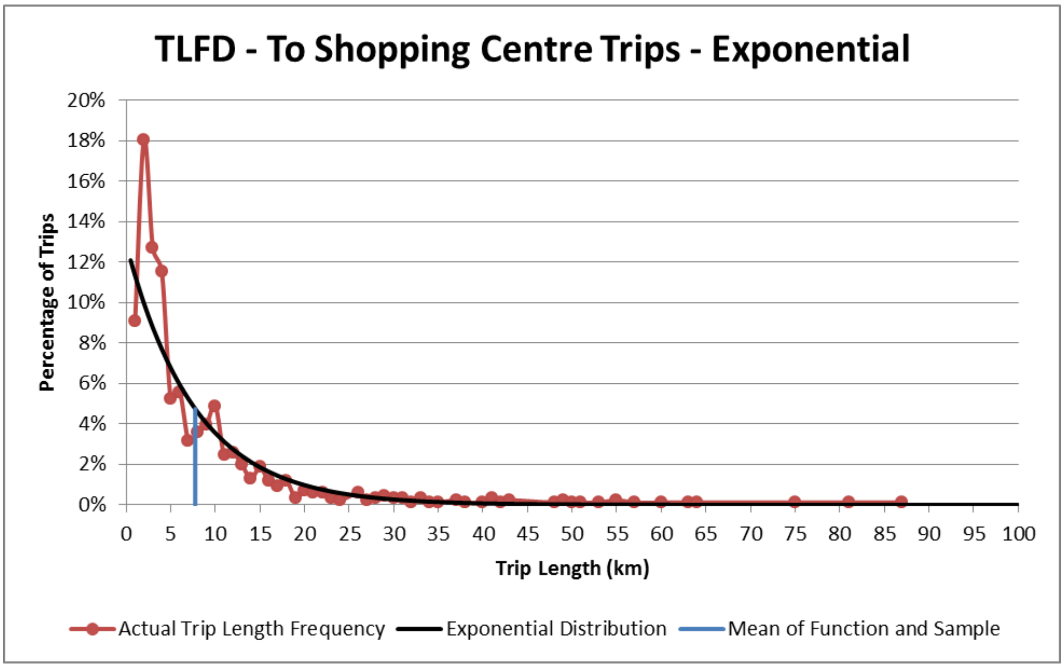
TLFD – To Shopping Centre Trips – Weibull Distribution		
Variable	**Function Mean Not Equal to Sample Mean**	**Function Mean Equal to Sample Mean**
Alpha	1.20	0.99
Beta	5.61	7.73
R^2	0.908	0.867
RMS Error	0.012	0.015
Error Squared	0.008	0.011
Mean Diff (km)	-2.5	0.0



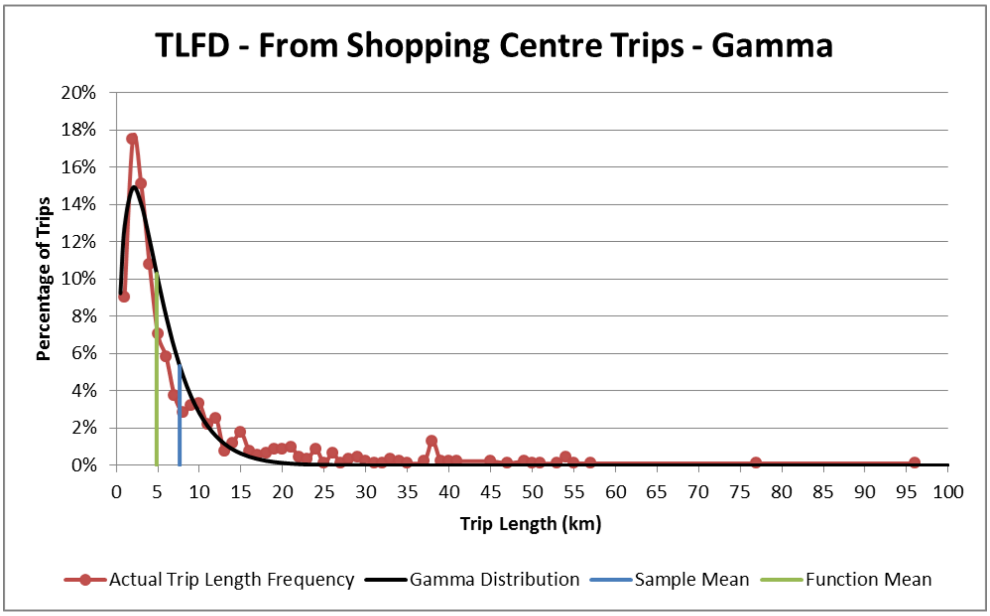


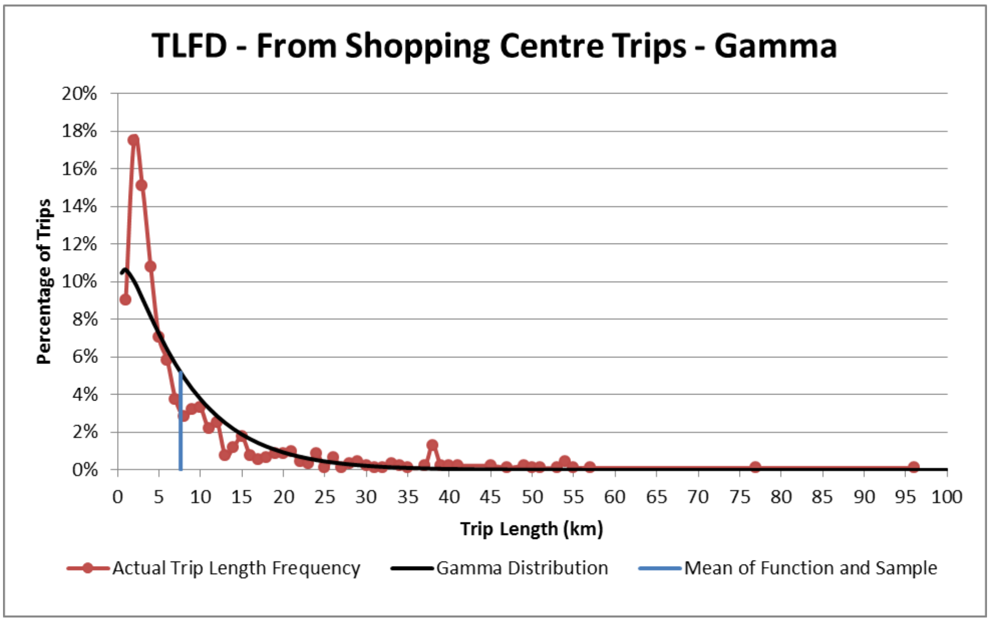
TLFD – To Shopping Centre Trips – Exponential Distribution		
Variable	**Function Mean Not Equal to Sample Mean**	**Function Mean Equal to Sample Mean**
Lambda	0.170	0.129
R^2	0.861	0.829
RMS Error	0.013	0.015
Error Squared	0.009	0.011
Mean Diff (km)	-1.9	0.0



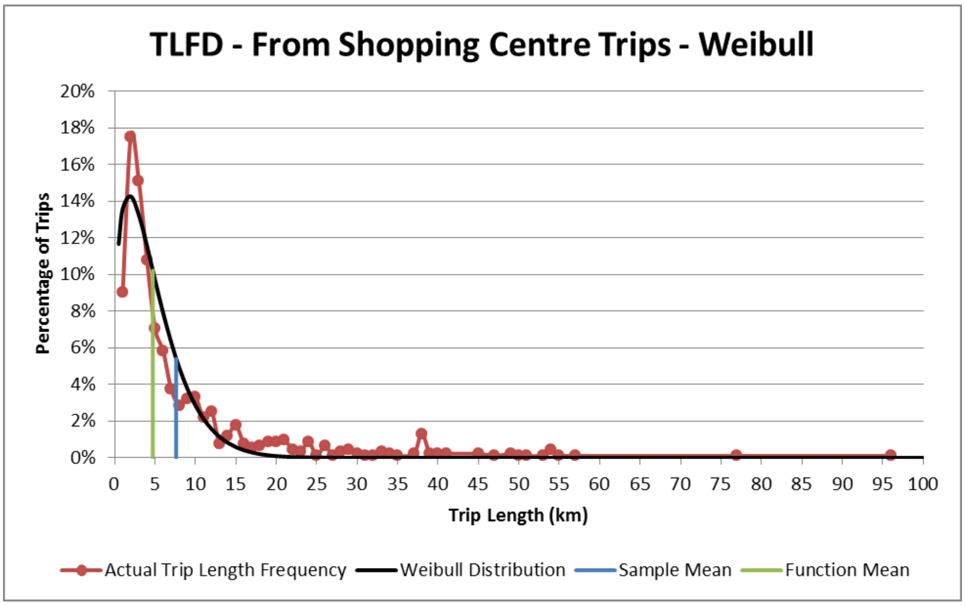


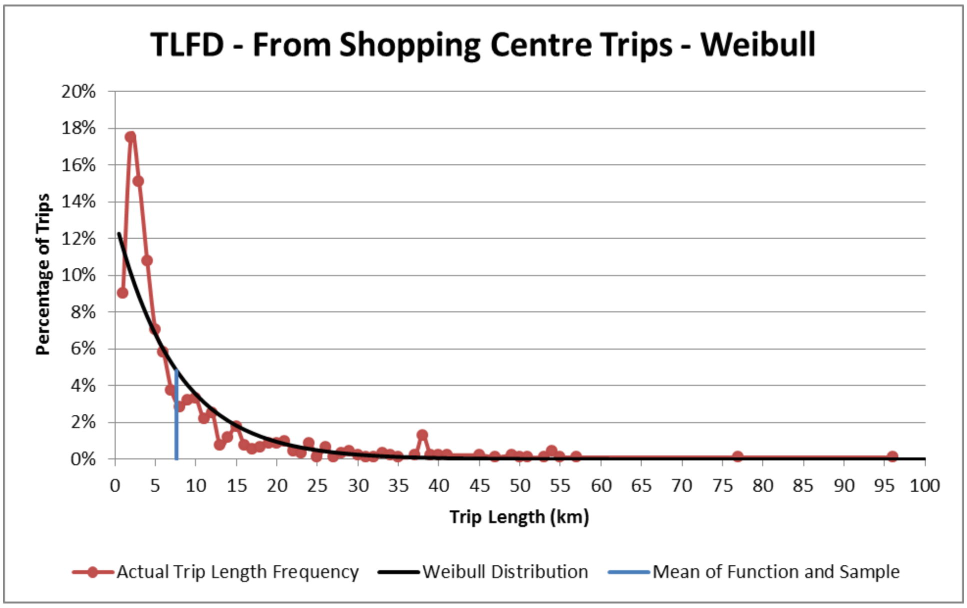
TLFD – From Shopping Centre Trips – Gamma Distribution		
Variable	**Function Mean Not Equal to Sample Mean**	**Function Mean Equal to Sample Mean**
Alpha	1.73	1.13
Beta	2.81	6.73
R^2	0.917	0.825
RMS Error	0.011	0.015
Error Squared	0.006	0.012
Mean Diff (km)	-2.8	0.0



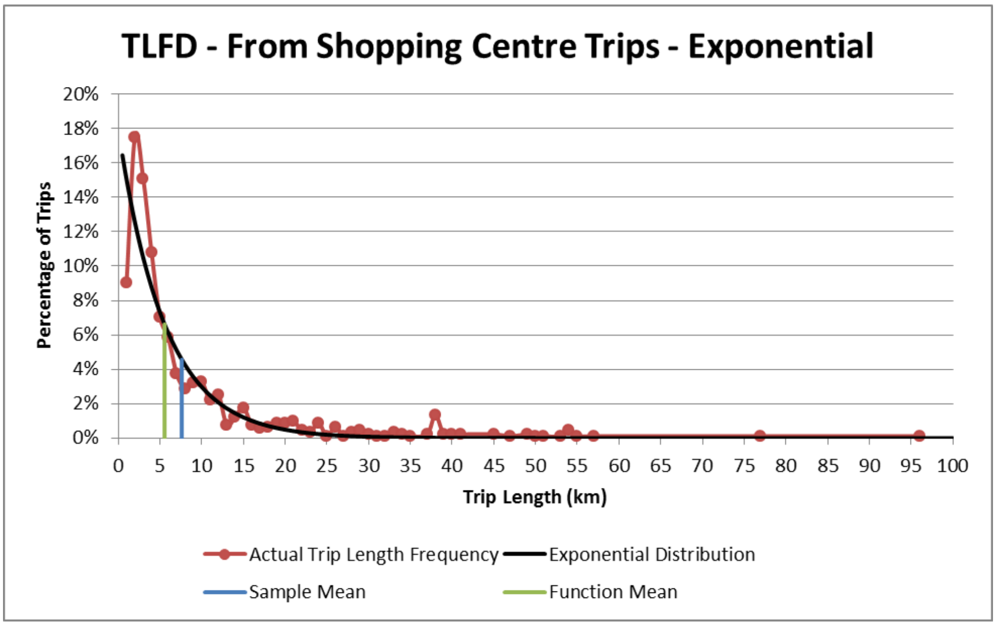


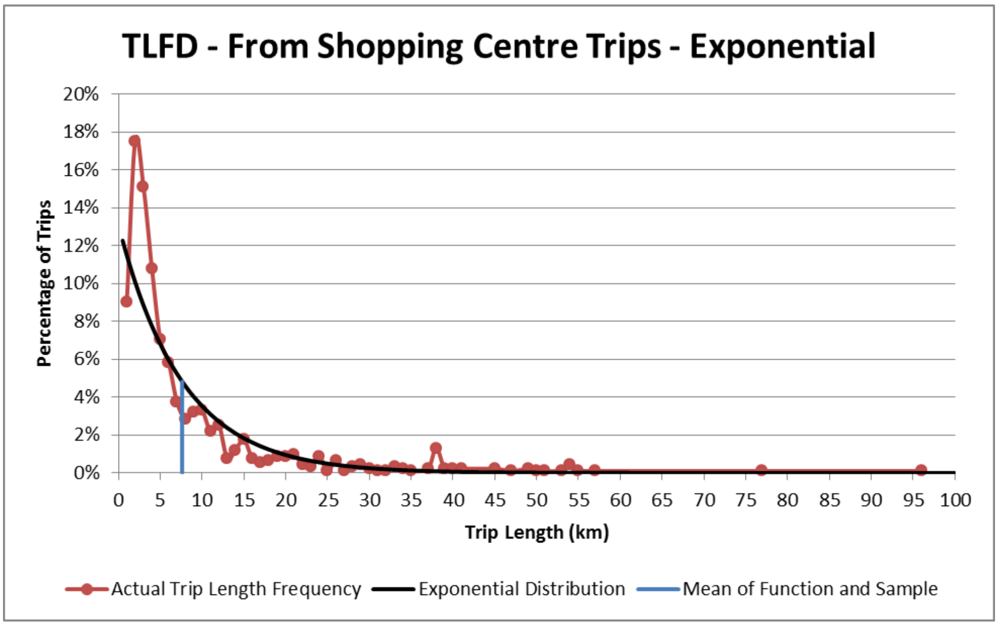
TLFD – From Shopping Centre Trips – Weibull Distribution		
Variable	**Function Mean Not Equal to Sample Mean**	**Function Mean Equal to Sample Mean**
Alpha	1.32	1.00
Beta	5.13	7.63
R^2	0.924	0.860
RMS Error	0.011	0.016
Error Squared	0.007	0.012
Mean Diff (km)	-2.9	0.0



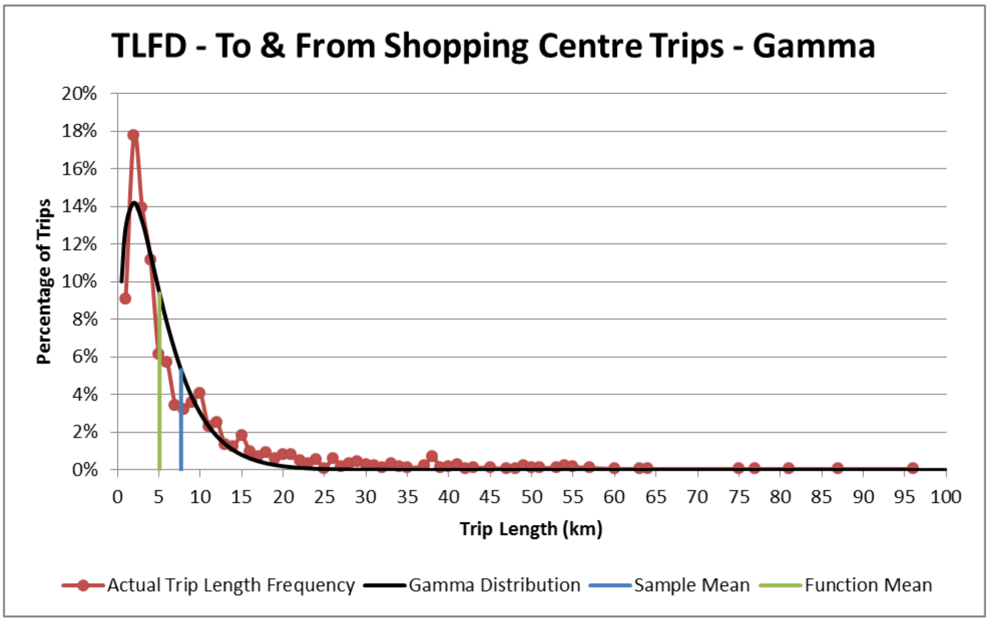


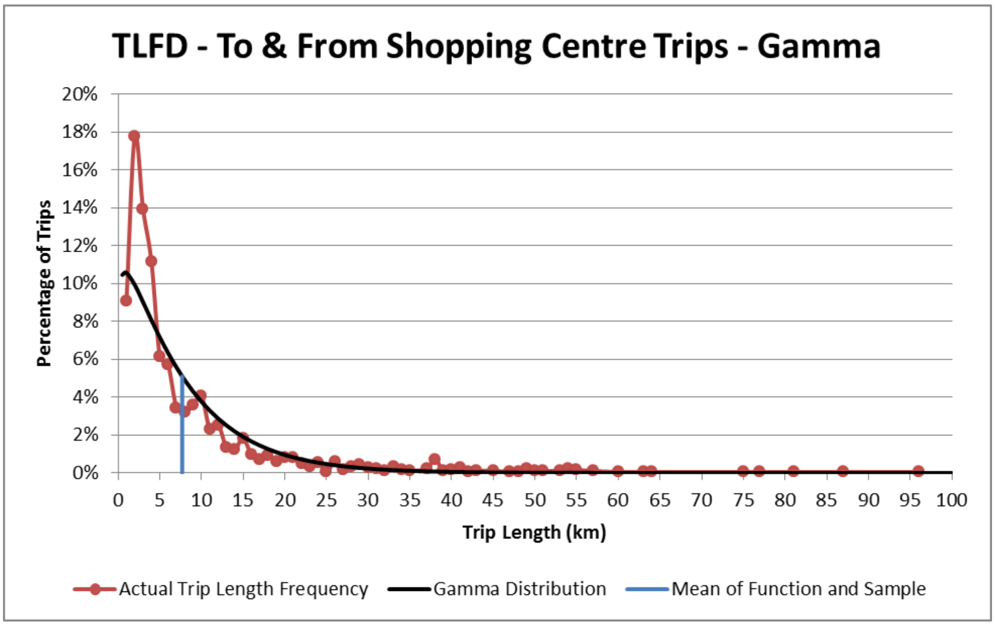
TLFD – To Shopping Centre Trips – Exponential Distribution		
Variable	**Function Mean Not Equal to Sample Mean**	**Function Mean Equal to Sample Mean**
Lambda	0.180	0.131
R^2	0.861	0.820
RMS Error	0.014	0.016
Error Squared	0.010	0.012
Mean Diff (km)	-2.1	0.0



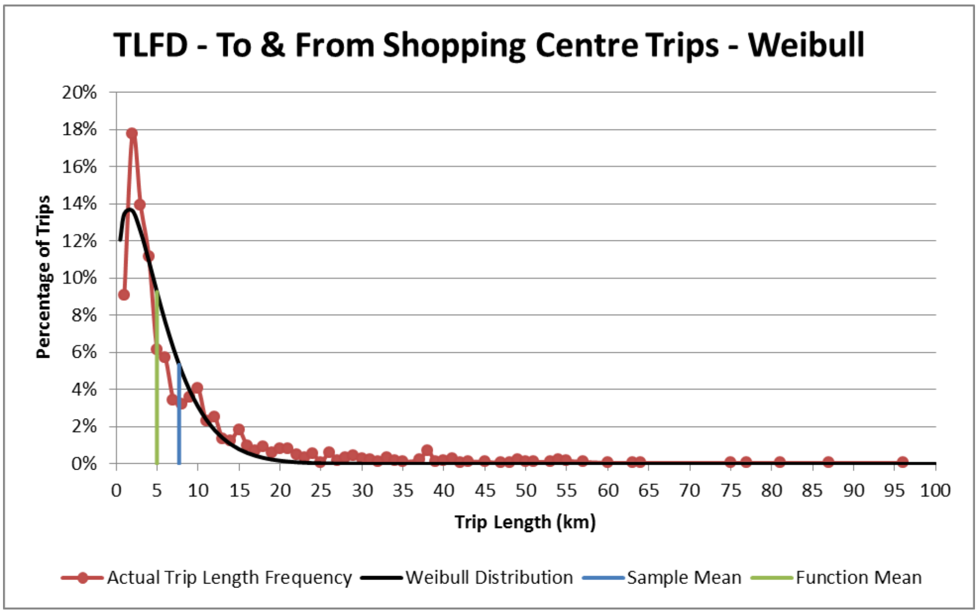


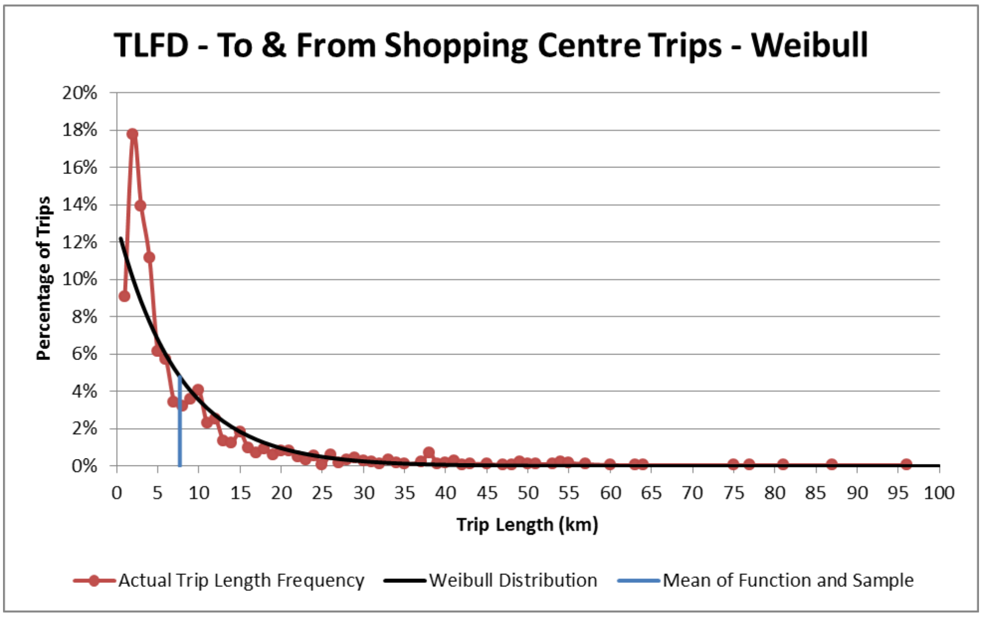
TLFD –To & From Shopping Centre Trips – Gamma Distribution		
Variable	**Function Mean Not Equal to Sample Mean**	**Function Mean Equal to Sample Mean**
Alpha	1.59	1.12
Beta	3.20	6.88
R^2	0.912	0.839
RMS Error	0.010	0.014
Error Squared	0.006	0.011
Mean Diff (km)	-2.6	0.0



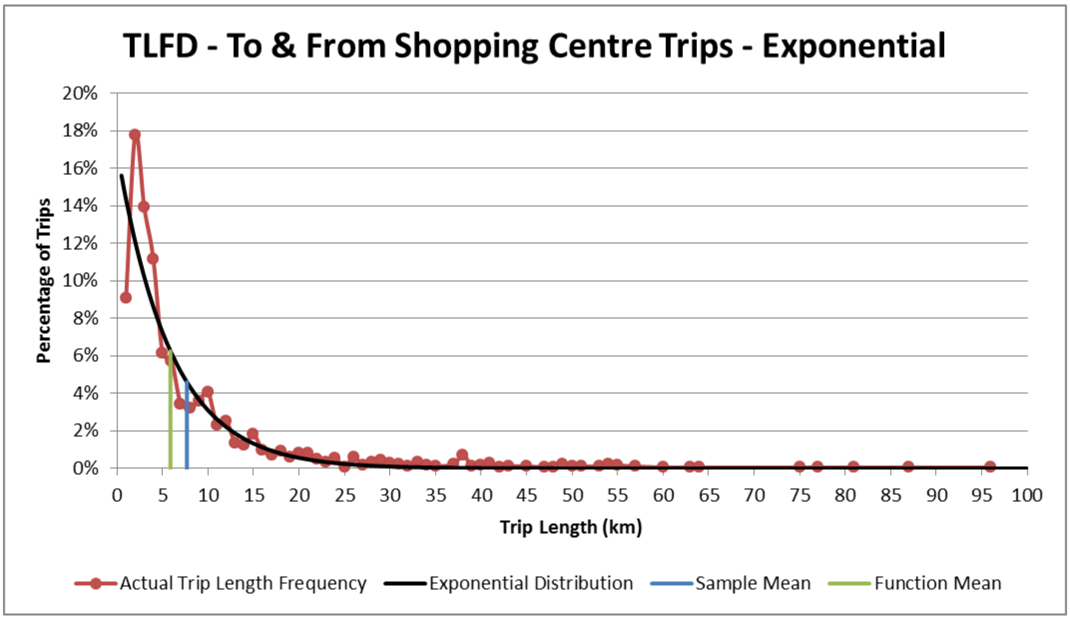


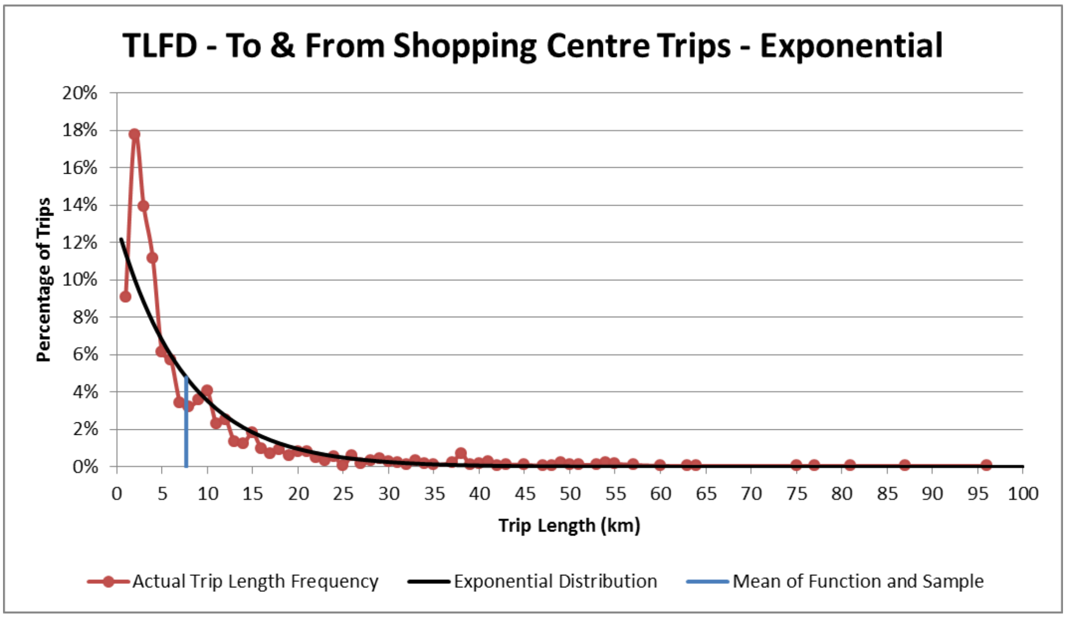
TLFD –To & From Shopping Centre Trips – Weibull Distribution		
Variable	**Function Mean Not Equal to Sample Mean**	**Function Mean Equal to Sample Mean**
Alpha	1.26	1.00
Beta	5.36	7.68
R^2	0.920	0.868
RMS Error	0.011	0.014
Error Squared	0.007	0.011
Mean Diff (km)	-2.7	0.0





TLFD –To & From Shopping Centre Trips – Exponential Distribution		
Variable	**Function Mean Not Equal to Sample Mean**	**Function Mean Equal to Sample Mean**
Lambda	0.170	0.130
R^2	0.872	0.836
RMS Error	0.012	0.014
Error Squared	0.009	0.011
Mean Diff (km)	-1.8	0.0

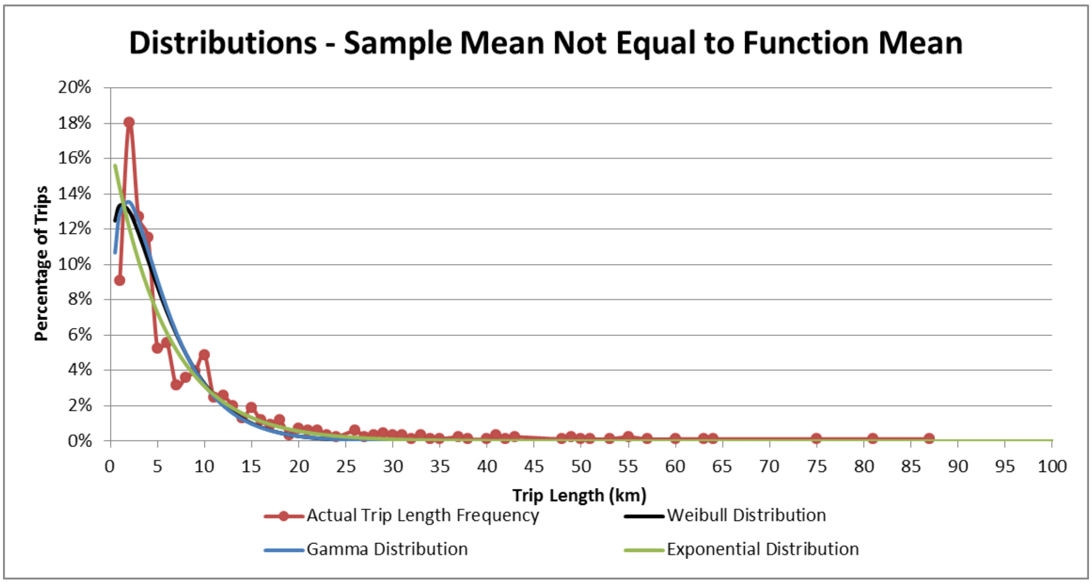


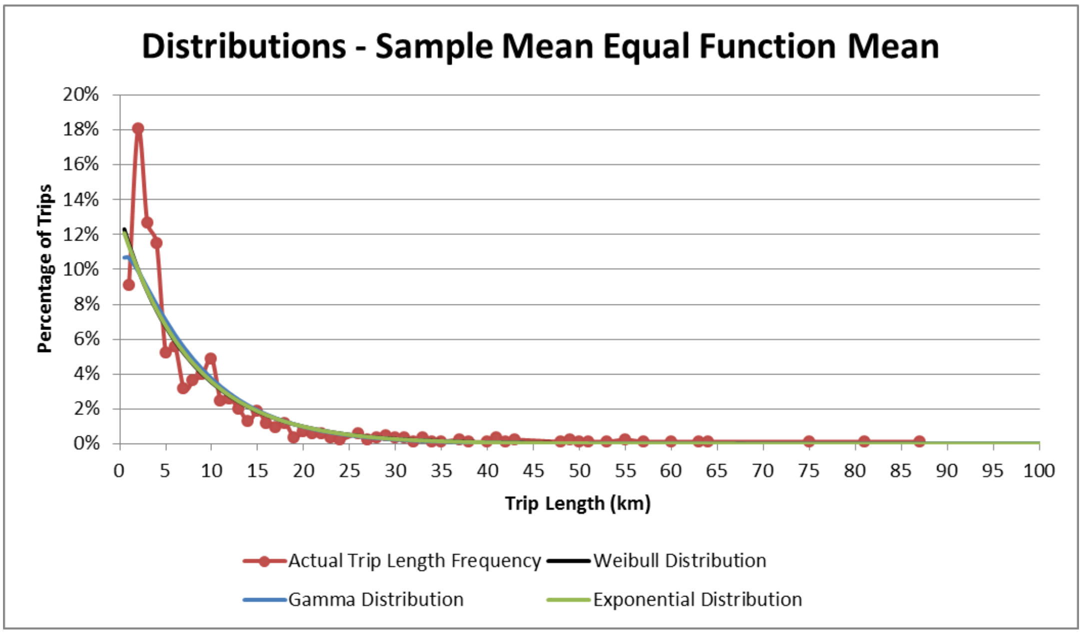


Appendix C

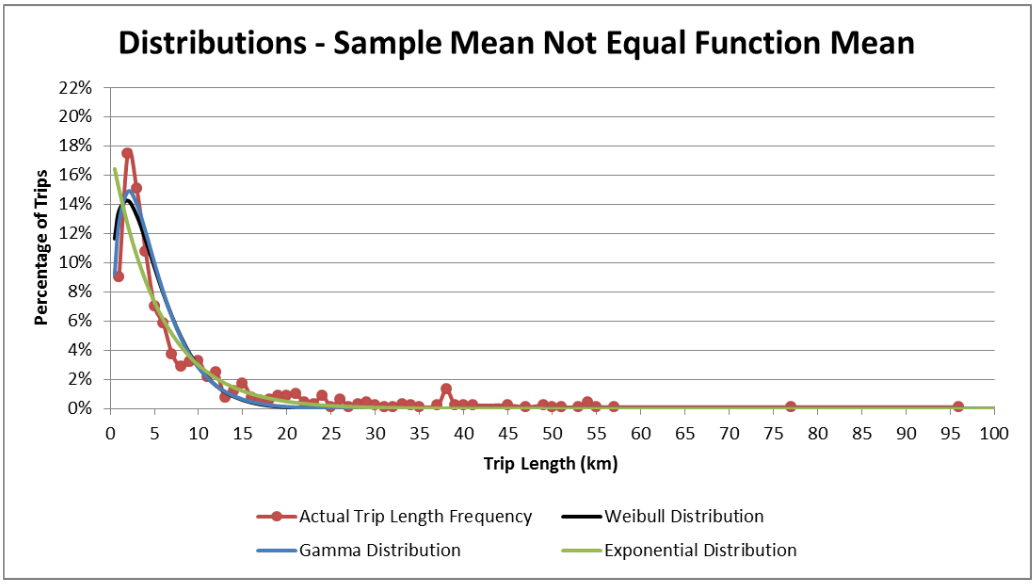
Comparisons of Trip Length Distributions

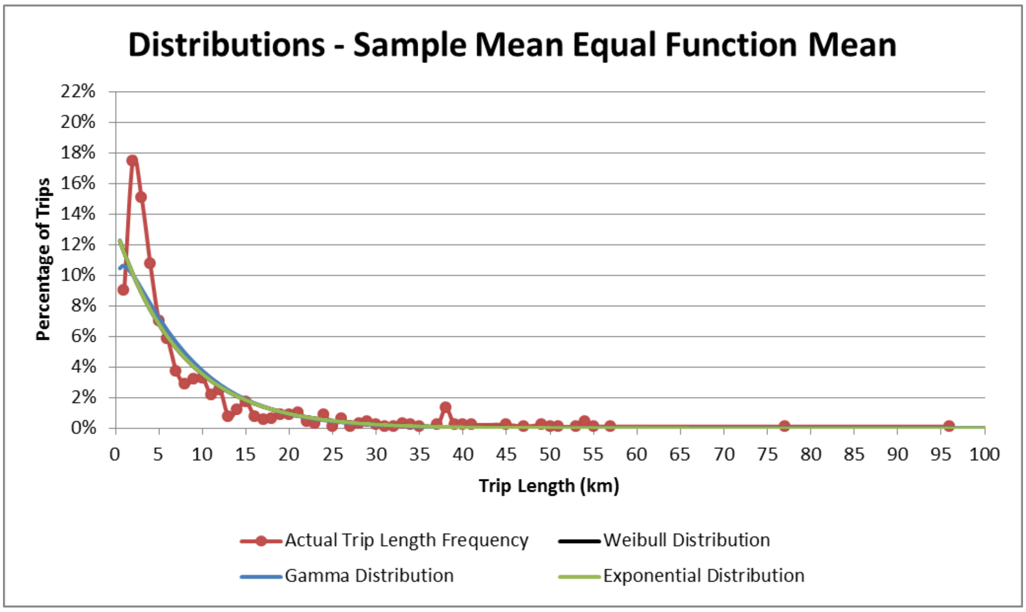
Trip Type: To Shop						
Variable	Function Mean Not Equal to Sample Mean			Function Mean Equal to Sample Mean		
	Gamma	Weibull	Exponential	Gamma	Weibull	Exponential
R^2	0.891	0.908	0.861	0.832	0.867	0.829
RMS Error	0.012	0.012	0.013	0.015	0.015	0.015
Error Squared	0.007	0.008	0.009	0.011	0.011	0.011
Mean Diff (km)	-2.4	-2.5	-1.9	0.0	0.0	0.0



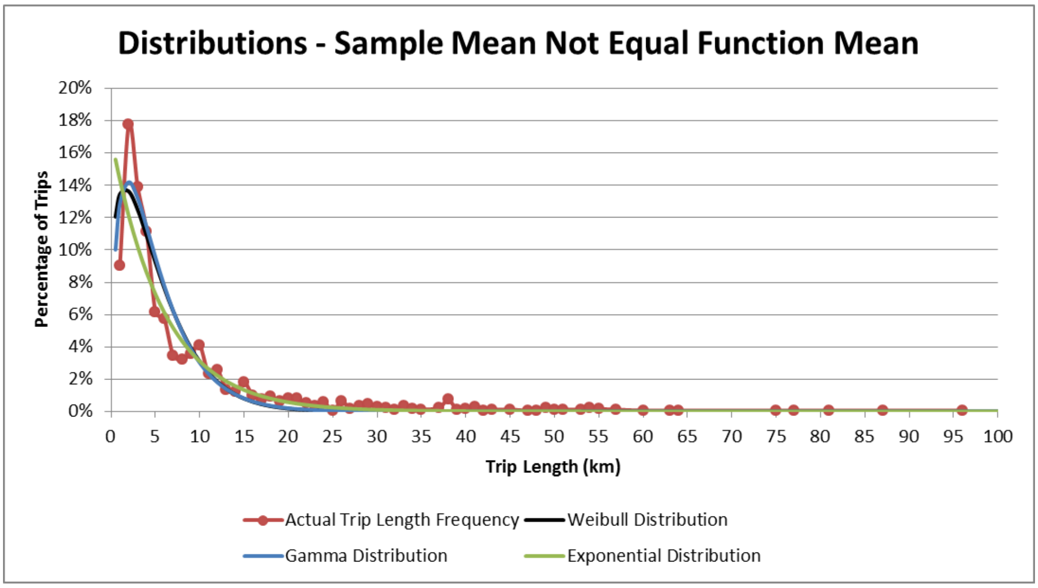


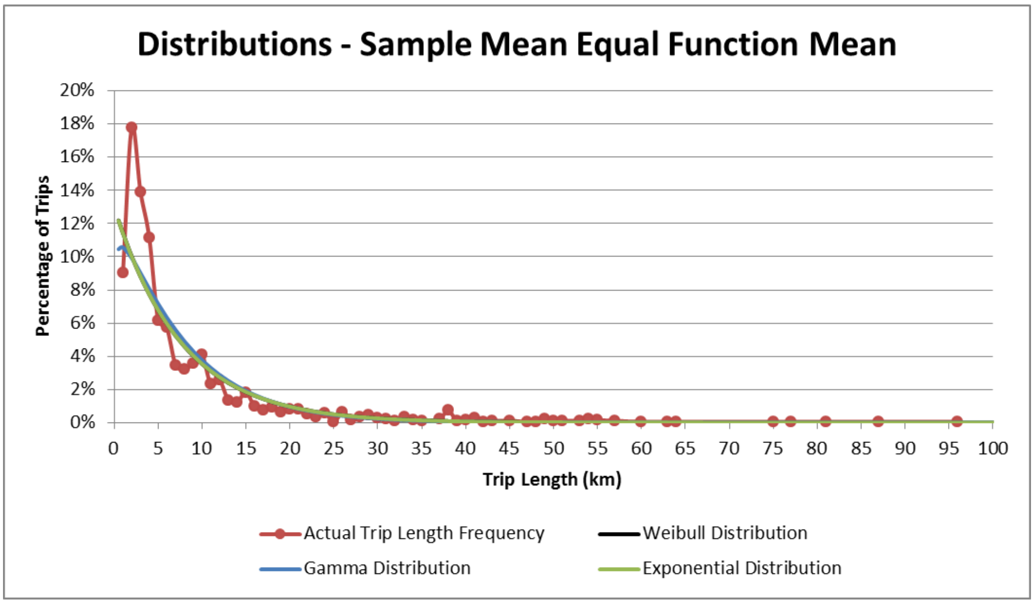
Trip Type: From Shop						
Variable	Function Mean Not Equal to Sample Mean			Function Mean Equal to Sample Mean		
	Gamma	Weibull	Exponential	Gamma	Weibull	Exponential
R^2	0.917	0.924	0.861	0.825	0.860	0.820
RMS Error	0.011	0.011	0.014	0.015	0.016	0.016
Error Squared	0.006	0.007	0.010	0.012	0.012	0.012
Mean Diff (km)	-2.8	-2.9	-2.1	0.0	0.0	0.0





Trip Type: To & From Shop						
Variable	Function Mean Not Equal to Sample Mean			Function Mean Equal to Sample Mean		
	Gamma	Weibull	Exponential	Gamma	Weibull	Exponential
R^2	0.912	0.920	0.872	0.839	0.868	0.836
RMS Error	0.010	0.011	0.012	0.014	0.014	0.014
Error Squared	0.006	0.007	0.009	0.011	0.011	0.011
Mean Diff (km)	-2.6	-2.7	-1.8	0.0	0.0	0.0

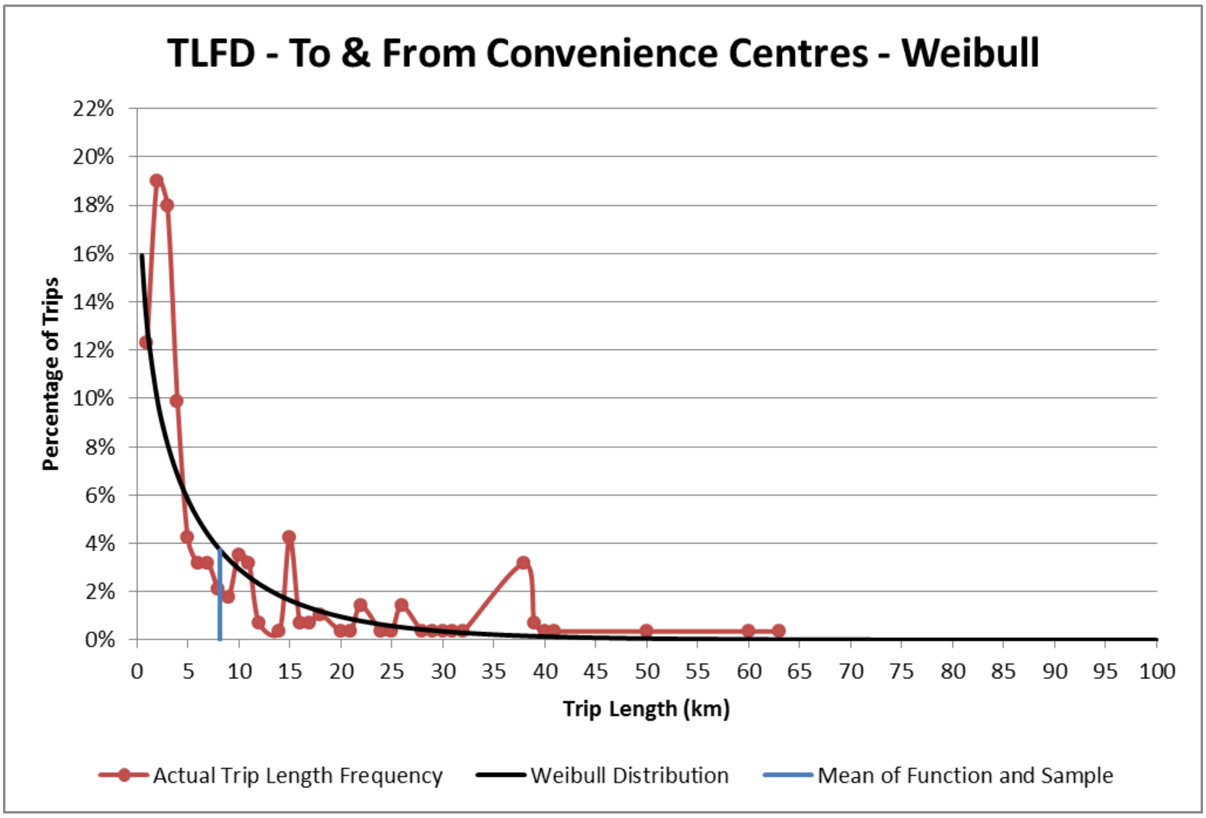




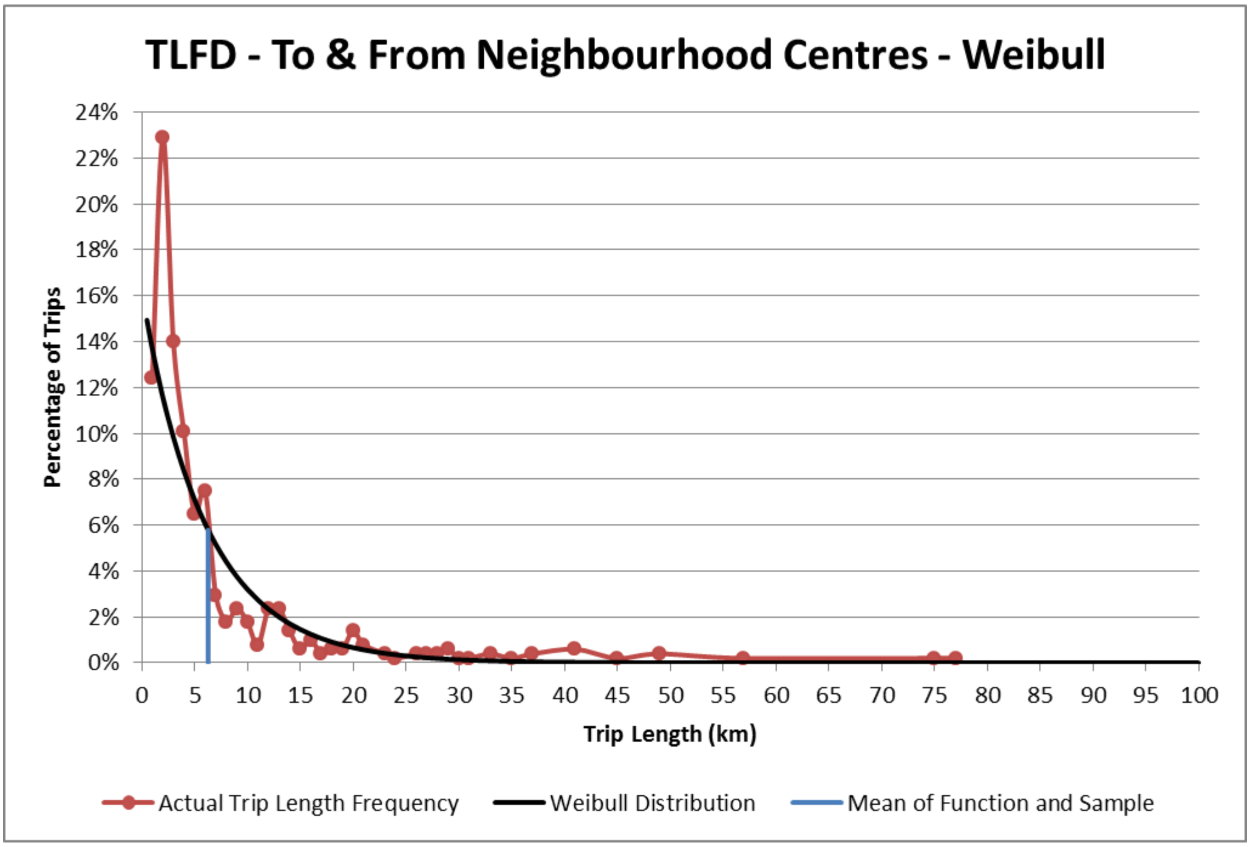
Appendix D

Weibull Distribution per Shopping Centre Type

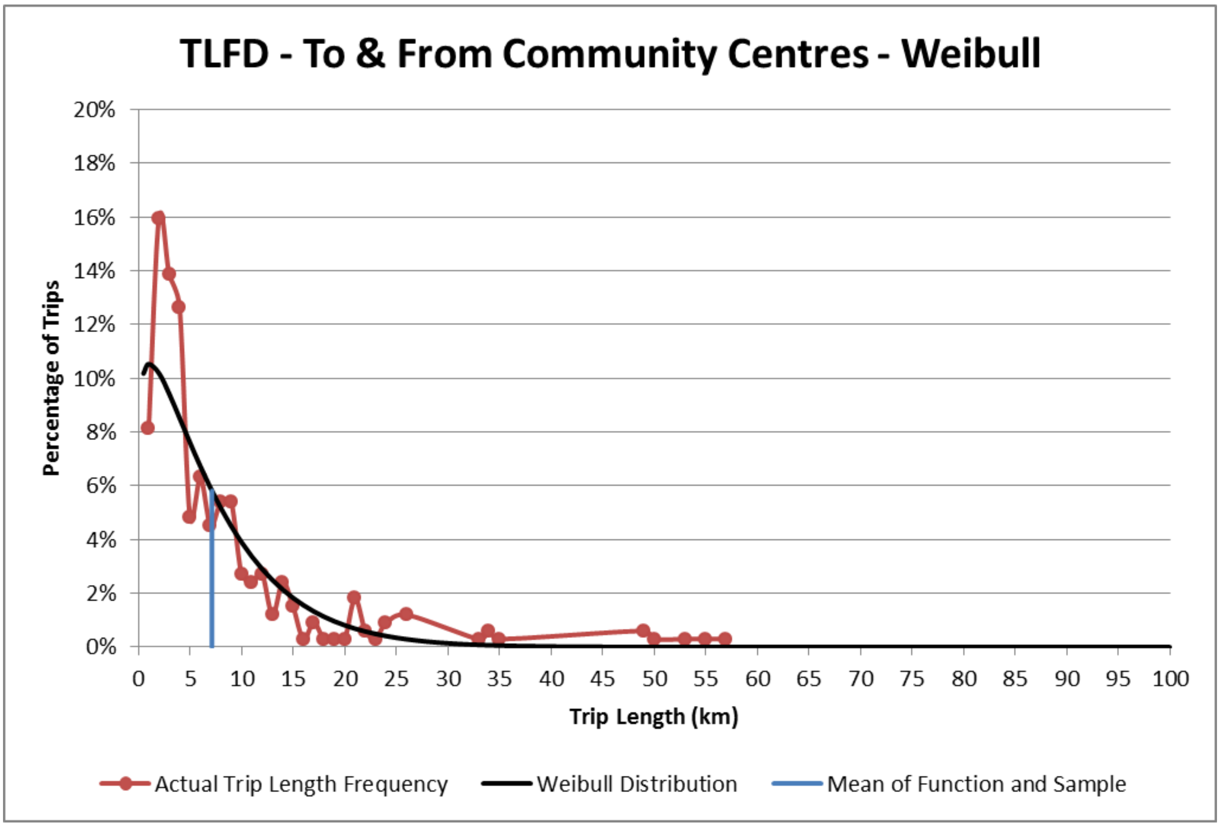
To & From Convenience Centres	
Variable	**Function Mean Equal to Sample Mean**
Alpha	0.83
Beta	7.41
R^2	0.702
RMS Error	0.025
Error Squared	0.022
Mean Diff (km)	0.0



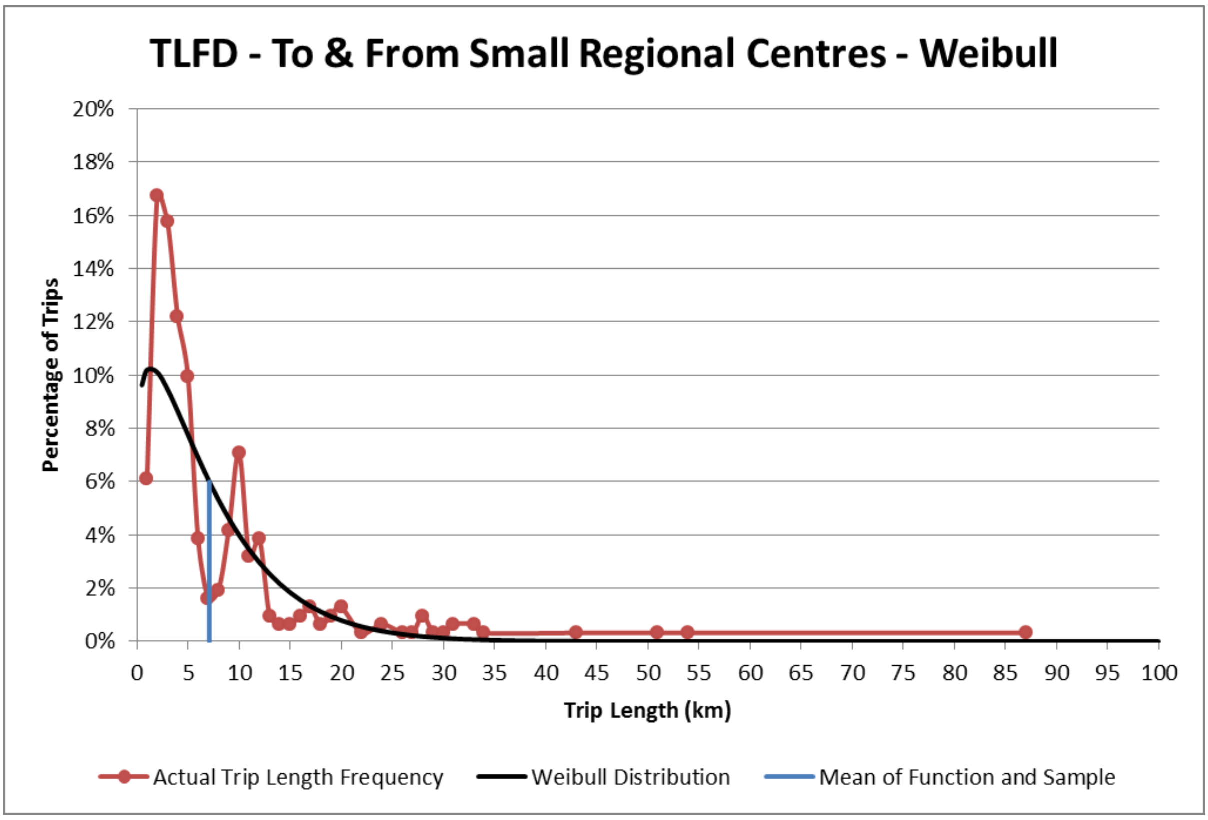
To & From Neighbourhood Centres	
Variable	**Function Mean Equal to Sample Mean**
Alpha	0.99
Beta	6.26
R^2	0.795
RMS Error	0.021
Error Squared	0.018
Mean Diff (km)	0.0



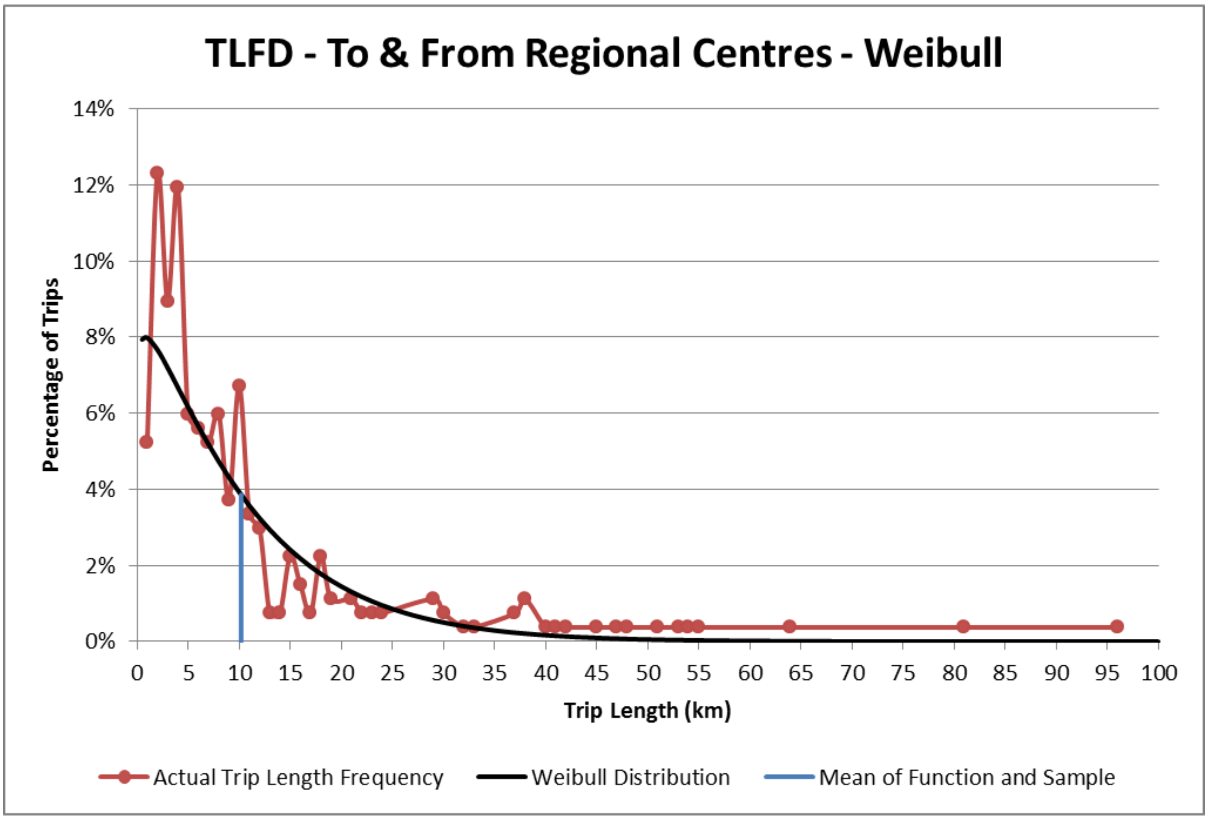
To & From Community Centres	
Variable	**Function Mean Equal to Sample Mean**
Alpha	1.13
Beta	7.45
R^2	0.823
RMS Error	0.017
Error Squared	0.010
Mean Diff (km)	0.0



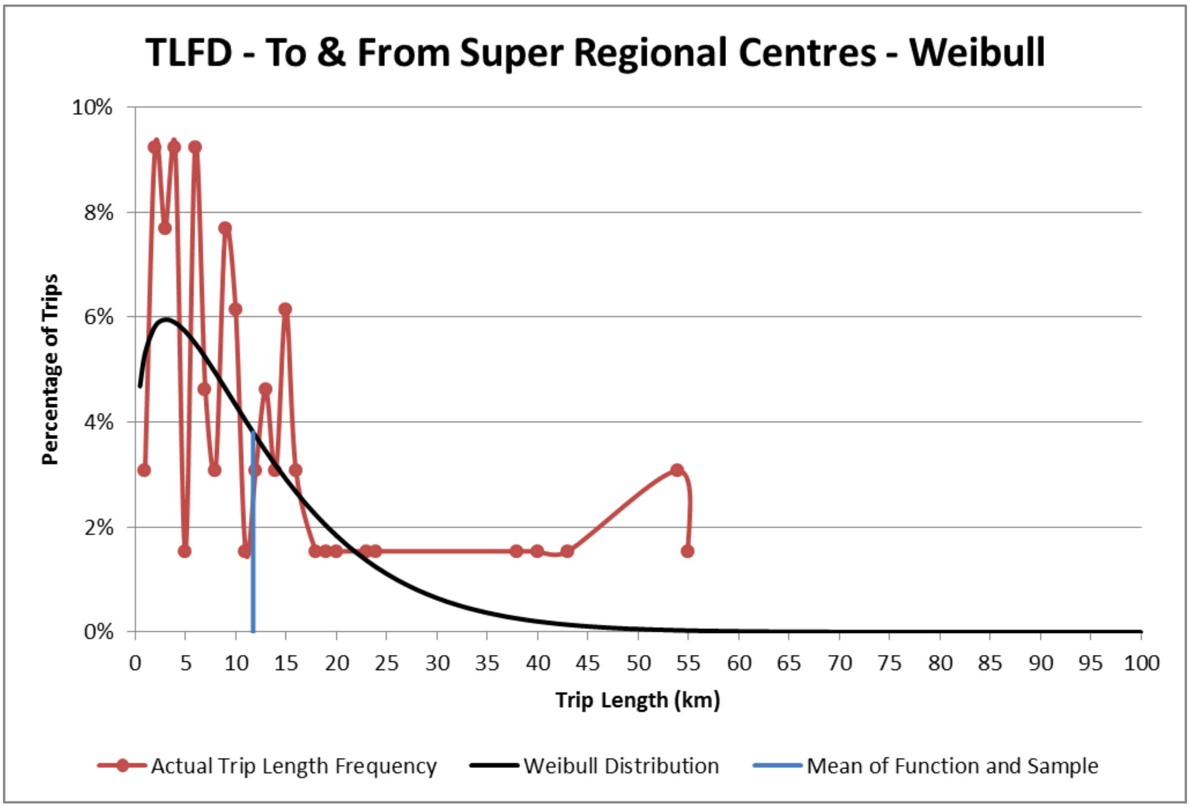
To & From Small Regional Centres	
Variable	**Function Mean Equal to Sample Mean**
Alpha	1.16
Beta	7.48
R^2	0.724
RMS Error	0.023
Error Squared	0.018
Mean Diff (km)	0.0



To & From Regional Centres	
Variable	**Function Mean Equal to Sample Mean**
Alpha	1.07
Beta	10.48
R^2	0.789
RMS Error	0.014
Error Squared	0.008
Mean Diff (km)	0.0



To & From Super Regional Centres	
Variable	**Function Mean Equal to Sample Mean**
Alpha	1.22
Beta	12.56
R^2	0.396
RMS Error	0.021
Error Squared	0.012
Mean Diff (km)	0.0

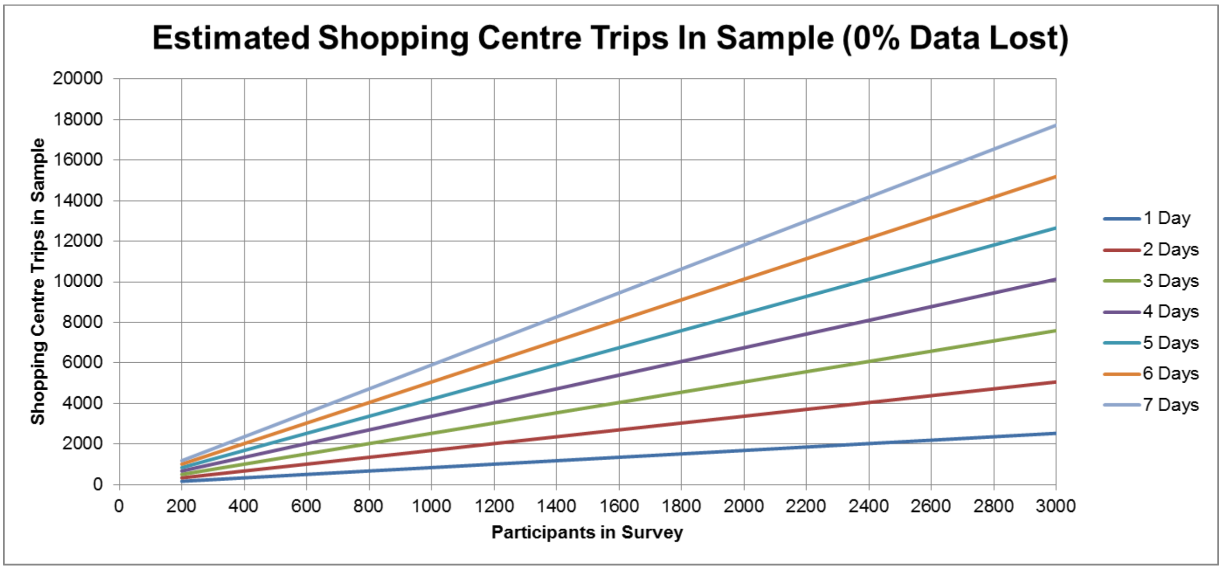


Appendix E

Estimated Shopping Centre Trips in Sample based on Participants and Duration of Survey

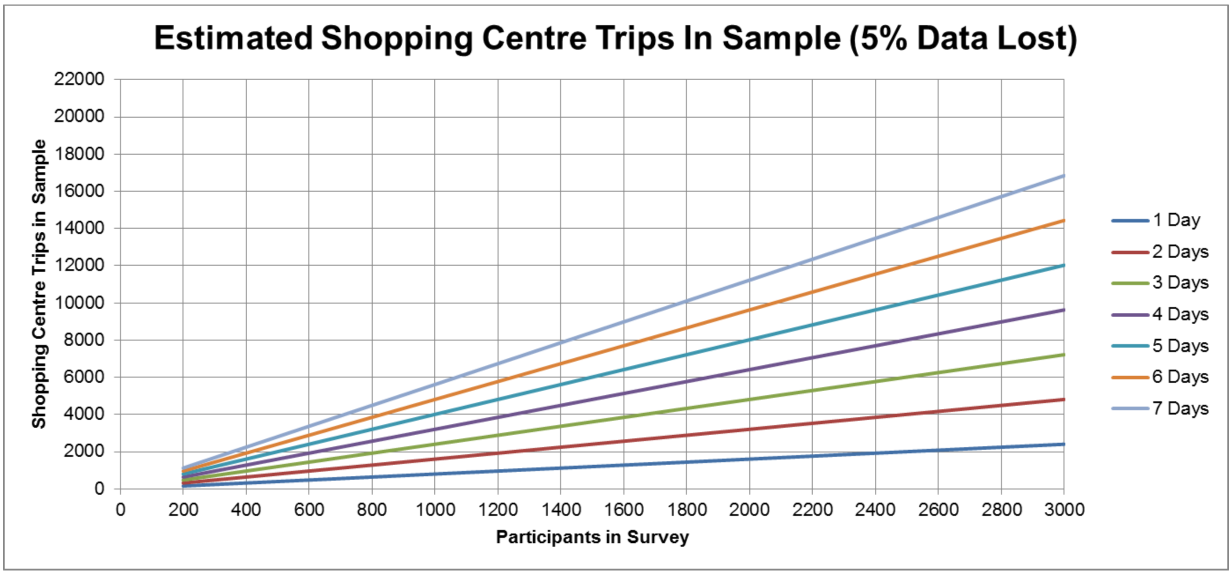
Estimated Shopping Centre Trips in Sample with 0% of Participants Data Lost

	Duration of Survey (Days)						
Participants	**1**	**2**	**3**	**4**	**5**	**6**	**7**
200	169	338	506	675	844	1013	1182
400	338	675	1013	1351	1688	2026	2363
600	506	1013	1519	2026	2532	3039	3545
800	675	1351	2026	2701	3376	4052	4727
1000	844	1688	2532	3376	4221	5065	5909
1200	1013	2026	3039	4052	5065	6078	7090
1400	1182	2363	3545	4727	5909	7090	8272
1600	1351	2701	4052	5402	6753	8103	9454
1800	1519	3039	4558	6078	7597	9116	10636
2000	1688	3376	5065	6753	8441	10129	11817
2200	1857	3714	5571	7428	9285	11142	12999
2400	2026	4052	6078	8103	10129	12155	14181
2600	2195	4389	6584	8779	10973	13168	15363
2800	2363	4727	7090	9454	11817	14181	16544
3000	2532	5065	7597	10129	12662	15194	17726



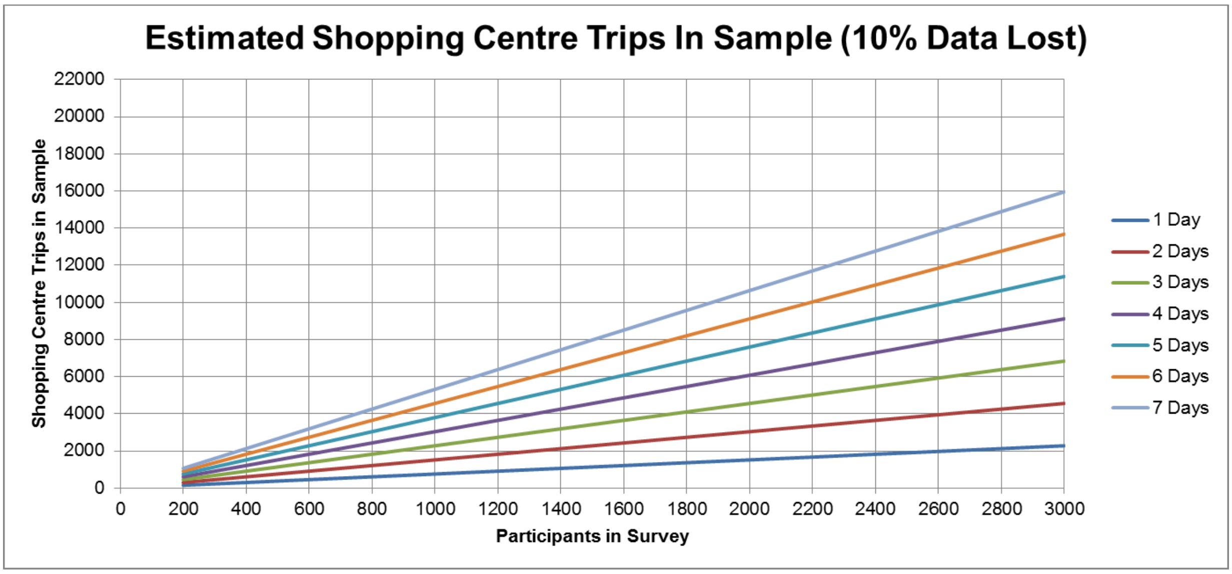
Estimated Shopping Centre Trips in Sample with 5% of Participants Data Lost

	Duration of Survey (Days)						
Participants	**1**	**2**	**3**	**4**	**5**	**6**	**7**
200	160	321	481	642	802	962	1123
400	321	642	962	1283	1604	1925	2245
600	481	962	1443	1925	2406	2887	3368
800	642	1283	1925	2566	3208	3849	4491
1000	802	1604	2406	3208	4010	4811	5613
1200	962	1925	2887	3849	4811	5774	6736
1400	1123	2245	3368	4491	5613	6736	7859
1600	1283	2566	3849	5132	6415	7698	8981
1800	1443	2887	4330	5774	7217	8661	10104
2000	1604	3208	4811	6415	8019	9623	11227
2200	1764	3528	5293	7057	8821	10585	12349
2400	1925	3849	5774	7698	9623	11547	13472
2600	2085	4170	6255	8340	10425	12510	14595
2800	2245	4491	6736	8981	11227	13472	15717
3000	2406	4811	7217	9623	12029	14434	16840



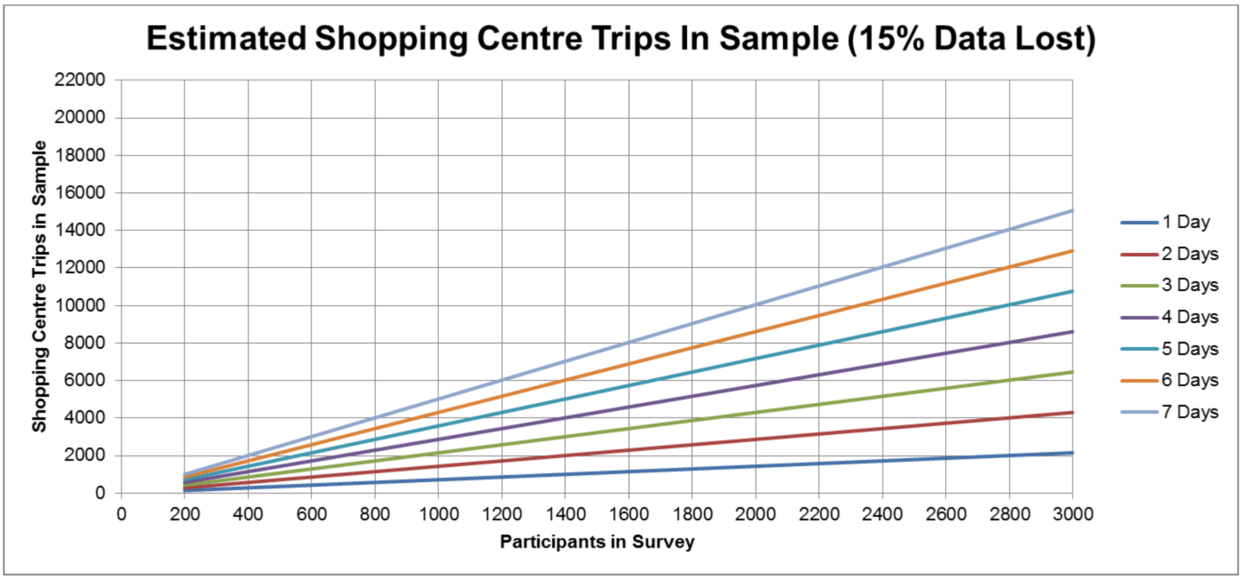
Estimated Shopping Centre Trips in Sample with 10% of Participants Data Lost

	Duration of Survey (Days)						
Participants	**1**	**2**	**3**	**4**	**5**	**6**	**7**
200	152	304	456	608	760	912	1064
400	304	608	912	1216	1519	1823	2127
600	456	912	1367	1823	2279	2735	3191
800	608	1216	1823	2431	3039	3647	4254
1000	760	1519	2279	3039	3798	4558	5318
1200	912	1823	2735	3647	4558	5470	6381
1400	1064	2127	3191	4254	5318	6381	7445
1600	1216	2431	3647	4862	6078	7293	8509
1800	1367	2735	4102	5470	6837	8205	9572
2000	1519	3039	4558	6078	7597	9116	10636
2200	1671	3343	5014	6685	8357	10028	11699
2400	1823	3647	5470	7293	9116	10940	12763
2600	1975	3950	5926	7901	9876	11851	13826
2800	2127	4254	6381	8509	10636	12763	14890
3000	2279	4558	6837	9116	11395	13675	15954



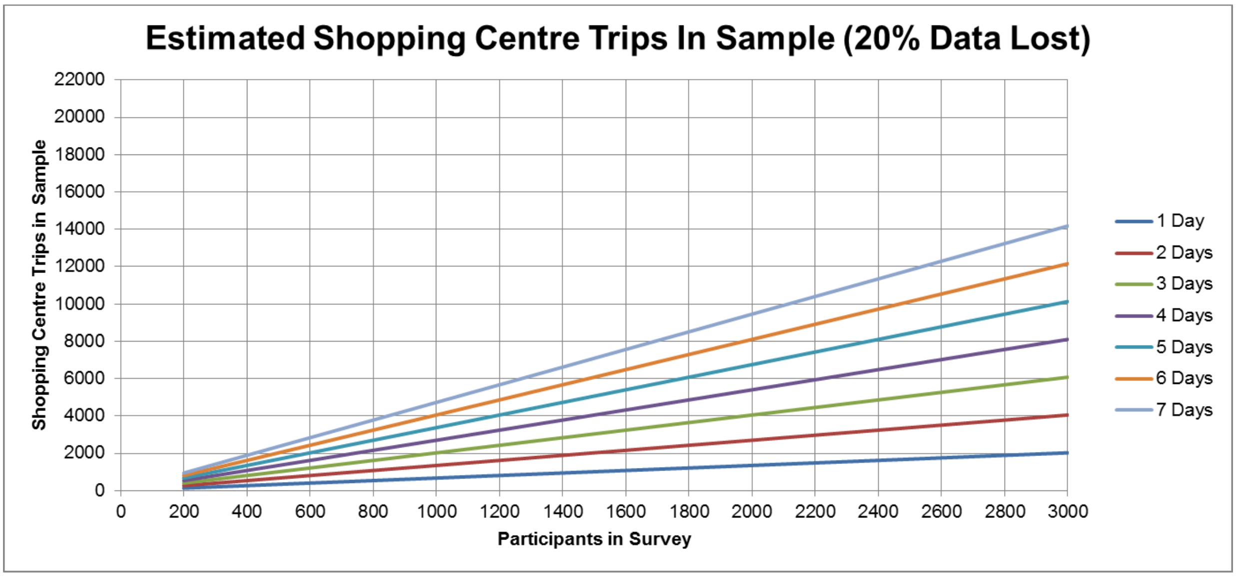
Estimated Shopping Centre Trips in Sample with 15% of Participants Data Lost

	Duration of Survey (Days)						
Participants	**1**	**2**	**3**	**4**	**5**	**6**	**7**
200	143	287	430	574	717	861	1004
400	287	574	861	1148	1435	1722	2009
600	430	861	1291	1722	2152	2583	3013
800	574	1148	1722	2296	2870	3444	4018
1000	717	1435	2152	2870	3587	4305	5022
1200	861	1722	2583	3444	4305	5166	6027
1400	1004	2009	3013	4018	5022	6027	7031
1600	1148	2296	3444	4592	5740	6888	8036
1800	1291	2583	3874	5166	6457	7749	9040
2000	1435	2870	4305	5740	7175	8610	10045
2200	1578	3157	4735	6314	7892	9471	11049
2400	1722	3444	5166	6888	8610	10332	12054
2600	1865	3731	5596	7462	9327	11193	13058
2800	2009	4018	6027	8036	10045	12054	14063
3000	2152	4305	6457	8610	10762	12915	15067



Estimated Shopping Centre Trips in Sample with 20% of Participants Data Lost

	Duration of Survey (Days)						
Participants	**1**	**2**	**3**	**4**	**5**	**6**	**7**
200	135	270	405	540	675	810	945
400	270	540	810	1080	1351	1621	1891
600	405	810	1216	1621	2026	2431	2836
800	540	1080	1621	2161	2701	3241	3782
1000	675	1351	2026	2701	3376	4052	4727
1200	810	1621	2431	3241	4052	4862	5672
1400	945	1891	2836	3782	4727	5672	6618
1600	1080	2161	3241	4322	5402	6483	7563
1800	1216	2431	3647	4862	6078	7293	8509
2000	1351	2701	4052	5402	6753	8103	9454
2200	1486	2971	4457	5943	7428	8914	10399
2400	1621	3241	4862	6483	8103	9724	11345
2600	1756	3511	5267	7023	8779	10534	12290
2800	1891	3782	5672	7563	9454	11345	13236
3000	2026	4052	6078	8103	10129	12155	14181



Estimated Shopping Centre Trips in Sample with 30% of Participants Data Lost

	Duration of Survey (Days)						
Participants	**1**	**2**	**3**	**4**	**5**	**6**	**7**
200	118	236	355	473	591	709	827
400	236	473	709	945	1182	1418	1654
600	355	709	1064	1418	1773	2127	2482
800	473	945	1418	1891	2363	2836	3309
1000	591	1182	1773	2363	2954	3545	4136
1200	709	1418	2127	2836	3545	4254	4963
1400	827	1654	2482	3309	4136	4963	5791
1600	945	1891	2836	3782	4727	5672	6618
1800	1064	2127	3191	4254	5318	6381	7445
2000	1182	2363	3545	4727	5909	7090	8272
2200	1300	2600	3900	5200	6500	7800	9099
2400	1418	2836	4254	5672	7090	8509	9927
2600	1536	3073	4609	6145	7681	9218	10754
2800	1654	3309	4963	6618	8272	9927	11581
3000	1773	3545	5318	7090	8863	10636	12408