Validation of TomTom historical average speeds on freeway segments in Gauteng, South Africa

by
Batsirai Gwara

Thesis presented in fulfilment of the requirements for the degree of Master of Engineering in the Faculty of Engineering at Stellenbosch University

Supervisor: Dr Johann Andersen

March 2017

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

March 2017

ABSTRACT

Traditional methods of traffic data collection, such as inductive loops and road sensors, continue to be the main source of traffic data. The advancement in technology and vehicle tracking methods has proved to be the impetus behind the emerging of alternative and innovative sources of traffic data, such as ITS data sources. ITS sources, such as vehicle probes, are becoming increasingly important due to their low cost and the vast amounts of traffic data produced. However, traffic data from ITS sources raise new concerns about data quality. The quality of probe data in South Africa and other developing countries is unknown. This study sets out to investigate the quality of TomTom historical average speeds on selected freeway segments in South Africa.

The study compared TomTom historical speed estimates and reference speeds on six directional segments on the N1 and R21 freeways. The reference data used was Automatic Number Plate Recognition (ANPR) data, a component of Open Road Tolling (ORT) in Gauteng. A freeway segment is the road section between two toll gantries. All 15-minute and 1-hour intervals between 05:00 and 20:00 during the weekdays (Monday – Friday) in February 2015 were grouped and aggregated. The quality measures evaluated were accuracy, completeness, validity, coverage and accessibility.

To evaluate accuracy, three error quantities were determined, namely signed error, average absolute speed error (AASE) and speed error bias (SEB). The allowable errors for the signed error, AASE and SEB were ±10 %, 10 km/h and ±7.5 km/h, respectively. TomTom speeds were highly consistent with the reference speeds. The error quantities for the combined freeway segments were less than the allowable errors. The signed errors and AASE for all the six individual freeway segments were also less than the allowable errors. In five of the six sections, the SEB was less than the allowable error. There were no significant differences between the error quantities derived from 15-minute and 1-hour interval speeds for the combined and individual freeway segments. On the other hand, validity was dependent on the selected measure. TomTom speeds were of very high quality based on the signed error and AASE, whereas the same data was of moderate quality based on the SEB.

Although the TomTom speeds were within the specified accuracy thresholds, the speed estimates were generally lower than the reference speeds throughout the analysis period. TomTom estimates were better at low speeds and the quality of TomTom estimates declined with an increase in speed. It is possible that the low TomTom speed estimates were due to a sample that was not a true representation of the traffic stream. Importantly, it is possible to enhance the accuracy of TomTom speed estimates by using certain percentile speeds instead of average speeds.

OPSOMMING

Tradisionele metodes vir die insameling van verkeerdata, soos byvoorbeeld induksie lusse en padsensors, is tans die hoofbron van verkeerdata. Die vooruitgang in tegnologie en voertuig monitering is tans die dryfkrag van alternatiewe en innoverende bronne van verkeerdata, soos byvoorbeeld Intelligente Vervoer Stelsels (IVS) databronne. IVS bronne, soos voertuigsondes, se toepaslikheid neem toe weens lae koste en die hoeveelheid data wat versamel word. Een bekommernis aangaande data vanaf IVS bronne is die data kwaliteit. Die kwaliteit van sondes se data in Suid-Afrika en ander ontwikkelende lande is nie geverifieer nie. Hierdie studie ondersoek die kwaliteit van spoed metings vanuit TomTom se historiese data vir deurpadsegmente in Suid-Afrika.

Hierdie ondersoek vergelyk TomTom se historiese snelhede verwysing snelhede op ses segmente op die N1 en R21 deurpaaie. Die verwysingdata was afkomstig van outomatiese nommerplaat identifisering (ANPR), ’n komponent van “Open Road Tolling” (ORT) in Gauteng. ’n Segment is gedefinieer as die seksie tussen twee tol stellasies. Alle 15-minuut en 1-uur intervalle tussen 05:00 en 20:00 tydens die weekdae (Maandag-Vrydag) in Februarie 2015 was gegroepeer en opgesom. Die kwaliteitsmaatstawwe wat geëvalueer is sluit akkuraatheid, volledigheid, geskiktheid, dekking en toeganklikheid in.

Om akkuraatheid te evalueer was drie foutmaatstawwe bepaal, naamlik getekende fout, gemiddelde absolute spoed fout (AASE) en spoed-fout-vooroordeel (SEB). Die toelaatbare foute vir die getekende fout, AASE en SEB was ±10 %, 10 km/h en ±7.5 km/h, respektiewelik. TomTom snelhede het uitstekend korreleer met die verwysingsnelhede. Die fout meting vir die gekombineerde deurpadsegmente was minder as die toelaatbare foute. Die getekende foute en AASE vir al ses die individuele deurpad segmente was ook minder as die toelaatbare foute. Vir vyf van die ses segmente was die SEB minder as die maksimum waarde. Daar was geen noemenswaardige verskille tussen fout maatstawwe tussen die 15-minuut en 1-uur interval snelhede vir die gekombineerde en individuele deurpad segmente nie. Die geskiktheid van die data was afhanklik van die gekose maatstaaf. TomTom snelhede was van hoë gehalte gebaseer op die getekende fout en AASE, maar was van matige kwaliteit gebaseer op die SEB.

Alhoewel die TomTom snelhede binne die voorgeskrewe perke was, was dit in die algemeen laer as die verwysingsnelhede vir die meeste van die analise periode. TomTom voorspellings was beter vir laer snelhede en die kwaliteit van TomTom data het afgeneem met ’n toename in snelheid. Dit is moontlik dat die lae TomTom snelheidvoorspellings ’n gevolg is van ’n monster wat nie ’n ware verteenwoordiging van die verkeerstroom is nie. Die akkuraatheid van snelheidvoorspellings verbeter word deur ‘n sekere persentiel snelheid te gebruik in plaas van gemiddelde snelhede.

DEDICATION

"To my mom, Mrs C. Gwara"

ACKNOWLEDGEMENTS

First and foremost, I would like to thank God for giving me the strength to complete this research project.

I would like to offer my sincerest gratitude to my supervisor, Dr Johann Andersen, who supported and guided me throughout this project. His encouragement, supervision and advice made it possible to complete this thesis. I also thank him for the opportunities to travel during the course of this degree program.

I also thank Mrs. Bruwer for her guidance and advice and Prof Kidd (Centre for Statistical Consultation) for his help with the statistical aspects of this study.

I also appreciate my family for their continuous support, encouragement and prayers.

My utmost appreciation and special thanks to my friend, Ronald Chinyemba, for his significant contribution in developing the Java program for data processing. His input and programming skills were of great value.

I would like to thank the students in the Stellenbosch Smart Mobility Lab (SSML), staff in the Civil Engineering department and friends for their support.

I also thank TomTom and SANRAL (in particular, Alan Robinson) for the data used in this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

AASE	-	Average Absolute Speed Error
ADT	-	Annual Daily Traffic
ANPR	-	Automatic Number Plate Recognition
BT	-	Bluetooth
CSV	-	Comma Separated Value
DMI	-	Distance Measuring Instrument
GB	-	Gigabyte
GFIP	-	Gauteng Freeway Improvement Project
GPS	-	Global Positioning System
GUI	-	Graphical User Interface
ITLS	-	Institute of Transport and Logistics Studies
ITS	-	Intelligent Transport Systems
MAC	-	Machine Access Control
NB	-	Northbound
NCHRP	-	National Cooperative Highways Research Program
ORT	-	Open Road Tolling
PND	-	Personal Navigation Device
RMSE	-	Root Mean Square Error
SANRAL	-	South African National Road Agency SOC Limited
SB	-	Southbound
SEB	-	Speed Error Bias
SSML	-	Stellenbosch Smart Mobility Laboratory
TIS	-	Traveller Information Service
TMC	-	Traffic Management Centre
VMS	-	Variable Message Sign
VPP	-	Vehicle Probe Project

LIST OF SYMBOLS USED

μ_{ORT}	=	mean ORT speed in km/h
μ_{TOMTOM}	=	mean TomTom speed in km/h
$AASE_{allowable}$	=	allowable AASE (average absolute speed error)
E_1	=	signed error in one 15-minute or 1-hour interval
E_2	=	AASE in one 15-minute or 1-hour interval
E_3	=	SEB in one 15-minute or 1-hour interval
L	=	length of a road segment in km
N	=	number of observations
$n_{AVAILABLE\ VALUES}$	=	the number of records or rows with available values present
N_{MIN}	=	minimum sample number
N_{ORT}	=	number of ORT observations
N_{TOMTOM}	=	number of TomTom observations
$n_{TOTAL\ EXPECTED}$	=	the number of records or rows expected.
n_{TOTAL}	=	the total number of rows subjected to validity criteria
n_{VALID}	=	the number of rows with values meeting validity criteria
PV	=	percent valid
S.E	=	signed error
$SEB_{allowable}$	=	allowable SEB (speed error bias)
S_{ORT}	=	reference speed (ORT) in a 15-minute or 1-hour interval
S_{TOMTOM}	=	TomTom speed estimate in a 15-minute or 1-hour interval
T_i	=	travel time of the i^{th} observation in seconds; and
$z_{\alpha/2}$	=	critical normal deviate for the desired confidence interval
θ	=	error tolerance level
σ_{ORT}	=	standard deviation of ORT speeds in km/h
σ_{TOMTOM}	=	standard deviation of TomTom speeds in km/h

CHAPTER 1 : PRELUDE

1.1 INTRODUCTION

Traffic data is a prerequisite in transportation operations and planning. For this reason, collecting sufficient, accurate and reliable traffic information, in a cost effective manner, is important for traffic engineers and authorities. Conventional methods of traffic data collection include inductive loops and road sensors. To date, the conventional methods continue to be the main source of traffic data. Newer and innovative traffic data sources have since emerged. An example thereof is Intelligent Transport Systems (ITS) sources.

ITS sources produce large amounts of data, which has potential use in transportation applications. The usefulness of ITS data has only begun to be realised. Although there is potential for ITS data to be used in a wide range of mainstream applications of transportation planning, engineering and operations, there are institutional, technical and possibly financial issues that still need to be resolved (Battelle, Cambridge Systematics Inc & Texas Transportation Institute, 2004). One of the technical issues is the concern over the quality of the data. The vast amounts of traffic data from ITS sources raise new concerns about the quality of the data. The issues pertaining to the quality of the data could hinder the potential of ITS data in fulfilling data requirements for transportation planning, engineering and operations.

1.2 BACKGROUND

Traffic data provides key indicators for various transportation operations and planning purposes. To meet the aforementioned purposes, the quality of the data must be assured. Collecting timely, accurate and reliable traffic data is essential for effective traffic management as it gives the transportation agencies confidence when using the data (Miles & Chen, 2004; Haghani, Hamedi, Sadabadi, Young & Tarnoff, 2010). Recent studies identified various issues regarding the quality of traffic data from ITS sources (Battelle *et al.*, 2004). These issues include accuracy and validity. To address these challenges and concerns, there is a need to conduct data quality assessments. As such, there is a need to develop methods and tools to assess the quality of the data.

A state of practice report by the National Cooperative Highways Research Program (NCHRP) looked at several data quality assessments of probe-based travel time estimation technology that have been conducted (National Cooperative Highways Research Program, 2009). The goal of the data quality assessments was to measure the accuracy of speed and travel time estimates from a data service provider or probe-based technology. The assessments aimed to address the data quality concerns. To address the issues regarding the quality of probe data, there are a number of "quality measures" that should be considered. Although different assessments investigate different quality measures, the quality measures of traffic data that are usually of interest are accuracy, completeness, validity, timeliness, coverage and accessibility (Battelle *et al.*, 2004). The definitions of these quality measures will be discussed in a later section.

To be able to assess data quality, one needs to understand the attributes of traffic data. These are travel time and speed. Travel time is defined as the time required to traverse a given road segment between any two points of interest. Speed is defined as the rate at which a vehicle traverses a given road segment. It is apparent that the two attributes are closely related as one attribute can be easily determined if the other is known. Although these two attributes are generally used, travel time is widely understood and communicated by a wide variety of audiences including transportation engineers, planners and data users.

Road users rely on travel time information to make informed decisions on travel choices in order to avoid unnecessary delay and minimise trip times (Liu & Ma, 2009). Transportation agencies use travel time data for monitoring and planning purposes as well as for emergency responses (Aliari & Haghani, 2012). As already mentioned, ITS technologies offer newer methods of obtaining data.

Traditionally, travel information is acquired from road-based data sources such as inductive loop detectors and traffic sensors, which include sonic, radar and infrared sensors. The main disadvantage of inductive loop detectors is the disruption of traffic flow during installation and maintenance whilst the traffic sensors are susceptible to bad weather conditions. A key factor in traffic management system is information about the transport networks. Obtaining information about a transport network involves capturing data on the links and interconnections by means of videoing the networks, aerial photographing and on-site surveys (Miles & Chen, 2004). As a result, collection, reduction and analysis of the data are cumbersome and labour intensive processes. The high cost of the road-based data sources makes it challenging to manage traffic in an effective and sustainable manner.

Travel time information can also be collected by means of vehicle-based data sources. Vehicles act as mobile sensor nodes that are equipped with a location aware device such as a GPS unit and a communication device such as a cell phone (Ayala, Lin, Wolfson, Rishe & Tanizaki, 2010). With the advancement in vehicle tracking, identification technology and the proliferation of location aware and connected devices, vast amounts of data are generated by the millions of vehicles on the traffic network (Zhang, Hamedi & Haghani, 2015). Evidently, this method of data collection is advantageous as it is cost-effective compared to the conventional traffic sensors which require road-based infrastructure for data collection.

1.3 RESEARCH PROBLEM

Probes are vehicles that anonymously provide traffic data, for example, vehicles equipped with personal navigation devices (PND) or in-dash GPS, smartphones and commercial vehicles with GPS devices. Vehicle probes are becoming more important because there has been comprehension and awareness of the advantages, such as low costs and vast amounts of traffic data generated. Therefore, it is not surprising that traffic data vendors such as TomTom and Inrix are promoting the collection and utilisation of probe data for traveller information.

The technological advancement in traffic detection systems and the growth of ITS infrastructure have provided innovative sources of traffic data. Although the new and innovative sources of traffic data suggest more data availability, concerns about the quality of the data are also raised. The quality of probe data is the primary focus of this study. It was established that the innovative ways of collecting data result in a new set of challenges. As such, there is a need to develop methods and tools to determine the data quality. If the data quality assessments are successful, it gives confidence to the transportation agencies and data users. Successful assessments can possibly make way for ITS data to be used in mainstream applications of transportation operations, engineering and planning.

Several data quality assessments have been conducted. However, most of the assessment projects were conducted in developed countries (National Cooperative Highways Research Program, 2009). The quality of probe data in South Africa and other developing countries is not verified. Hence, the data should not be immediately used for traffic management and planning purposes. The penetration rates and different sources of raw data are reasons for possible differences in the quality of probe data in different parts of the world. Where probe data is used to monitor current traffic conditions and predict future traffic conditions, the accuracy of the predicted traffic conditions is not guaranteed, given that the accuracy of the probe data is unknown in the first place. It is not ideal to use unsubstantiated probe data for traffic management, as it is not reliable.

For probe data to be used for transportation operations and planning purposes in South Africa, validation assessments have to be conducted to determine the quality of the data. Quality measures of probe data have to be checked against benchmark data. Benchmark data, which is also known as reference data, is the data that is assumed to be correct, against which estimates are compared. In the South African context, Open Road Tolling (ORT) data or data from Bluetooth (BT) readers can be used as a source of reference data. Bluetooth has been identified as a cost-effective source of reference data. Studies that have used Bluetooth as reference data include Haghani *et al.* (2009), Haghani *et al.* (2010) and Aliari & Haghani (2012).

The ORT system uses Automatic Number Plate Recognition (ANPR) for the purposes of electronic toll collection. Therefore, ANPR is a component of Open Road Tolling on the Gauteng freeways. Currently, ORT data is used in network modelling, trip generation and origin-destination studies. Due to the high quality and high-density data, the ORT system is an ideal source of benchmark data (Robinson, 2014, 2016).

1.4 RESEARCH QUESTION

The research question for this study is as follows:

Is the quality of TomTom historical average speeds on freeway segments consistent with reference speeds?

1.4.1 RESEARCH GOALS

In response to the research question defined above, the goals of this study are:

1. To identify the quality measures of probe data based on international studies; and
2. To determine the quality of TomTom historical average speeds using the quality measures developed in international studies as a basis for comparison.

1.4.2 RESEARCH OBJECTIVES

The research objectives for this study are as follows:

i. To develop a suitable data processing plan to process and manipulate the reference data.

ii. To investigate the quality measures of TomTom historical average speeds. The quality measures include:

    1. Accuracy
    2. Completeness
    3. Validity
    4. Coverage
    5. Accessibility

iii. To identify ways of explaining and reducing the differences between TomTom speed estimates and reference speeds.

1.5 SIGNIFICANCE OF RESEARCH

One of the key conclusions from the 1st National Congress and Exhibition on Intelligent Transport Systems for Highways was that society can no longer afford to operate non-intelligent transportation systems (Mahmassani, 2014). Non-intelligent transport systems are detrimental to the economy and the environment due to inadequate responses to increased traffic congestion. Traffic congestion negatively affects the economy due to high fuel consumption whilst increased exhaust gas emissions are harmful to the environment.

The Institute of Transport and Logistics Studies (ITLS Africa) at the University of Johannesburg conducted a survey (State of Transport Opinion Poll) between October and November 2012 to determine the critical issues facing the South African society. A sample of 1000 adults aged 18 years and above, representing the distribution and demographics of South Africa participated in the poll. Results of the poll identified transport problems as the third highest priority issue in South Africa, behind education and health (Luke & Heyns, 2013a, 2013b, 2016).

South Africa and other developing countries have limited budgets for development and maintenance of transport infrastructure. It is therefore important to find effective, sustainable and environmentally friendly transport systems that are within the allocated budgets.

Although road-based data sources have been the main source of traffic data for a long time, there are still issues regarding the cost of these systems. It is worthwhile to note that the initial capital investments and maintenance costs for road-based data sources are quite substantial. The operating and maintenance components of the cost are often overlooked. As such, the life cycle costs should be considered as opposed to only the initial capital investments.

The coverage of road-based data sources is often limited. Urban areas typically have data collection devices whilst most rural areas do not have road-based infrastructure for data collection. The study demonstrates how probe data overcomes the challenge of coverage, as probe data does not require any stations or road-based infrastructure for data collection.

The study sets out to investigate the quality of probe data as an alternative source of traffic data. The study is approached from the context of developing countries such as South Africa. Therefore, this investigation is significant for developing countries as it seeks to promote sustainable and effective transportation systems and ultimately, contribute in addressing the critical issues facing South Africa. This is an opportunity to supplement the existing traffic data from road-based sources with probe data and unlock the wider use of probe data as a key source of traffic data.

1.6 LIMITATIONS/SCOPE

The traffic data service provider used in this study is TomTom and the data used was historical traffic data. The decision to use TomTom data was as a result of availability of the data through a licencing agreement between the Stellenbosch Smart Mobility Laboratory (SSML) of Stellenbosch University and TomTom Africa (Pty) Ltd.

ORT data was used as the reference data. The selection of ORT data was based on the fact that the ORT system is a state-of-the-art system that is already calibrated, from which reliable, high density and high-quality traffic data is obtained.

In South Africa, the ORT system is operational in Gauteng. Therefore, the area of study was limited to Gauteng. Furthermore, only the national routes under electronic tolling were investigated in this study. The freeway segments that were investigated were segments that start and end on the same freeway. No attempt was made to investigate sections that originate from one freeway and end on another freeway i.e. merging and diverging sections were not considered. The length of the sections investigated varied in relation to the permanent ORT infrastructure.

Only 15-minute and 1-hour time intervals were investigated. In addition, the data in this study is for 15 hours a day, from 05:00 to 20:00 for all the weekdays (Monday – Friday) in the month of February 2015. Other time intervals such as 30-minute interval were not investigated in this study. A detailed explanation on the time elements of the study is presented in Chapter 3.

Traffic information can be classified into real-time and historical information. Historical information is the data on the conditions of the road network for a period in the past; whereas, real-time data is the information on the traffic conditions as they occur. The quality measure, timeliness, is useful when assessing real-time traffic information to get an indication of the delay between collecting the data in the field, relaying the data to the Traffic Management Centre (TMC) and sending back the information to the Variable Message Signs (VMS) on the road network. The fundamental measures of traffic data quality are accuracy, completeness, validity, timeliness, coverage and accessibility (Battelle *et al.*, 2004). Of these six measures, timeliness was not investigated in this study.

ORT data gives a complete representation of the traffic that traversed on the road network, consisting of date and time, traffic volume, vehicle classification and vehicle identification (by means of an anonymised ID) of the vehicles on the road network. Information on the individual vehicles sampled by the service provider and distribution of the sample in terms of vehicle classification was not provided. Only aggregated data of the proportion of the traffic stream sampled was provided. TomTom data was derived from various sources such as GPS devices, GSM probes and incident data. There was no indication of the percentage of the data obtained from the different sources.

1.7 ASSUMPTIONS

The following assumptions were made:

i. The variance of the sample used by the service provider in deriving the speed estimates is equal to the variance of the entire traffic stream.

ii. The traffic conditions do not change significantly in a 15-minute time interval.

iii. The entire traffic stream on a given freeway segment is captured on the ORT system.

1.8 RESEARCH OVERVIEW

In **Chapter 2 (Literature review)**, a comprehensive literature survey on the data collection techniques and international guidelines and standards on assessing traffic data quality is presented. An analysis of the available literature on a number of selected validation projects is conducted. To conclude this chapter, the lessons learnt from these projects are examined.

The procedure for conducting a data quality assessment is explored further in **Chapter 3 (Research design and methodology).** The chapter begins by discussing the research design and rationale for the research design. The methodology section explores the experimental set-up and the different elements of data collection, processing and reduction. The analysis conducted and the limitations of the project are discussed in the final sections of the chapter.

The findings of the study are presented in **Chapter 4 (Results)**. The results chapter start by presenting the findings from speed profiles for both the 15-minute and 1-hour intervals. Subsequently, the results of the quality measures are presented.

An in-depth discussion of the findings is presented in **Chapter 5 (Discussion)**. In this chapter, the issues relating to the differences between the TomTom speed estimates and reference speeds are explored.

The outstanding issues from Chapter 4 and Chapter 5 are explored further in **Chapter 6 (Further analysis)**. A comparison between the findings of this study and other data quality assessments is presented, discussing possible similarities, and differences thereof, of the results. Issues pertaining to daily variation of speeds and possible correction of TomTom speed estimates are also discussed.

To conclude this study, the conclusions and recommendations are presented in **Chapter 7 (Conclusions and recommendations for future research).** This chapter highlights how the research objectives were achieved. The conclusion and areas recommended for future research are presented in the final sections of the chapter.

CHAPTER 2 : LITERATURE REVIEW

2.1 INTRODUCTION

The purpose of this chapter is to review the literature on assessing the quality of traffic data from a service provider. The chapter begins by discussing the available guidelines and standards on evaluating the quality measures of traffic data. Thereafter, the different components of a data quality assessment such as developing an evaluation plan, collecting reference data and computing the quality measures are discussed. To end the chapter, a discussion on the selected validation projects and lessons learnt from these projects is presented.

2.2 GUIDELINES FOR ASSESSING THE TRAFFIC DATA QUALITY

In recent years, there has been an interest in finding alternative sources of traffic data. Due to the technological advancement, vast amounts of traffic data have been obtained from Intelligent Transport Systems (ITS) sources. However, the drawback of ITS data is the uncertainty over the quality of the data (Battelle *et al.*, 2004).

Formal and informal evaluations of traffic data that have been conducted differ in the methodologies used in comparing the data (Turner, Richardson, Fontaine & Smith, 2011). In order to address the issue of different methodologies and procedures in evaluations, guidelines and standards have been developed to allow fair assessment of the data quality measures from different data sources, including ITS and traditional data sources. However, the major drawback of most guidelines is the fact the guidelines were developed for real-time data, possibly because of its wide use in everyday applications. Although most of the guidelines are based on real-time data, recent studies have applied the same evaluation procedures in assessing the quality of historical traffic data (Texas Transportation Institute, 2012).

The available guidelines include "*Travel time data collection handbook*", "*Guidelines for evaluating the accuracy of travel time and speed data*", "*Traffic data quality measurement*" and "*Traffic information benchmarking guidelines*" (Turner, Eisele, Benz & Douglas, 1998; Turner *et al.*, 2011; Battelle *et al.*, 2004; North Amercian Traffic Working Group, 2010). These guidelines were intended to be informative and not prescriptive, with the intention of providing guidance and consistency on the evaluation parameters. Furthermore, the guidelines outline the procedures for collecting and reducing traffic data (Turner *et al.*, 2011). However, the guidelines do not establish the level of accuracy that a traffic data supplier must achieve. As such, the guidelines states that the quality level must be defined based on the intended application of the data. This suggests that a user or purchaser of the data has the flexibility to select the quality level. This approach makes sense because different applications require different levels of accuracy. However, a serious weakness associated with this approach is the fact that it leads to inconsistent assessment results even for the same intended application.

The previous paragraph introduced the idea of different users of traffic data and the different types of data. In this section, a detailed discussion on the different uses of the data, types of traffic data and the type of data likely to be used by the different data consumers is presented. Table 2.1 shows the various entities that typically use the guidelines to evaluate the quality of traffic data.

Table 2.1: Different types of data evaluators (Turner *et al.*, 2011)

Entity	**Uses for the Guidelines**
Public agencies or third party evaluators	Evaluate the accuracy of traffic data prior to purchasing, on-going quality assurance process, ensuring that contractual specifications are met.
Private companies or third party evaluators	Assess the accuracy of data provided by other companies
Commercial data provider	Validate the data provided meets a specified quality level
Public agencies and private companies	Internal assessment of traffic data, enhance their own data quality

After identifying the different entities that conduct data quality assessments, it is important to understand the various types of data that are available for evaluations. The main types of traffic data that were identified by Battelle *et al.* (2004) were:

i. Original source data – data from data collection devices, can be raw, real-time or historical data.
ii. Archive data – historical data stored in a database. Archive data can be either stored as raw data or aggregated data.
iii. Traveller information – information disseminated to travellers, usual as real-time data.

The various types of data and data users are considered in a data quality framework. The data quality assessment is dependent on the intended application and the type of data. Due to the vast range of applications of traffic data, the level of quality is different for the various data users and uses. A summary of the data users, types of data and typical applications of traffic data is shown in Table 2.2.

Table 2.2: Typical applications of the various data types and data users (Battelle *et al.*, 2004)

Data consumers or users	**Types of data**	**Applications or uses**
Traffic operators	Original source data	Traffic management
	Archived source data	Incident management
Archived data administrators	Original source data	Database administration
Archived data users	Original source data	Analysis
	Archived source data	Planning
	Archived processed data	Modelling (development and calibration)
Traffic data collectors	Original source data	Traffic monitoring
	Archived source data	Equipment calibration
		Data collection planning
Information service providers	Original source data (real-time)	Dissemination of information
Travellers	Traveller information	Pre-trip planning

2.3 TRAFFIC DATA QUALITY ASSESSMENTS

A traffic data assessment typically consists of three major parts, namely:

1. Developing an assessment plan;
2. Reference data collection and reduction; and
3. Computing and reporting of data quality measures.

Alternatively, Battelle *et al.* (2004) presented the framework for data quality measurement as a sequence of steps. The steps provide recommendations for determining the data quality measures for various applications in transportation projects. The steps are as follows:

i. Define data customers;
ii. Define data measures;
iii. Establish acceptable data quality targets;
iv. Calculate data quality measures;
v. Identify data quality deficiencies;
vi. Assign responsibility and automate reporting; and
vii. Perform periodic assessment.

Turner *et al.* (2011) defined methodologies for assessing the accuracy of traffic data for the three specialised cases. These are congestion level, link speed and route travel time. The fundamental data quality measures are accuracy, validity, completeness, coverage, accessibility and timeliness. Turner *et al.* (2011) only proposed guidelines for evaluating the accuracy of traffic data. In light of this limitation, the procedures by Battelle *et al.* (2004) were adapted to develop the methodologies for evaluating the other data quality measures. As previously explained, most of the guidelines and standards were developed for the assessment of real-time data. However, for the purposes of this study, the guidelines and standards were adapted for the assessment of historical probe data.

A somewhat generic approach to collecting data for evaluation purposes was presented by Turner, Eisele, Benz & Douglas (1998). This data collection procedure is shown in Figure 2.1.

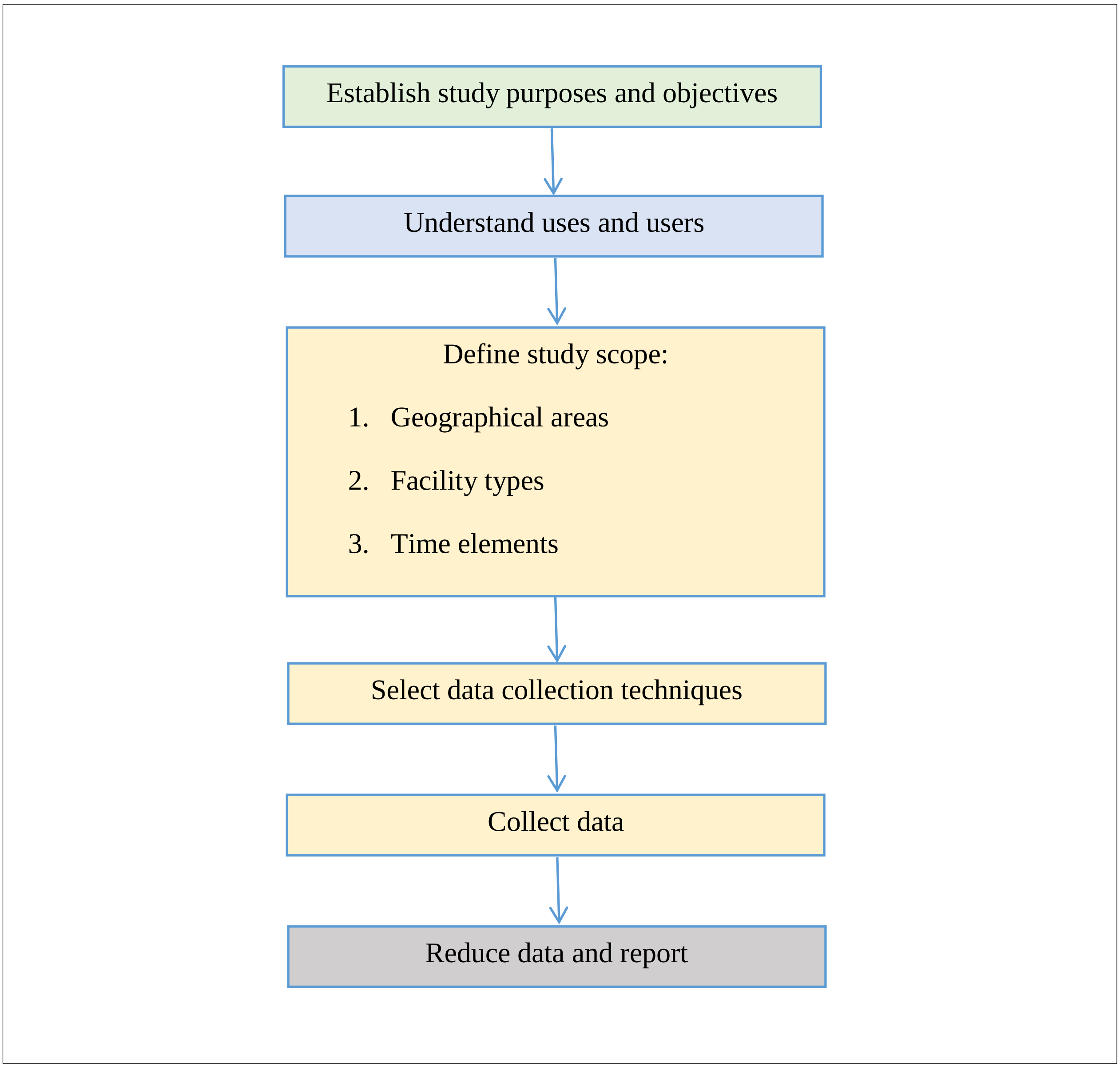


Figure 2.1: Data collection plan, adapted from Turner *et al.* (1998)

Essentially, the methodology is similar to what was presented by Battelle *et al.* (2004) and Turner *et al.* (2011). This methodology focuses on the data collection and not establishing the data quality measures or level of quality. However, it gives the user the tools to delineate the scope of the study. Therefore, it helps in developing an assessment plan and collecting benchmark data, with the emphasis on emerging technologies and ITS data.

2.3.1 DEVELOPING ASSESSMENT PLAN

An assessment plan describes how decisions about the basic parameters of a data quality assessment are made. The question that an assessment plan asks is "What information is being evaluated?" In its simplest form, an evaluation plan seeks to address issues pertaining to the evaluation measures, methods for collecting benchmark data, the type of facility, evaluation times, frequency and duration of the evaluation.

It is worthwhile to note that developing an assessment plan is a vital phase that defines the scope of a project. A carefully crafted evaluation plan clearly defines the context or scope. The advantage of a clearly defined plan often results in savings of resources, such as time and money. Turner *et al.* (2011) reported that there is no one single best solution for selecting evaluation parameters, as the parameters are dependent on the context and scope of the evaluation.

In addition, linking the parameters back to the intended application can help in defining the various parameters of the evaluation plan. Although certain elements should be addressed in the plan, it was observed that the order of the steps does not significantly affect the plan but it gives a logical development of the evaluation plan.

Table 2.3 shows a summary of the procedure adapted for developing a plan to evaluate the quality of historical probe data. In order to illustrate the process of defining and delineating the context and scope of the evaluation, there are evaluation parameters that should be taken into account. These are roadway segmentation, time intervals, links to evaluate, time periods to evaluate, the method for benchmark data collection, data quality measures, frequency and duration of evaluation and evaluation objectivity and transparency.

Table 2.3: Overview of a data quality assessment (Turner et al., 2011)

Element	**Points to consider**
Determine evaluation scenario(s) and available resources	What type(s) and form of traveller information is to be evaluated? What resources (i.e. budget, equipment, staff capabilities etc.) are available?
Identify relevant roadway segmentation	Typically based on the way(s) traveller information is currently provided Group together links shorter than 0.5 to 1 mile
Select time interval	Select a time interval basis for evaluation (5-minute interval is common) Consider longer time interval if sample sizes are inadequate
Select routes or links to evaluate	Stratify network by facility type (e.g. freeways, arterials) Stratify and select links/routes by level of expected variance
Select the time periods to evaluate	Collect evaluation data during peak traffic periods Attempt to include non-recurring events in evaluation data
Select benchmark data collection method	Re-identification is the preferred method for statistically valid evaluations Test vehicle could be used for initial screening or other limited circumstances
Select accuracy measures	Accuracy measures are selected based on the evaluation scenario Multiple measures tell a more complete story (i.e. no single best measure)
Determine frequency and duration of evaluations	Evaluation frequency and duration depends on several factors Increasing duration of evaluation (through identification) gives a better chance to evaluate TIS accuracy in non-recurring congestion
Evaluation objectivity and transparency	Evaluator is typically a third party hired by purchaser of TIS data Make entire evaluation method and results publicly available

2.3.1.1 ROADWAY SEGMENTATION

Roadway segmentation is dependent on the type of evaluation scenario and it is should be defined accordingly. The same road segmentation used to report the speed and travel time estimates should be used in collecting reference data and calculating quality measures. A typical road segment for an urban street is the link between two intersections whilst the point where the number of lanes changes is a good segmentation point for a freeway.

2.3.1.2 TIME INTERVAL

The evaluation scenario defines the time interval used to collect and compare the estimate data to the reference data. The main consideration when selecting a time interval is sample size. A time interval should be long enough to collect sufficient sample size. A sufficient sample size is essential in establishing "reliable" speed and travel time estimates.

In general, increasing the time interval increases the sample size per time interval and thereby enhances the statistical confidence. For real-time data, one would typically make the time interval as small as possible to identify any sudden change in traffic speeds. A 5-minute interval is a common time interval used in real-time traffic data assessments. For historical traffic data, longer time intervals such as 15-minute, 30-minute and 60-minute intervals are often used.

2.3.1.3 EVALUATION SEGMENTS

When selecting the segments to evaluate, it is important to include those with high variance in speeds or travel times. The literature reveals that sections with high variance in speeds are challenging for a data service provider or monitoring system to measure (Turner *et al.*, 2011). The authors added that if a system or data provider can meet the specified conditions on high variance sections, it stands to reason that the provider or system is likely meet the requirements on the low variance sections.

2.3.1.4 NETWORK STRATIFICATION

It is possible to use the facility type to stratify the road network. A typical stratification by facility type might result into two strata, freeways and arterials (Turner *et al.*, 2011). Furthermore, arterials can be categorised into urban and non-urban arterials (North Amercian Traffic Working Group, 2010). Typical stratification for freeways can be based on the average daily traffic (ADT), access point density and length of the links.

On arterial roads, traffic flow is interrupted on signalised and un-signalised intersections. On freeways, there is limited traffic flow interruption because access points are limited to ramps and interchanges. The scope of this study was limited to only freeway segments.

2.3.1.5 ASSESSMENT TIMES

The morning peak, midday, evening peak and overnight periods should be reflected in a data quality assessment (North Amercian Traffic Working Group, 2010). The exact times and duration for peak periods are specific to an area or region, hence the definition of these periods should be user defined.

Non-recurring events such as incidents and special events are usually the most challenging times for service providers to accurately measure. The traffic characteristics for a normal weekday and a day when a non-recurring event occurs are significantly different. Similarly, the traffic characteristics for a weekday are significantly different from typical weekend traffic. Unlike the case for historical data when non-recurring events have already occurred, it is difficult to predict when non-recurring events will occur for real-time evaluations.

2.3.1.6 REFERENCE DATA COLLECTION

Factors such as available funds, equipment and required sample size often determine the method of collecting reference data. Collecting the reference data is one of the most important tasks in a traffic data assessment. Due to the numerous considerations and level of detail required, reference data collection is presented in a separate section (see section 2.4).

2.3.1.7 DATA QUALITY MEASURES

The selection of data quality measures is typically based on the type of evaluation. It is widely accepted that there is no single best data quality measure for all possible scenarios (Turner *et al.*, 2011). Due to the numerous considerations that have to be taken into account, the data quality measures are presented in a separate section (see section 2.5).

2.3.1.8 DURATION OF ASSESSMENTS

A data quality evaluation is similar to a point estimate. The evaluation results indicate the data quality at a point in time (Turner *et al.*, 2011). It should be noted that the results could be different if the assessment is repeated. Often, the question that is asked is, how often should an evaluation be repeated and over how long of a period should an evaluation be conducted?

Data quality evaluations can be short, medium or long term. Short term assessments, such as a peak period or a day, is nearly a "point estimate" and can be unreliable. Medium term assessments, such as a week or a month, are more accurate due to large sample. Assessments conducted for periods such as one year or continuous evaluations, are long term. The length of the evaluation period is a function of cost and average sample size. In general, a longer evaluation period cost more and a large sample is collected.

Business requirements, changes in the network and changes in local travel patterns are probable reasons for re-evaluating data quality, which makes the issue of repeating an evaluation even more complex.

2.3.1.9 OBJECTIVITY IN ASSESSMENTS

It is important to maintain objectivity and transparency during the evaluation process, especially for performance-based data service contracts. Even the perception of impropriety or bias could call into question the evaluation results.

2.4 REFERENCE DATA COLLECTION AND REDUCTION

Reference data is defined as “any observations of speed or travel time over some spatial and temporal extent that is used as the benchmark data for evaluating a given estimation method” (National Cooperative Highways Research Program, 2009). Alternatively, Turner *et al.* (2011) defined reference data as the definitive set of speed and travel time considered to be the standard, against which estimates are compared. From the above definitions, it is clear that care must be taken when collecting, reducing and reporting reference data. The two basic methods for collecting benchmark data are re-identification and test vehicle. The two methods are discussed in the sections that follow.

2.4.1 TEST VEHICLE

Test vehicle methods have been used for benchmark data collection for several decades. Test vehicle methods involve the use of a specifically designated vehicle that is driven in the traffic stream for the sole purpose of collecting benchmark data (Turner *et al.*, 2011). Floating car runs are the most common test vehicle approach. A location aware device, such as a GPS device, automatically record time and location at periodic intervals.

2.4.2 RE-IDENTIFICATION

Re-identification involves identifying a unique vehicle attribute then re-identifying the same vehicle attribute at another point on the roadway at a later time (National Cooperative Highways Research Program, 2009). According to the authors, attributes commonly used for re-identification are Machine Access Control (MAC) addresses from Bluetooth-enabled devices, license plates or electronic toll tag identifiers. In addition, it is important to note that re-identification methods are based on vehicles that are already in the traffic stream for purposes other than evaluation.

Table 2.4 compares the characteristics of re-identification and test vehicle methods in terms of cost, experimental control, technology maturity, depth and breadth of the coverage.

Table 2.4: Characteristics of reference data collection methods (Turner *et al.*, 2011)

Characteristics	**Re-identification**	**Test vehicle**
Depth of coverage	Typically provides many more samples per travel time interval than test vehicle on high-volume roads and times	Typically provides very limited samples
Breadth of coverage	Breadth of coverage limited to where equipment can be positioned	Breadth of coverage much greater than re-identification, but at the cost of depth of coverage. Not limited by positioning of equipment
Technology maturity and adoption	Depends on technology used for re-identification. For example, Bluetooth is a mature communication standard but relatively new for traffic monitoring applications.	GPS and DMI technologies mature and widely used in transportation applications.
Overall experimental control	Must establish outlier filtering to remove vehicles that deviate from the designated link/route. Also must filter slower modes (buses, bikes, etc.) and vehicles with multiple passengers (e.g., transit vehicles). Should ensure adequate sensor spacing to minimize error.	Must establish driving protocol for test vehicles and enforce its standard application.
Cost ranges	Dependent on re-identification technology. Bluetooth devices can provide at least several hundred more data points per data collection dollar than test vehicle.	Cost is usually the constraining factor and typically limits test vehicle runs to a statistically inadequate sample size.

2.4.2.1 BLUETOOTH TECHNOLOGY

Bluetooth technology is an innovative and cost-effective means of collecting benchmark travel time. The use of Bluetooth technology has considerably improved the cost and practicality of using re-identification for accuracy evaluations (Turner *et al.*, 2011). According to Haghani *et al.* (2010), Bluetooth is a telecommunication industry specification that defines the manner in which mobile phones, computers, personal digital assistants, car radios and other devices can be easily interconnected using short-range wireless communications. From the previous work carried out using the Bluetooth sensors, it is evident that high-quality travel time data is obtained using this method (Haghani *et al.*, 2009, 2010; Young & Haghani, 2010; Aliari & Haghani, 2012; Friesen & McLeod, 2015).

The concept of acquiring traffic data using Bluetooth technology is illustrated in Figure 2.2. Bluetooth devices all have a unique identifier called the Machine Access Control (MAC) ID. Bluetooth sensors can identify the MAC addresses of Bluetooth enabled devices within a range of 1 to 100 m. Bluetooth sensors are placed at two ends of a road segment and the time the vehicle drives past the sensors is recorded. Travel time between the two points can be easily determined. In addition, vehicle speed can be determined if the distance between the sensors is known.

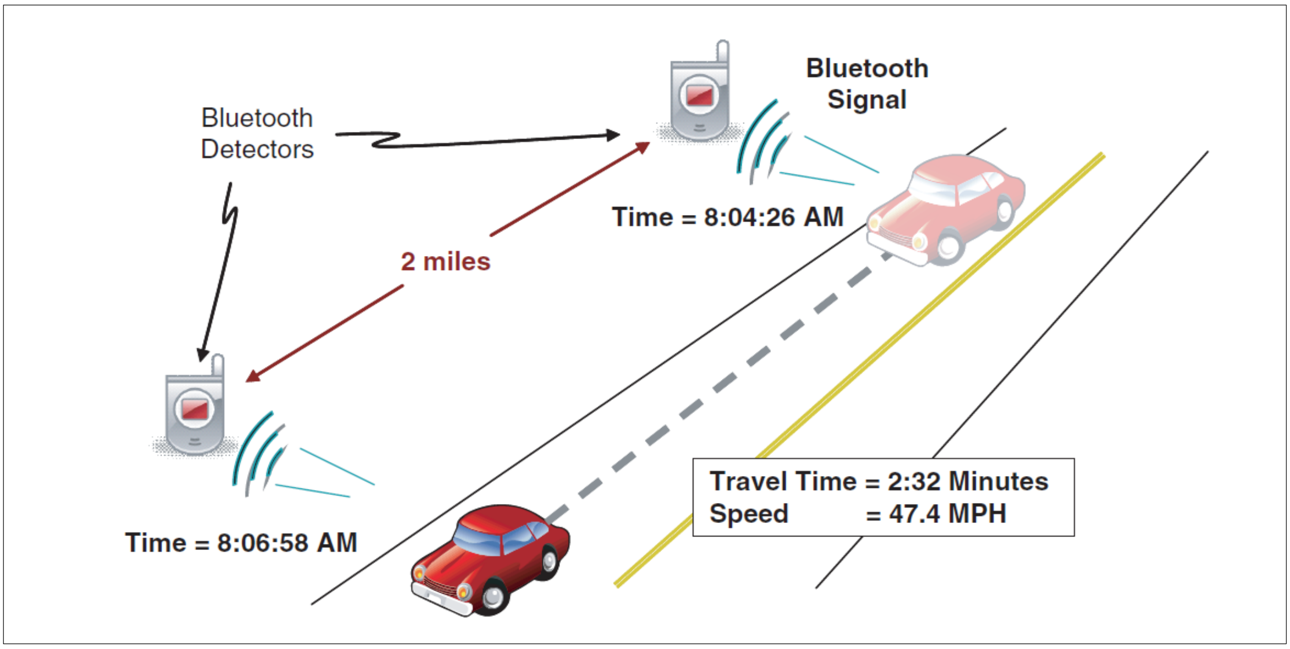


Figure 2.2: Bluetooth traffic monitoring system (Young, 2008; Haghani *et al.*, 2009, 2010)

2.4.2.2 ELECTRONIC TOLL TAG IDENTIFIERS

The toll tags used for electronic toll collection can be used to obtain travel time and speed data. A typical toll tag used in South Africa is shown in Figure 2.3. In South Africa, the toll tags are used for electronic toll collection on the Gauteng freeways (SANRAL, 2016). The toll tag is identified at a gantry, such as the one shown in Figure 2.4, and then re-identified at another gantry at a further point along the road. From these observations at two different gantry locations, the travel time and speed are determined. Similar to the Bluetooth benchmark data, the sample size for the reference data derived from toll tags depends on the market penetration of the toll tags. It is worthwhile to note that toll tags are primarily used for electronic toll collection. The toll tag reference data was used in a number of studies including Wright & Dahlgren (2001), Ferman, Blumenfeld & Dai (2005), Haas, Carter, Perry, Trombly, Bedsole & Margiotta (2009) and Texas Transportation Institute (2012).

Figure 2.3: SANRAL e-tag (SANRAL, 2016)

2.4.2.3 AUTOMATIC NUMBER PLATE RECOGNITION (ANPR)

Bluetooth and toll tags are used to determine travel times and speeds for a sample of the traffic stream. It is possible to obtain data for the entire traffic stream using Automatic Number Plate Recognition (ANPR). ANPR uses optical character recognition on an image to read the licence plates on vehicles (Kranthi, Pranathi & Srisaila, 2011). On the Gauteng freeways, ANPR is a component of Open Road Tolling (Robinson, 2014). The use of ANPR at a gantry on the Gauteng freeways is shown in Figure 2.4. Unlike the toll tags and MAC ID, the licence plate is linked to the personal information of the owner. To avoid privacy issues, ANPR data is anonymised i.e. the licence plate numbers are replaced by an anonymised numeric identifier that cannot be traced back to the owner. Kennedy, Cantrell, Varney, Czyzewski & Smith (2004) and Li (2008) used ANPR data for determining travel time and speeds in their respective studies.

Figure 2.4: A typical toll gantry on the Gauteng freeway network (Majangaza, 2015)

2.4.3 PROBE VEHICLE SAMPLE SIZES

Most data quality assessments focused on the available budgets in their approach to data collection rather than statistical significance (Schneider IV, Turner, Roth & Wikander, 2010). As a result, the evaluators provide no confidence interval to reinforce the accuracy of the reference data. A more comprehensive approach would be to give a confidence interval for both the reference data and the speed estimates from the service provider.

From the work done by Schneider IV *et al.* (2010), Quiroga & Bullock (1998) and National Cooperative Highways Research Program (2009), the minimum sample size is determined using t-statistic ($t_{\sigma/2}$) or z-statistic ($z_{\alpha/2}$), relative allowable error (θ), sample mean (μ) and standard deviation (σ), as shown in Equation 2.1.

$$N_{MIN} = \left(\frac{z_{\alpha/2} \times \frac{\sigma}{\mu}}{\theta} \right)^2 \tag{2.1}$$

Table 2.5 shows the various research projects conducted to investigate sample sizes. The research conducted on the sample sizes of reference data is summarised in Table 2.5. Presented in Table 2.5 are the findings or objectives achieved for the various research projects conducted to test, verify or define sample sizes. In the context of this study, the sample size refers to the number of probes used in deriving the speed estimates by the service provider.

Table 2.5: Literature on probe vehicle sample size (Schneider IV *et al.*, 2010).

Research Title	**Objective or Findings**	**References**
Probe sampling strategies for traffic monitoring systems based on wireless location technology	The study reviews probe vehicle sample size requirements	Fontaine, Yakkala & Smith (2007)
Travel time estimation using cell phone (TTECP) for highways and roadways	The study reviews accuracy results of various cell phone based probe vehicle tests and demonstrations	Wunnava, Yen, Babij, Zavaleta, Romero & Archilla (2007)
Penetration requirements for real-time traffic information from probe vehicles	The study concludes sample numbers required depend upon accuracy demand	Ferman & Blumenfeld (2006)
An analytical evaluation of a real-time traffic information system using probe vehicles	The study concludes 3 % penetration on freeways and 5 % penetration on surface streets is required	Ferman *et al.* (2005)
Factors affecting minimum number of probes required for reliable estimation of travel time	The study estimates travel time error for various probe sample sizes	Cetin, List & Zhou (2005)
Investigation of dynamic probe sample requirements for traffic condition monitoring	The study concludes sample size vary from 2 to 78 vehicles every 5 minutes based on various factors	Green, Fontaine & Smith (2004)
Extended floating car data: Traffic information and necessary penetration rates	The study concludes floating car penetration should be 2-20 % of traffic volume and varies for road type	Breitenberger, Gruber, Neuherz & Kates (2004)
Probe vehicle population and sample size for arterial speed estimation	This study proposes a methodology for reducing the bias in probe vehicle reports using on stratified sampling techniques	Long Cheu, Xie & Lee (2002)
Bias in probe-based arterial travel time estimates	Study proposes a method for reducing bias in probe vehicle reports	Hellinga & Fu (2002)

Research Title	**Objective or Findings**	**References**
Dynamic freeway travel time prediction with probe vehicle data	The study estimates travel time error based on probe vehicle simulation results	Chen & Chien (2001)
Determining the number of probe vehicles for freeway travel time by microscopic simulation	The study estimates required probe sample sizes based on simulation results	Chen & Chien (2000)
Travel time estimation on the San Francisco Bay Area network using cellular phones as probes	The study concludes that freeway link travel times could be estimated to within 10 % of their actual value if there is at least 5 % of wireless devices in the traffic stream	Ygnace, Drane, Yim & de Lacvivier (2000)
Assessing expected accuracy of probe vehicle travel time reports	The study examines the effect of sampling bias on the accuracy of the probe vehicle travel time estimates	Hellinga & Fu (1999)
The Grand Draw	The study concludes that approximately 10 % probe penetration of the total vehicle population is required for accuracy	Hoogenboom (1999)
Determination of the number of probe vehicles required for reliable travel time measurement in urban networks	The study concludes that approximately 5 % probe penetration of the total vehicle population is required for accuracy	Srinivasan & Jovanis (1996)
Probe vehicle sample sizes for real-time information: The Houston experience	Study concludes that probe samples between 1 and 6 vehicles per 5-min period is required for 95% confidence level and 10% error	Turner & Holdener (1995)
Vehicle as probes	The study concludes that approximately 4 % probe penetration of the total vehicle population is required for accuracy	Sanwal & Walrand (1995)

2.5 COMPUTING AND REPORTING DATA QUALITY MEASURES

A generally accepted definition of data quality was given by Strong, Lee & Wang (1997) as "fit for use by an information consumer". A further definition is given by English (1999), who described data quality as "fitness for all purposes in the enterprise processes that require it". Data quality is also defined as the "fitness of data for all purposes that require it" (Turner, 2002). It is evident that data quality is relative and it is possible that different consumers can have a different meaning of data quality. It is important to understand "all intended uses" of the data before conducting any evaluations. The data uses and users were discussed in section 2.2.

Whereas Turner *et al.* (2011) presented the methodology for evaluating the accuracy of traffic data, Battelle *et al.* (2004) took a step further and defined the procedures to evaluate the other quality measures of traffic data, namely completeness, validity, timeliness, coverage and accessibility. These quality measures are discussed in the sections that follow. It is important to note that these data quality measures constitute reasonable "categories".

Battelle *et al.* (2004) explained that the calculation of the data quality measures could vary from user to user. The authors further highlighted that it is conventional and even necessary to use a slightly different calculation procedure for different applications or users. A probable reason for allowing this deviation is due to the numerous transformation the original traffic data from the source undergoes or possible changes to the data when it moves from the field to the end user.

2.5.1 ACCURACY

Accuracy is "the measure or degree of agreement between a data value or set of values and a source assumed to be correct" (Turner, 2002; Battelle *et al.*, 2004). Accuracy measures the extent of closeness between the estimated parameter(s) and the reference value(s). Alternatively, accuracy refers to "a quality of that which is free of error" or an assessment of freedom from error (Cykana, Paul & Stern, n.d.; Shafer, 2002).

The selection of accuracy measures should be based on the evaluation scenario. In addition, the literature reveals that there is no single best accuracy measure for all possible scenarios (Kandarpa, Sangillo, Burgess & Toppen, 2010; Turner *et al.*, 2011). From experience, the commonly used measures for evaluating link speeds are *mean absolute percent error*, *signed percent error*, *root mean square error*, *average absolute error* and *average error (bias)*.

Accuracy can be tested statistically by determining if the speeds measured from one system, such as probe data from a service provider, are "equal" or within specified statistical parameters to the speeds measured on the system assumed to be correct, as later presented in Chapter 3.

2.5.2 COMPLETENESS

Completeness is defined as "the degree to which data values are present in the attributes that require them" (Cykana *et al.*, n.d.; Shafer, 2002). It is important to note that completeness can refer to both the temporal and spatial aspects of data quality, in the sense that completeness measures how much data is available compared to how much data should be available. Typically, completeness is described in terms of percentages. It is the available number of data values to the total number of expected data values, expressed as a percentage.

2.5.3 VALIDITY

Validity is a measure of the extent to which the "data is founded on an adequate system of classification and is rigorous enough to compel acceptance" (Cykana *et al.*, n.d.; Shafer, 2002). It measures the degree to which the data falls within the range of acceptable thresholds. Validity criteria often range from a simple rule to complex ones, be based on established theory, scientific facts or can be a "rule of thumbs". As a result, it is not surprising that validation criteria for different applications are often very different. Validity is typically expressed as the percentage of data passing the validity criteria.

2.5.4 TIMELINESS

Timeliness is the measure of the degree to which data is provided at the specified time. Timeliness is particularly useful when dealing with real-time data since it gives an indication of how quickly information is disseminated, for example, from the road sensors to the Traffic Management Centre (TMC) or TMC to end user. Timeliness gives an indication in the delay of obtaining data. Typically, timeliness is expressed as a percentage of the data received on time to the total datasets received. It is worthwhile to note that it is difficult to get insightful information from timeliness in cases where historical archived data is used in the analysis.

2.5.5 COVERAGE

Coverage is the quantity that gives an insight to the proportion of the system being measured. It is defined as the degree to which data values in a sample represent the population (Turner, 2002; Battelle *et al.*, 2004). According to the authors, this definition of coverage leaves several quantities open for interpretation. For example, what is a sample that accurately represents the whole population? Even with the highlighted issues, it must be emphasised that coverage is an important quality measure that helps in explaining the other data quality measures and possibly explain any variation in these measures.

2.5.6 ACCESSIBILITY

To obtain meaningful information from the data collected, a considerable amount of effort goes into manipulating and processing the data. Accessibility is the quantity that describes the easiness of using the data. Accessibility is the easiness with which data is obtained and processed by the data consumer. Alternatively, accessibility is defined as "the relative ease with which data can be retrieved and manipulated by data consumers to meet their needs" (Turner, 2002; Battelle *et al.*, 2004). Accessibility is also known as usability and can be expressed as a qualitative or quantitative quality measure.

2.5.7 ESTABLISH ACCEPTABLE DATA QUALITY TARGETS

In order to compare the estimate and reference values of a traffic attribute and make a call on the result, threshold values for the data quality measures have to be set. As established earlier, the threshold values for different applications or data users are likely to be different. The set targets should reflect the acceptable quality based on the data user's needs and applications. Battelle *et al.* (2004) stated that the data quality measures falling outside the thresholds reflect one of two things. First and the obvious one, the data could be unacceptable for the intended use and secondly, the data ought to be used with caution.

2.5.8 CALCULATE DATA QUALITY MEASURES

After establishing the acceptable data quality targets, the data quality measures can be computed. The procedures for determining the quality measures were presented in section 2.5.1 to section 2.5.6.

2.5.9 IDENTIFY DATA QUALITY DEFICIENCIES

After comparing the estimates to the reference data and computing the different data quality measures, it is important to understand the reasons why some datasets did not meet the specified criteria. This process of identifying data quality deficiencies is often overlooked. It should be kept in mind that if traffic data does not meet the specified criteria for one application, it does not necessarily imply that the data is worthless. Perhaps, it suggests the data has to be used with extra caution or for a different application. More information about the data is usually obtained from identifying and querying data quality deficiencies.

2.5.10 ASSIGN RESPONSIBILITY AND AUTOMATE REPORTING

This step is usually a key area in organisations such as transportation agencies that conduct long-term projects. To ensure that good quality data is obtained, it might be necessary to appoint individuals who focus solely on validating and enhancing the data quality and automated reporting.

2.5.11 PERFORM PERIODIC ASSESSMENT

The framework for assessing the traffic data quality is cyclical. As a result, it is important to complete the cycle by re-assessing the data user and quality targets for their applications on a regular basis.

2.6 SELECTED VALIDATION PROJECTS

This section presents a number of case studies to illustrate the important aspects of the data quality evaluations conducted in the past. The evaluations described here used real-time or historical traffic data from selected data service providers such as Inrix, Traffic.com and TomTom. The various reference data used in these studies are also explored. The reference data used include Bluetooth, toll tag and floating car. The three evaluations discussed in this section used error metrics and methodologies consistent with best practices recommended in Turner *et al*. 2011 and Battelle *et al.* 2004.

2.6.1 I-95 CORRIDOR COALITION PROJECT EVALUATION

2.6.1.1 PROJECT SUMMARY

The I-95 Corridor Coalition is the world's largest ongoing data validation project. Several partners combined to collect and assess real-time travel time and speed data. Data collection and validation was conducted on approximately 1500 miles (2400 km) of freeways and 1000 miles (1600 km) of arterials in six USA states, namely New Jersey, Pennsylvania, Delaware, Maryland, Virginia and North Carolina (Haghani *et al.*, 2009, 2010; Haghani, Hamedi & Parvan, 2013).

2.6.1.2 DATA

Inrix data was evaluated. The data was primarily derived from GPS-equipped fleet vehicles and in some situations, complemented by sensor and detector-based data. Benchmark data used for this project was collected using Bluetooth readers. Bluetooth readers identify the MAC addresses of Bluetooth enabled devices and matched observations were then used to compute the travel times and speeds.

2.6.1.3 METHODOLOGY

The evaluation of Inrix data was performed by the University of Maryland. The reference speeds from the Bluetooth readers were compared to the speed estimates from the service provider. Equivalent Inrix speed estimates were based on the start and end time of an observed Bluetooth match. The speeds were grouped in four speed bins, 0-30 mph, 30-45 mph, 45-60 mph and 60+ mph. The contract requirements were based on two error metrics, average absolute speed error (AASE) less than 10 mph (16 km/h) and speed error bias (SEB) less than 5 mph (8 km/h).

2.6.1.4 RESULTS

The study showed that the Inrix travel time and speed estimates across the system and by individual state generally meet the specified validation criteria of the contract i.e. AASE less than 10 mph and SEB less than 5 mph (Haghani *et al.*, 2009). Another interesting finding of the project was that the travel time and speed estimates from the service provider improved with an increase in speed.

2.6.2 EVALUATION OF INRIX AND TRAFFIC.COM

2.6.2.1 PROJECT SUMMARY

The evaluation of traffic data quality from two data service providers, Inrix and Traffic.com, was conducted in Philadelphia, Providence and Washington DC. The study was designed to evaluate the accuracy and coverage of real-time traffic data (Inrix, 2006; National Cooperative Highways Research Program, 2009).

2.6.2.2 DATA

Floating car data from 141 car runs was used as the benchmark for this study. Data from Traffic.com and Inrix data was derived from loop detectors and GPS-equipped fleet vehicles.

2.6.2.3 METHODOLOGY

The study was carried out by an independent researcher, Frost and Sullivan. For this particular study, the data quality measures used to evaluate the service providers were accuracy and coverage. The error metric used was the root mean square error (RMSE).

2.6.2.4 RESULTS

It was shown that both Inrix and Traffic.com provided strong and comparable levels of accuracy in reporting travel time on the analysed routes (Inrix, 2006). However, Inrix had a slightly better evaluation due to the significantly broader coverage.

2.6.3 EVALUATION OF TOMTOM DATA

2.6.3.1 PROJECT SUMMARY

The previous two projects discussed in this section focused on real-time data. The evaluation of TomTom historical traffic data was conducted by the Texas Transportation Institute in Houston, Texas (Texas Transportation Institute, 2012).

2.6.3.2 DATA

Houston TranStar provided the reference data used in this study. The two reference data sources were toll tag-based traffic monitoring system and Bluetooth-based arterial street monitoring system. TomTom data was derived from several sources such as fleet GPS, personal navigation devices, road sensors, third party data and GSM probes.

2.6.3.3 METHODOLOGY

The evaluation of TomTom historical traffic data was carried out on 11 directional segments. Although historical traffic data was evaluated, the methodology for this study was consistent with best practices and the I-95 Corridor Coalition project. The error metric used was the AASE.

2.6.3.4 RESULTS

The results demonstrated how closely TomTom estimated the traffic speeds. For the heavily congested freeway, the AASE was less than 5 mph for most of the segments, across all the speed bins. A slightly higher AASE of 8 mph was obtained for nearly all the major arterial road sections that were investigated in the study.

2.7 LESSONS FROM PREVIOUS PROJECTS

Although there is currently no agreement on the quality levels and best methods for determining the data quality measures, important lessons were noted from the case studies that were reviewed in this section (National Cooperative Highways Research Program, 2009):

1. Although a number of traffic data evaluations use point estimates of travel time and speed to measure accuracy, it was shown that a single reference measurement may not be sufficient. This was observed in cases where the variation in speed is high. Therefore, speed and travel time should be viewed as a distribution.

2. Link level analysis may not be sufficient to measure accuracy at a route level. Route travel time is more meaningful to users compared to link travel time. Therefore, route level analysis may be needed.
3. Measuring accuracy on signalised arterials is more challenging due to the interrupted nature of traffic flow on arterials. This area requires special attention.
4. Data from one source could be complemented by other data sources to improve reliability and address coverage issues. However, care must be taken when fusing data from different sources.
5. More resources should be directed towards measuring accuracy during transition periods, i.e. from free flow to congested traffic, as these periods are typically challenging for a service provider or monitoring system to accurately measure. In addition, links with high variance in speeds were also found to be challenging for a data service provider or monitoring system to measure.
6. The level of variance typically dictates the data collection methods. Classifying links by the level of variance in speed and travel time helps to determine a suitable data collection method as well as sample size issues. For example, freeways, urban and non-urban arterials with high variance in speeds should be sampled using re-identification methods. Test vehicle or re-identification methods are suitable for non-urban arterials and freeways with low variance in speeds.
7. The road network should be stratified by facility type into freeways, urban arterials and non-urban arterials.
8. Where the identification link length differs from the link length of the estimates speeds, the differences between the two lengths should be clearly reported.
9. Adjustments to the travel time should be made (or use average speed) to account for the differences in length between the reference link and the service provider link.

2.8 CONCLUSION

In this chapter, the different methodologies for evaluating the quality of traffic data were discussed. Because the guidelines and standards were developed for real-time data, ways in which these guidelines and standards were adapted for historical probe data were explored. The outcomes and main findings from the selected evaluation projects were discussed. The chapter ended with a discussion on the lessons learnt from the previous data evaluation projects.

CHAPTER 3 : RESEARCH DESIGN AND METHODOLOGY

3.1 INTRODUCTION

The study focuses on the evaluation of the quality of TomTom historical data. This chapter introduces the research design and methodology used to answer the research question. Details and justification for the various aspects of the study such as the type of road links, time intervals, geographical area and data are discussed. Included in this chapter is the discussion on the data processing and reduction procedure. The data analysis conducted to obtain results is discussed in the final section of the chapter.

3.2 RESEARCH DESIGN

The goal of the research project is to assess the quality of TomTom historical data. The traffic data attributes of interest are travel time and speed. These two attributes are expressed as numerical data. For this reason, a quantitative approach was adopted for this study. More specifically, travel times and speeds derived from historical probe data were compared to the benchmark data.

The quantitative research method used was a field experiment. A field experiment is distinguished from a laboratory experiment by the fact that they occur in a natural setting rather than an artificial setting. The quantitative research method used allows the assessment of the quality measures of TomTom historical data.

3.3 RESEARCH METHODOLOGY

The field experiment adopted for this study was designed to investigate the quality measures of probe data (Bailey, 2008). The natural setting of field experiments increases generalisability of results and often decreases the possibility of laboratory errors such as experimental errors (Mouton, 2012).

The natural experiment aspect of this method has one obvious limitation. The level of control is often very low. For example, the fixed gantry positions meant that the researcher had no control over the length of the freeway segments.

The main sources of error often arise because of less control of extraneous variables which, in turn, leads to weaker causal claims (Mouton, 2012). Common sources of error are measurement error, sampling error and sample size. Issues regarding sampling and sample sizes are discussed in section 3.4.3.2. The issues pertaining to measurement errors with respect to the reference data are addressed in section 3.5.1.

Before going into detail on the data collection, the critical aspects of the data collection plan adopted for this study are discussed. In Chapter 2, the reader was introduced to the data collection plan. Figure 2.1 shows the data collection plan, modified from Turner *et al.* (1998).

3.3.1 STUDY GOALS AND OBJECTIVES

The study purposes and objectives establish the need for data. As such, the objectives should be defined as the first step in any data collection exercise. The purpose of this study is to assess the quality measures of TomTom historical data. For this study, data quality measures refer to accuracy, validity, completeness, coverage and accessibility of TomTom historical probe data. A detailed discussion of study goals and objectives was presented in Chapter 1.

3.3.2 DATA USES AND USERS

When identifying and understanding the uses and users of the data, it is important to ask who is interested in using the data and the typical range of applications for which the data can be used. Transportation agencies are particularly interested in knowing if it is feasible to use probe data to supplement existing traffic data from conventional data sources. Better still, transportation agencies are interested in knowing if probe data can be used as the principal source of traffic information for mainstream applications in engineering, planning and operations. An in-depth discussion on the uses and users of traffic data was presented in Chapter 2.

3.3.3 SCOPE

In delineating the boundaries of a data collection exercise, there are key questions that should be asked (Turner *et al.*, 1998). These questions are:

1. Where is the data collected?
2. On which facilities is the traffic data collected?
3. When is the data collected?

To answer these questions, the aspects that should be investigated are geographical areas, facility types and time elements, as shown in Figure 2.1 (Turner *et al.*, 1998). An in-depth discussion of these elements and the rationale for the choices of these aspects is discussed in the sections that follow.

3.3.3.1 GEOGRAPHICAL AREAS AND TYPE OF FACILITY

The Gauteng Freeway Improvement Project (GFIP) was designed as the commencement of Open Road Tolling in South Africa. Although the project was primarily for tolling reasons, valuable high quality and high-density data of the entire traffic stream is collected on the ORT system. The ORT system is mature, extensively tested and has been operational since 2013. ORT data has already been calibrated (Robinson, 2014, 2016). Therefore, ORT system is an ideal source of reference data. Figure 3.1 shows the Gauteng freeway network and the locations of the existing gantries on these freeways. The gantry positions were not meant to cover each freeway segments between access interchanges but locations where most traffic can be matched for toll collection purposes.

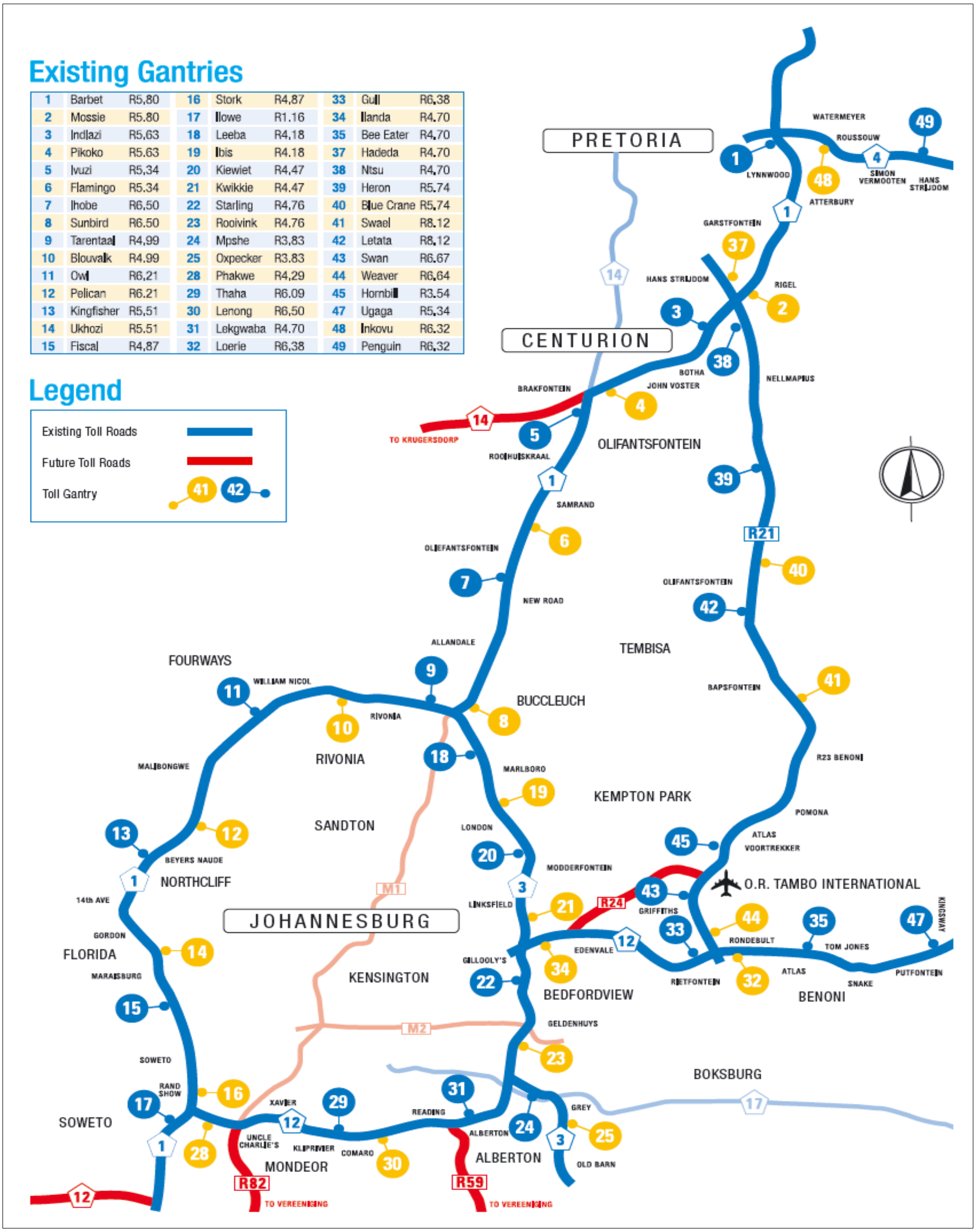


Figure 3.1: Location of gantries on the Gauteng freeway network (Robinson, 2016)

Spatial sampling is the process of selecting a subset of the spatial elements in a geographical area. In general, spatial sampling for a data quality assessment reflects the type of roads under investigation. This study was limited to freeways on which open road tolling is implemented as the facility type and Gauteng as the geographical area.

3.3.3.2 TIME INTERVAL

Another critical aspect of a data quality assessment is temporal sampling. Temporal sampling relates to the selection of the time elements of the study. These elements include analysis period, peak hours, off-peak hours and length of the data collection.

Studies on the assessment of traffic data quality typically use a 5-minute time interval (Haghani *et al.*, 2009). A 5-minute time interval ensures that any variations of speeds with time are easily observed. Ideally, a time interval must be small enough to observe any fluctuations in average speeds. In addition, the time interval is proportional to the average sample size of the probe data, the number of data points and the computational workload in processing the data.

TomTom historical data can be obtained in 1 hour, 30 minute and 15-minute intervals. It is possible to use other time intervals. The lowest available time interval is 15 minutes. If a 15-minute time interval is used, four average speeds are obtained in an hour, whereas only one average speed is obtained using a 1-hour time interval. Using a 1-hour interval could possibly lead to low speeds and high speeds averaging. The level of aggregation for a 1-hour interval is greater than aggregation for a 15-minute interval. In general, the smaller the time interval, the less the likelihood of extreme traffic conditions cancelling out. Time interval is a trade-off between accuracy and computational workload. It was decided to enhance the accuracy of the study. Therefore, a 15-minute time interval was selected. A 15-minute time interval is also a basic time unit used in the Highway Capacity Manual (National Research Council, 2010). In addition, a 1-hour interval was also used so that the speeds derived from both time intervals can be compared.

For any traffic data assessment, the morning peak, afternoon peak, evening peak and off-peak times should be included. However, it must be mentioned that not all off-peak times result in valuable traffic information, for example, during the night. This is because traffic is typically moving at free-flow conditions during the night. Additionally, the possibility of obtaining sufficient probes during the night is low. In light of the above discussion, the limits of the analysis period were set from 05:00 to 20:00.

The traffic characteristics for weekdays and weekends are different. For this reason, only weekdays (Monday to Friday) were considered in this study. The evaluation period for this study was February 2015. The month of February was selected because there were no exceptional days such as holidays.

3.3.4 EXPERIMENTAL SET-UP

The experimental set-up, shown in Figure 3.2, illustrates the research context used for data collection. Figure 3.2 shows a typical freeway segment between two gantries. For simplicity, no off-ramps, on-ramps or access interchanges are shown in Figure 3.2.

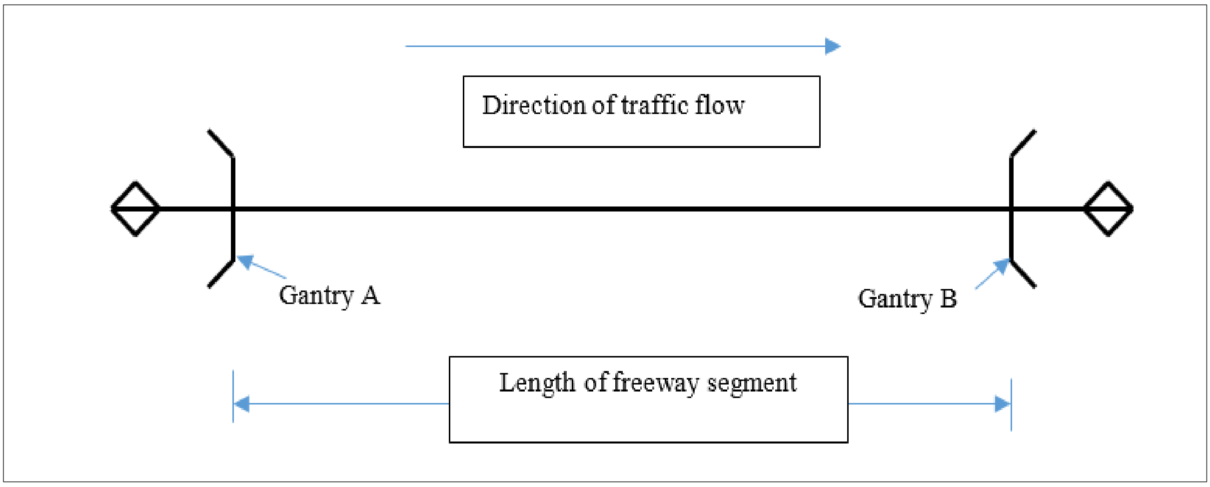


Figure 3.2: Experimental set-up

As the traffic traverses from left to right, vehicles cross gantry A and then gantry B. The number plate is the unique attribute that is identified at gantry A and re-identified at gantry B. The time stamps at both gantries are captured and recorded on the ORT system. The time stamps are accurate to the nearest second, therefore highly reliable and accurate traffic data is obtained. From the time stamps and length of the freeway segments, travel times and speeds for that particular freeway segment were then computed. This is the reference data, which is the control of the experiment.

The notation for naming the freeway segments uses the freeway name, direction and the two gantries. A freeway segment between gantry 1010 and gantry 1012 on the N1 in the southbound (SB) direction, with gantry 1010 as the first gantry crossed by the vehicle and gantry 1012 as the second gantry crossed, is described using this notation, N1_SB_1010_1012.

3.4 DATA

This section introduces the reader to the data used in the study. A discussion on TomTom historical data and ORT data is presented. The issues on sampling and sample sizes are also addressed.

3.4.1 TOMTOM DATA

TomTom data is derived from various sources. These sources include live personal navigation devices, in-dash GPS, business solutions, third party data, GSM probes, road sensors and incident data, as depicted in Figure 3.3 (Dannehy & Krootjes, 2013; Ressler, Thomas & Dannehy, 2013).

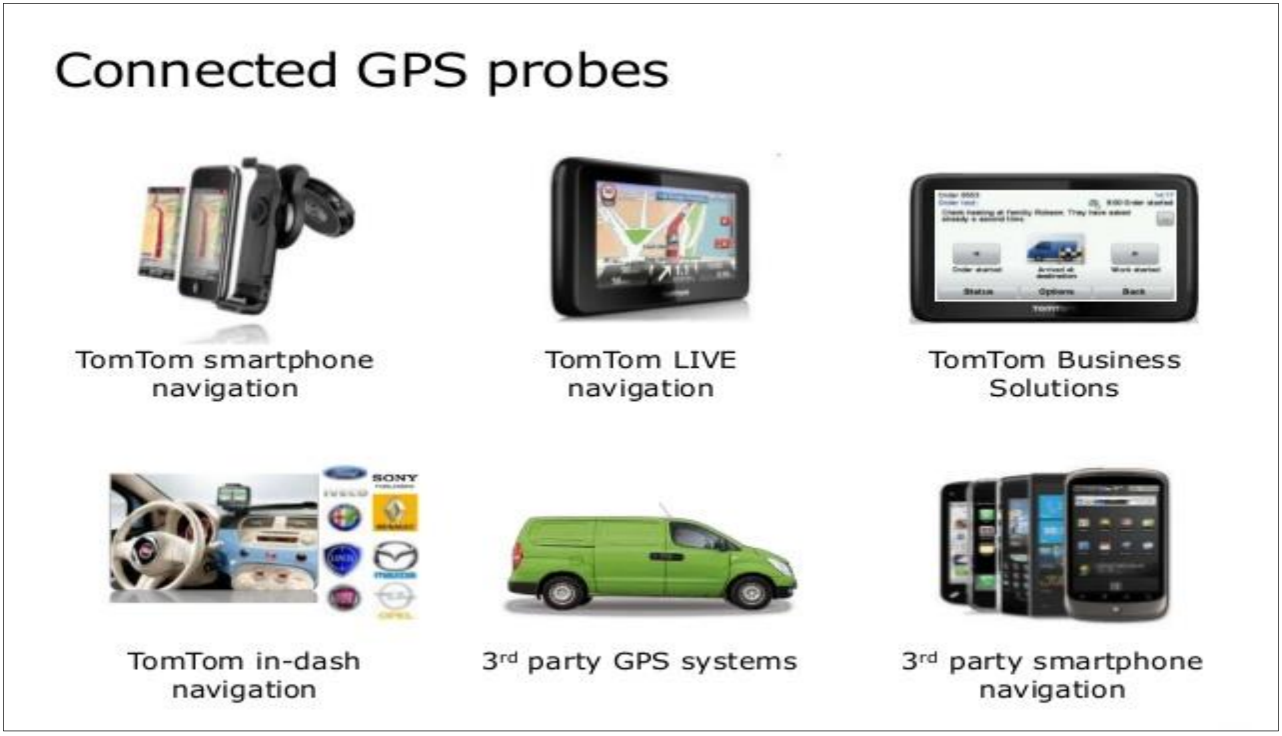


Figure 3.3: TomTom data sources (Ressler *et al.*, 2013)

TomTom historical data was downloaded from the TrafficStats portal on the TomTom website (Dannehy & Krootjes, 2013; TomTom, 2016). The data is already processed and no further data processing or manipulation was required.

Checks were carried out when inputting the coordinates of the start and end locations of the freeway segments and entering dates and time intervals on the portal. In this way, the researcher confirmed that the input parameters for the reference data were the same as the inputs for the probe data.

3.4.2 ORT DATA

ORT data was requested from South African National Road Agency Limited (SANRAL). ORT provides high density and high-quality traffic data for the population of the traffic stream. ORT is a state-of-the-art system that uses Automatic Number Plate Recognition (ANPR) to capture useful traffic data. The system has been extensively calibrated to give accurate information (Robinson, 2014, 2016). ORT data is currently used in transportation planning projects, such as network modelling, origin-destination studies and trip generation. ORT data was used as the reference data for this study.

ORT data is secondary quantitative data, as it was not gathered by the researcher. The limitations of using secondary data are acknowledged. In order to gain control and understanding of the data processing, raw data was requested from SANRAL. Quality assurance checks were conducted to ensure that accurate speeds were derived from the raw data. In section 3.5, the data processing and reduction plan is explained in detail and the quality checks implemented are discussed.

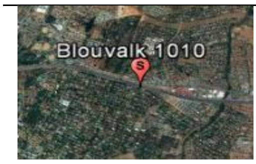
3.4.3 SELECTION OF FREEWAY SEGMENTS

It was not practical to investigate all the freeway segments due to resource constraints. Considerably more effort is required to collect, manipulate and process data for the population compared to a representative sample. For this reason, a sample of the freeway segments was investigated. The speed profiles for the ORT data were examined to analyse the different patterns and characteristics of traffic on the different freeways. As expected from the speed profiles, four distinct patterns were observed on the different freeway segments, namely:

i. Morning peak;
ii. Evening peak;
iii. Morning and evening peak; and
iv. No pronounced peak.

During the process of selecting the sample of freeway segments, it was decided to investigate freeway segments that show the aforementioned patterns. Six freeway segments were selected using an arbitrary sampling technique. The motivation for choosing arbitrary sampling was that it was required to select freeway segments showing all the four aforementioned patterns. Selecting the freeway segments was based on judgement that guaranteed that segments with the desired patterns were chosen. The six freeway segments are on the N1 Western Bypass Highway, N1 Ben Schoeman Highway and R21 Albertina Sisulu Highway, as shown in Figure 3.4 to Figure 3.9. The locations of the freeway segments relative to each other were also shown in the Gauteng freeway network map, in Figure 3.1. Table 3.1 describes the features used in Figure 3.4 to Figure 3.9.

Table 3.1: Legend describing the features used on the freeway segment location figures

Key	Description
	A balloon marker is the exact position of a toll gantry.
	This represents the TomTom link segment. The different colours represent the different speed regimes during a given time interval.
	The donut marker represents an access interchange.

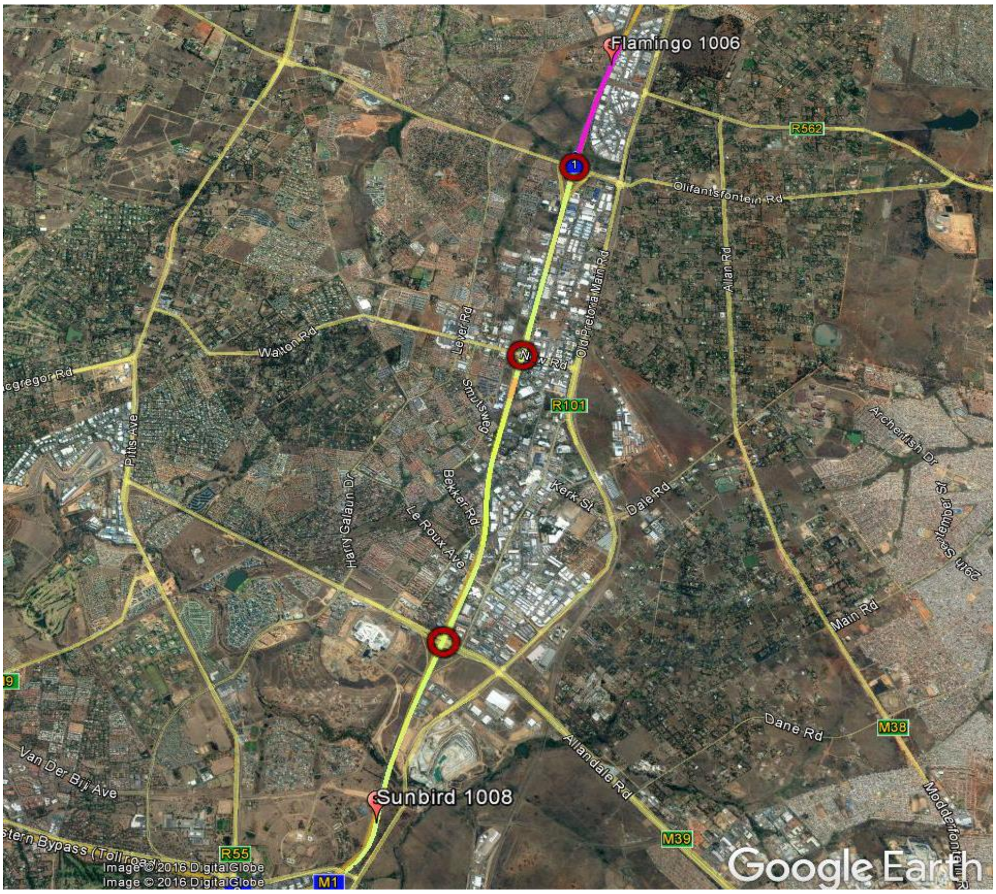


Figure 3.4: N1 Southbound (Ben Schoeman) between Flamingo (1006) and Sunbird (1008)

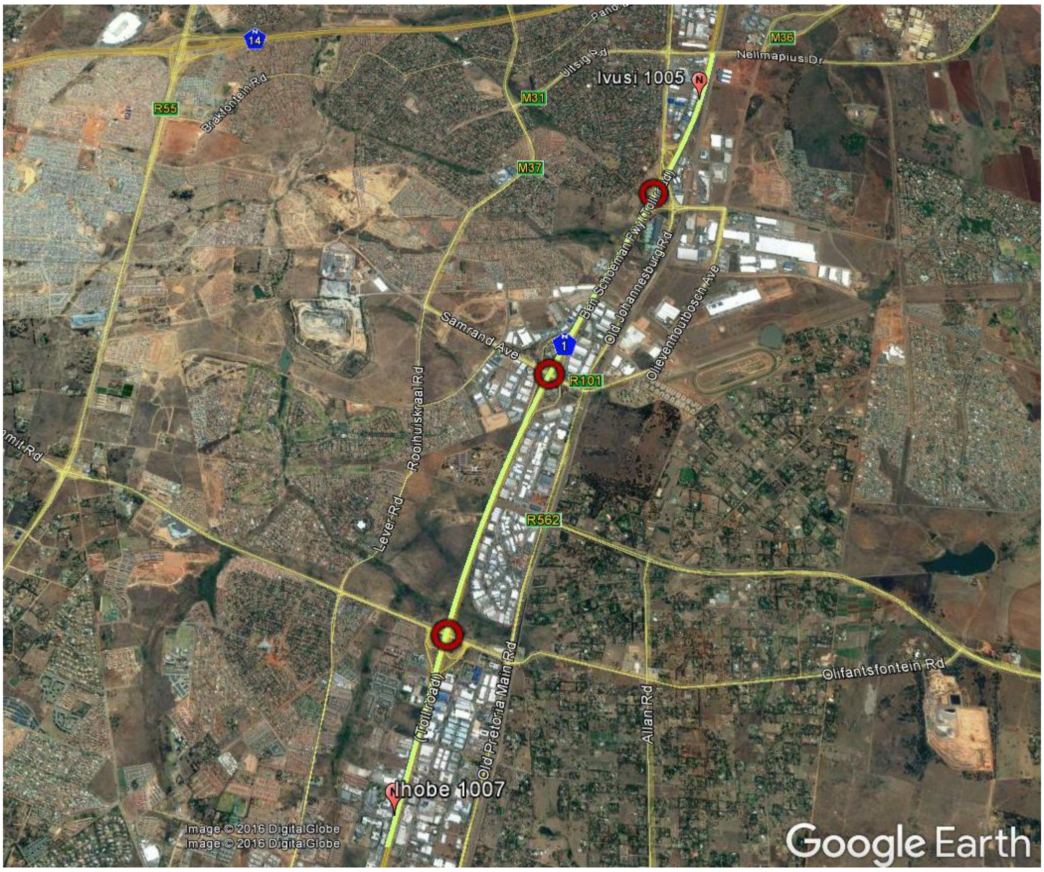


Figure 3.5: N1 Northbound (Ben Schoeman) between Ihobe (1007) and Ivusi (1005)

Figure 3.6: N1 Southbound (Western Bypass) between Blouvalk (1010) and Pelican (1012)

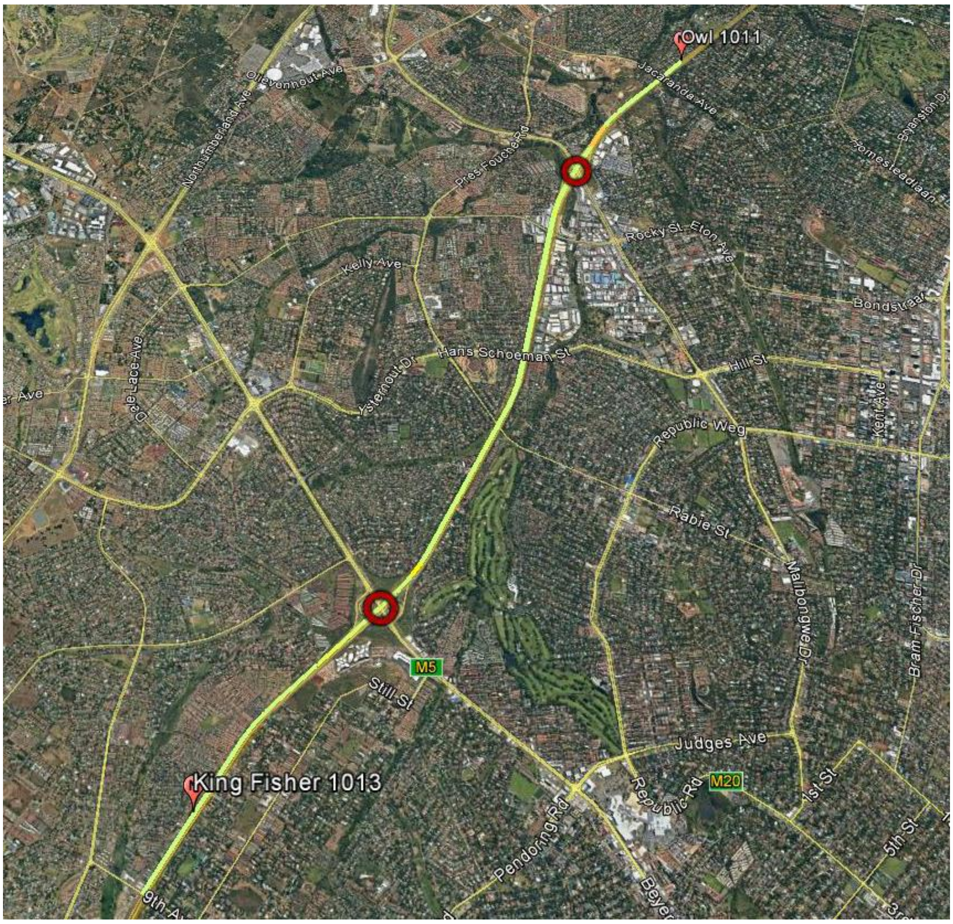


Figure 3.7: N1 Northbound (Western Bypass) between King Fisher (1013) and Owl (1011)

Figure 3.8: R21 Southbound (Albertina Sisulu) between Bluecrane (1040) and Swael (1041)

Figure 3.9: R21 Northbound (Albertina Sisulu) between Letata (1042) and Heron (1039)

3.4.3.1 LENGTH OF THE FREEWAY SEGMENTS

The lengths of the freeway segments are shown in Table 3.2. The literature revealed that the benchmark link lengths and probe data link lengths are not equal in most cases and the differences should be clearly indicated (Turner *et al.*, 2011).

Table 3.2: Freeway segment lengths for the different data sources

Freeway segment	**ORT link**	**Google Earth link**	**TomTom link**	**% difference**
N1_SB_1006_1008	11.3	11.3	12.4 (38)	-9.7
N1_NB_1007_1005	8.8	8.8	9.7 (28)	-10.9
N1_SB_1010_1012	11.0	11.0	11.4 (70)	-3.4
N1_NB_1013_1011	9.5	9.5	10.2 (43)	-8.2
R21_SB_1040_1041	10.7	10.7	13.5 (34)	-26.3
R21_NB_1042_1039	8.3	8.3	9.7 (13)	-16.4

Note: All the link lengths are in km. The number in brackets is the number of TomTom sub-segments for each freeway segment.

The lengths of the freeway segments were requested from SANRAL whilst probe data link lengths were obtained from the TomTom data files. To enhance the accuracy and reduce the measurement errors cited in section 3.3, the lengths of the freeway segments were also measured using Google Earth as an additional check.

A general rule for the differences in length recommended by Turner *et al.* (2011) is 10 %. The percentage differences show that the TomTom link lengths were greater than the ORT link lengths for all the six freeway segments. The percentage differences for the two R21 sections show that the TomTom links were longer than the ORT links by 26.3 % and 16.4 % in the southbound and northbound directions, respectively.

3.4.3.2 POSSIBLE EXPLANATIONS FOR THE UNEQUAL LINK LENGTHS

To establish the TomTom link length, the start and end coordinates of the freeway segment were entered on the TrafficStats portal. The start and end coordinates of a freeway segments are the coordinates of the two toll gantries that make up the freeway segment. The balloon markers, as shown in Figure 3.10, represent the toll gantries. It was noted that the start and end coordinates for the freeway segments automatically shifted to new locations, as indicated in Figure 3.10. As a result, the lengths of the ORT link and TomTom links were unequal.

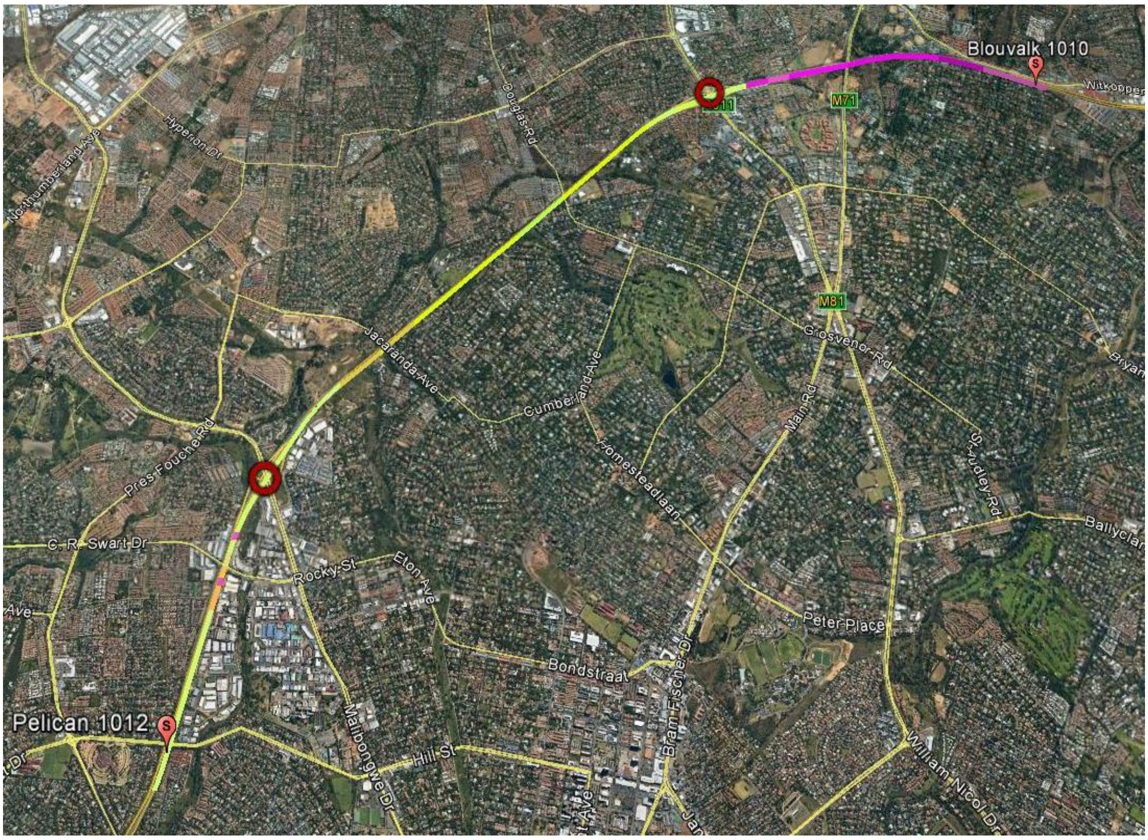


Figure 3.10: Illustration of the TomTom and ORT links

Considering only the TomTom link, it appears as if the link one continuous link. However, this is not the case. The TomTom link consists of several sub-segments. A sub-segment is a section of the road that has constant geometric characteristics and uniform speeds. The TomTom link in Figure 3.10 has different colours, which represents the different speed regimes during a given time interval. This also suggests that the sections with different speeds belong to different sub-segments. The TomTom link shown in Figure 3.10 consists of 70 sub-segments.

The coordinates of the start and end of the freeway segment shifted because the coordinates did not coincide with any of the nodes on the TomTom's link segmentation. The point moved to the nearest node of that sub-segment. The points shifted out in each case so that the specified gantry coordinates were included in the resultant TomTom link. As a result, the TomTom links were greater than the ORT links. TomTom's network segmentation depends on the speeds and geometric characteristics of the freeway section. This could explain why the differences in the ORT and TomTom links were not uniform.

3.4.4 SAMPLE SIZE AND SAMPLING ISSUES

A sample represents any subset of the elements of a population (Montgomery & Runger, 2007). It is important that a sample is representative of the characteristics of the population. In this study, sample sizes were considered for the following:

3.4.4.1 NUMBER OF REFERENCE OBSERVATIONS

The travel times and speeds for all the vehicles traversing along a given freeway segment were determined. Some matched observations that were deemed unacceptable were filtered out so that "reliable" reference data can be established. A discussion on data filtering is presented in section 3.5.1.4. Because of the observations were filtered, the ORT observations were also a sample, albeit a highly representative sample since only a few unacceptable observations were discareded. Due to the sample size of reference observations, as later shown in Table 3.5, there was no need of establishing the minimum sample size for the ORT data.

3.4.4.2 NUMBER OF PROBES FROM THE SERVICE PROVIDER

TomTom historical data gives an indication of the average sample size for custom travel time analysis. This is the number of probes considered in deriving the average TomTom speed estimates. Probes are vehicles that anonymously provide traffic data, for example, vehicles equipped with personal navigation devices (PND) or in-dash GPS, smartphones and commercial vehicles with GPS devices.

The determination of the minimum sample sizes for TomTom historical data is discussed in the section that follows. The researcher has no influence on the proportion of the traffic stream used by the service provider to derive the speed estimates. The individual vehicles that make up the sample of the traffic population and distribution of this sample in terms of heavy vehicles, passenger vehicles and motorcycles used by the service provider were unknown to the researcher.

3.4.4.3 MINIMUM SAMPLE SIZE

From the literature review, it was stated that the z-statistic or t-statistic, relative allowable error, sample mean, standard deviation are used to establish the minimum sample size. Ideally, the minimum sample size should be based on the standard deviation and mean speed in each 15-minute interval. However, this results in many computations. A simplified and conservative approach that uses the maximum standard deviation across all the intervals and average speed for each segment was used to determine the minimum sample size.

The standard deviation of the traffic proportion used by TomTom is unknown. The speed estimates from the service provider were derived from a proportion of the traffic stream. Assuming that the TomTom speed estimates were derived from a true representative sample of the entire traffic stream, the standard deviation of the traffic sample used by TomTom is approximately equal to that of the entire population. The standard deviation for the entire traffic stream was obtained from the ORT data. Table 3.3 shows the standard deviation and mean speeds for the six freeway segments.

Table 3.3: Standard deviation and mean speeds for the freeway segments for TomTom data

Freeway segment	15-minute μ	15-minute σ	1-hour μ	1-hour σ
N1_SB_1006_1008	87.4	5.0	86.9	11.6
N1_NB_1007_1005	95.1	4.6	94.9	16.4
N1_SB_1010_1012	79.6	6.1	86.3	16.9
N1_NB_1013_1011	92.1	8.2	91.7	25.6
R21_SB_1040_1041	106.3	4.9	106.2	17.1
R21_NB_1042_1039	108.1	5.0	108.1	16.0

Equation 3.1 shows the calculation of the minimum sample size using the z-statistic, relative allowable error, sample mean and standard deviation.

$$N_{MIN} = \left(\frac{z_{\alpha/2} \times \frac{\sigma}{\mu}}{\theta} \right)^2 \qquad (3.1)$$

Where:

N_{MIN} = minimum sample number;

$z_{\alpha/2}$ = critical normal deviate for the desired confidence interval;

σ = standard deviation;

μ = mean speed; and

θ = error tolerance level.

The minimum sample sizes for the segments were determined using Equation 3.1 and data in Table 3.3. A tolerance error of 5 % and a 99 % confidence interval were selected (Midwestern Consulting, 2008). Table 3.4 shows the required and observed minimum sample sizes for each freeway segment. The results in Table 3.4 show that there is a 99 % degree of confidence that the TomTom average speeds are within ±5 % of the mean speeds given in Table 3.3. The observed minimum sample sizes were greater than the required minimum sizes. This shows that the TomTom speeds were based on sufficient number of probes and can be regarded as “reliable” speed estimates. In addition, the ORT sample sizes were also significantly larger than the required sample sizes.

Table 3.4: Prescribed and observed minimum sample sizes

Freeway segment	15-minute – required	15-minute - observed	1-hour - required	1-hour - observed
N1_SB_1006_1008	9	67	48	522
N1_NB_1007_1005	7	60	80	335
N1_SB_1010_1012	16	37	102	298
N1_NB_1013_1011	22	79	207	386
R21_SB_1040_1041	6	54	69	384
R21_NB_1042_1039	6	50	59	306

The total ORT and TomTom sample sizes for the six freeway segments for all the weekdays (Monday – Friday) in February 2015 are given in Table 3.5. Although the TomTom samples were only 2.8 % of the ORT observations, the number of observations is statistically significant, based on a 5 % tolerance error and 99 % confidence interval that only requires at most 207 observations.

Table 3.5: Sample sizes for the freeway segments (February 2015)

Freeway segment	ORT 15-min	TomTom 15-min	ORT 1-hour	TomTom 1-hour
N1_SB_1006_1008	961173	25967	961173	25997
N1_NB_1007_1005	974895	23278	974895	23305
N1_SB_1010_1012	524421	18819	524421	18847
N1_NB_1013_1011	542775	15424	542775	15456
R21_SB_1040_1041	549093	15613	549093	15642
R21_NB_1042_1039	610009	15093	610009	15118

3.5 DATA REDUCTION AND PROCESSING

The ORT data requested from SANRAL was in its raw format. The data was processed and reduced before conducting any comparisons. TomTom historical data was already processed and reduced, hence no further processing was required. This section outlines the procedure for processing the raw benchmark data.

3.5.1 ORT DATA PROCESSING

The raw benchmark data for February 2015 was provided as a 2.5 GB text file, containing over 70 million counts recorded on all the gantries shown in Figure 3.1. A screenshot of the text file containing the benchmark data is shown in Figure 3.11.

TimeStamp|Gantry|Class|Vehicle
2015-02-02 06:15:29|1002|2|1081750
2015-02-02 06:19:08|1008|2|2631494
2015-02-02 06:14:39|1035|2|5138626
2015-02-02 06:14:00|1040|2|5099741
2015-02-02 06:13:10|1002|2|2930995
2015-02-02 06:14:19|1038|2|4712232
2015-02-02 06:15:26|1008|2|1205656
2015-02-02 06:15:44|1023|2|1569532

Figure 3.11: A screenshot of the ORT raw data

The four columns represent the following attributes:

i. Date and time stamp;
ii. Gantry number (1000 + gantry number);
iii. Vehicle class (1 – motorcycle, 2 – car, 3 – small heavy vehicles, 4 – heavy vehicles); and
iv. Vehicle Licence Number ID (random number assigned to Vehicle Licence Number to anonymise the data).

The first row of Figure 3.11 describes a class 2 vehicle, with an anonymised identifier 1081750, which crossed gantry 1002 on the 2^{nd} of February 2015 at 06:15:29. In the form shown in Figure 3.11, the information is not very useful. The first step to processing and reducing the benchmark data was to identify and define the freeway segments. A freeway segment was introduced in Figure 3.2 and it was defined as the road section between two toll gantries. The freeway segments used in this study were discussed in section 3.4.3.

Turning now to the method for determining the travel times and speeds on the freeway segments, the procedure consists of a number of steps, namely *matching*, *15-minute interval allocation*, *1-hour interval allocation*, *filtering* and *aggregation*. These steps are discussed in detail in the sections that follow.

3.5.1.1 MATCHING

On a freeway segment, as shown in Figure 3.2, a vehicle can cross:

i. First gantry but not second gantry – vehicle that exited the freeway via off-ramps or a systems interchange;

ii. Second gantry but not first gantry – vehicle that entered the freeway via on-ramps or a system interchange; and

iii. First gantry and then second gantry – vehicle that traversed the entire freeway segment.

The first two types of observations cannot be used to calculate travel times and speeds. With the third type of observations, it is possible to compute travel times and speeds. It should be noted that the third type of observation could be executed as a broken trip or as an unbroken trip.

A broken trip describes a vehicle that crosses the first gantry then exits the freeway and re-enters the freeway before crossing the second gantry. An unbroken trip is one that is executed without exiting and re-entering the freeway. Broken trips are undesirable as these are typically executed in longer travel times and results in low speeds. Including the speeds from broken trips can distort the reference speeds. The procedure for eliminating broken trips is discussed in the data-filtering plan (section 3.5.1.4).

A match is obtained when a unique traffic attribute, such as a licence plate, is identified at one point on the roadway and the same attribute is re-identified at a later stage. An example of a matched observation is illustrated in Table 3.6, where a vehicle with an anonymised ID 156816 crossed gantry 1010 on the 2nd of February 2015 at 13:14:16. The same vehicle then crossed gantry 1012 at 13:19:31 of the same day. The travel time for this vehicle was 315 seconds, as shown in Table 3.7. The length of the freeway segment was 11.0 km. Therefore, the speed of this vehicle was 125.7 km/h.

Table 3.6: Example of a matched observation, adapted from (Haghani *et al.*, 2009)

ID	**Gantry**	**Date and time**
156816	1010	2015/02/02 13:14:16
156816	1012	2015/02/02 13:19:31

Table 3.7: Determination of travel time from matched observations

ID	**First gantry**	**Second Gantry**	**Road segment**	**Travel time (s)**
156816	1010	1012	N1_SB_1010_1012	315

3.5.1.2 METHOD OF ALLOCATING TRIPS TO A 15-MINUTE INTERVAL

After completing the matching step, the benchmark observations were used to compute travel times and speeds. TomTom historical data was based on a time interval, hence reference data was also classified into the respective time intervals.

Two time intervals, 15-minute and 1-hour, were used in this investigation. The observations were grouped based on the 15-minute and 1-hour intervals they belong. However, trips that start in one 15-minute interval do not always end in the same time interval. This section describes the procedure used to allocate the various types of trips into their respective 15-minute and 1-hour intervals.

3.5.1.2.1 Type 1: One 15-min time interval

Type 1 trips were the simplest to deal with. These represent the observations that start and end in one 15-minute time interval. For example, consider a trip that starts at 07:02:00 and ends at 07:10:30. The travel time for this observation is 510 seconds and this observation belongs to the 07:00-07:15 interval.

3.5.1.2.2 Type 2: Two consecutive 15-min time intervals

The second type of trip is one that starts in one 15-minute time interval and ends in the next 15-minute time interval. This trip belongs to the 15-minute interval in which the vehicle spent more time. For example, consider a trip that starts at 06:56:00 and ends at 07:10:30. The travel time is 870 seconds. The time intervals the vehicle spent time in, are 06:45-07:00 and 07:00-07:15. The first 240 seconds are spent in the 06:45-07:00 interval and the other 630 seconds are spent in the 07:00-07:15 interval. For this trip, more time was spent in the 07:00-07:15 interval compared to the 06:45-07:00 interval. Therefore, this observation belongs to the 07:00-07:15 interval.

3.5.1.2.3 Type 3: Multiple 15-min time intervals

The last type of trip is one that is completed in multiple 15-minute intervals. Such a trip originates in one time interval and completed after a number of 15-minute intervals. This trip belongs to the first full 15-minute interval it travelled in. To illustrate the methodology, consider a trip that starts at 06:56:00 and ends at 07:29:30. The travel time for this trip is 2010 seconds. The time intervals the vehicle spent time in, are 06:45-07:00, 07:00-07:15 and 07:15-07:30. The first 240 seconds are spent in the 06:45-07:00 interval, the next 900 seconds are spent in the 07:00-07:15 and the last 870 seconds are spent in the 07:15-07:30 interval. The first complete 15-minute time interval that the vehicle travelled in is the 07:00-07:15 interval. As a result, this observation belongs to the 07:00-07:15 interval, which is the first complete 15 minutes of the trip.

Trips that are completed in multiple 15-minute time intervals are not typical for a freeway segment of 8 to 12 km in length. Such travel times can only be experienced during extremely congested times and in cases of incidences on the freeway. As a result, these type of trips should be minimised. The data-filtering plan (section 3.5.1.4) describes the procedure for addressing trips completed in multiple 15-minute time intervals.

Note: The analysis period was set from 05:00 to 20:00. It should be noted that observations falling into the 05:00-05:15 and 19:45-20:00 intervals, by virtue of vehicles spending more time in these time intervals, were included in these time intervals although the start or end times of the trips were outside the lower (05:00) and upper limits (20:00).

3.5.1.3 METHOD OF ALLOCATING TRIPS TO A 1-HOUR INTERVAL

In the previous sections, the method for allocating the various type of trips into 15-minute intervals was discussed. To illustrate the method of allocating trips into hourly intervals, consider the 1-hour time interval between 05:00 and 06:00. The trips for the 1-hour interval, 05:00-06:00, are all the trips that belong to the four 15-minute intervals between 05:00-06:00, i.e. all the unique trips in the 05:00-05:15, 05:15-05:30, 05:30-05:45 and 05:45-06:00 intervals. The same procedure was applied to allocate trips to other 1-hour intervals.

3.5.1.4 DATA FILTERING PLAN

In section 3.5.1.1, the method for computing the travel time and speed from a matched observation was illustrated. However, not all travel times and speeds from matched observations are acceptable. It was emphasised that broken trips were undesirable because of the additional travel time that is not due to congestion. Other reasons why some of the observations were deemed unacceptable were (Haghani *et al.*, 2009):

1. Observations falling outside the limits of the analysis period;
2. Illogical observations and observations with unreasonably large travel times; and
3. Observations with speeds that are significantly different from the average speeds observed in that particular period.

For the I-95 Corridor Coalition Project, points 2 and 3 above were used to develop filters for discarding outlying Bluetooth observations (Haghani *et al.*, 2009). Similarly, the same filters are applicable to the ORT data. However, point 1 was added in this study to take into account only the observations that occur between 05:00 and 20:00, the set analysis period.

Filters were designed to eliminate undesirable observations. Ideally, a filter should be wide enough to include observations during congested times and strict enough to exclude observations executed via broken trips and trips completed in multiple 15-minute time intervals for all the different freeway segments. Three filters were designed to address the aforementioned issues and were applied to the pool of raw unfiltered data that resulted from the matching step (see section 3.5.1.1). These filters satisfied the three conditions, namely:

i. Including speeds during congested times;
ii. Minimising the speeds derived from broken trips; and
iii. Minimising the speeds derived from multiple 15-minute intervals.

3.5.1.4.1 Filter 1

The first filter considered the period in which a trip was made. Earlier in the scope section (see section 3.3.3.2), the reader was introduced to the temporal aspects of a data collection exercise where it was observed that no valuable traffic information could be obtained during the night. This is because traffic is typically travelling at free-flow speeds. In light of this, only observations executed in the analysis period (between 05:00 and 20:00) were kept. The other observations falling outside the analysis period were discarded.

3.5.1.4.2 Filter 2

Broken trips typically result in abnormal travel times. Broken trips are defined as those in which a vehicle crosses the first gantry, exits the freeway for some time, re-enters the freeway and crosses the second gantry. There are legitimate cases where vehicles do not exit and re-enter the freeway but still take longer than normal to traverse a given freeway segment, for example, extreme congestion during incidents and peak periods.

The threshold is a trade-off between vehicles that exit and re-enter the freeway in a short period and the vehicles that take longer to traverse a freeway segment due to congestion. After considering both cases, observations with travel time of more than 3600 seconds that survived the first filter were discarded. Illogical trips and trips that were executed in a number of 15-minute intervals were also discarded using this filter. Illogical trips may have originated from the anonymisation of the number plates, vehicles without number plates and any cloning of number plates.

3.5.1.4.3 Filter 3

The third filter addressed the outliers in each 15-minute interval. This filter checked the validity of an observation against the speeds of other vehicles in the same period. For the I-95 Corridor Coalition Project, a similar filter was applied to discard outliers from Bluetooth observations in each 5-minute interval (Haghani *et al.*, 2009).

In each 15-minute interval, the average speed and standard deviation of the speeds were calculated. Speeds falling outside ± 1.5 times the standard deviation from the mean were discarded. This is one of the algorithms recommended for identifying and discarding outliers (Turner *et al.*, 2011).

3.5.1.5 DATA AGGREGATION

For each 15-minute and 1-hour interval, a speed estimate from the service provider was obtained, which was the aggregated speed for that interval. In order to have a basis for comparison for the two datasets, the reference data for each corresponding 15-minute and 1-hour interval were aggregated. Space mean speed in each 15-minute and 1-hour interval is a typical way to aggregate reference data. Space mean speed is the harmonic mean of the speeds of vehicles traversing the freeway segment in that time interval (Garber & Hoel, 2010). TomTom historical data is already aggregated and ready to be used for analysis. This section describes the method for aggregating data.

3.5.1.5.1 TomTom Data Aggregation

For each interval, the speed estimate from TomTom was obtained in its aggregated form, which made it convenient to use. Additionally, data on the variation of speeds for the different sub-segments of the freeway segment was also provided.

3.5.1.5.2 ORT Data Aggregation

For each 15-minute and 1-hour interval, the space mean speed was calculated as follows:

$$S_{ORT} = \frac{NL}{\sum_{i=1}^{n} T_i} \times 3600 \qquad (3.2)$$

Where:

S_{ORT} = ORT speed in km/h;

L = length of a freeway segment in km;

T_i = travel time of the ith observation in seconds; and

N = number of observations.

3.6 SOFTWARE DESIGN

As previously mentioned, vast amounts of data are collected by the ORT system. In general, the more data that is available indicates that the likelihood of obtaining accurate results is high. However, vast amounts of data also suggest that processing the data is cumbersome and labour intensive. Table 3.8 shows the raw traffic counts on the ORT system from January 2015 to June 2015.

Table 3.8: Monthly traffic counts recorded on the ORT system

Year	**Month**	**Traffic counts**
2015	January	68 163 784
2015	**February**	**70 466 183**
2015	March	78 672 333
2015	April	71 644 727
2015	May	75 716 814
2015	June	73 878 018

From Table 3.8, no less than 68 million raw counts per month were recorded over that period and over 70 million raw counts were recorded for February 2015 alone (shown in bold). It was not possible to process and reduce the raw data using packages such as MS Excel.

For this reason, it was decided to develop a Java program to process the data. The program used the matching, filtering, processing and aggregation methods explained in the data reduction and processing section (see section 3.5). A summary of the steps taken in developing the Java program is presented in Appendix C.

Considering the ORT database, about 200 000 observations of the 70 million did not contain an ID. These were discarded because it was not possible to use these observations in computing reference speeds. The observations with missing IDs were identified across all the gantries and not a specific area. It is highly unlikely that discarding these observations, which account for about 0.28 % of the total observations, would cause a considerable difference in the magnitude of the reference speeds.

Figure 3.12 shows the graphical user interface (GUI) of the data processing program. The user enters the first gantry number, second gantry number and the distance between the two gantries then clicks the search button.

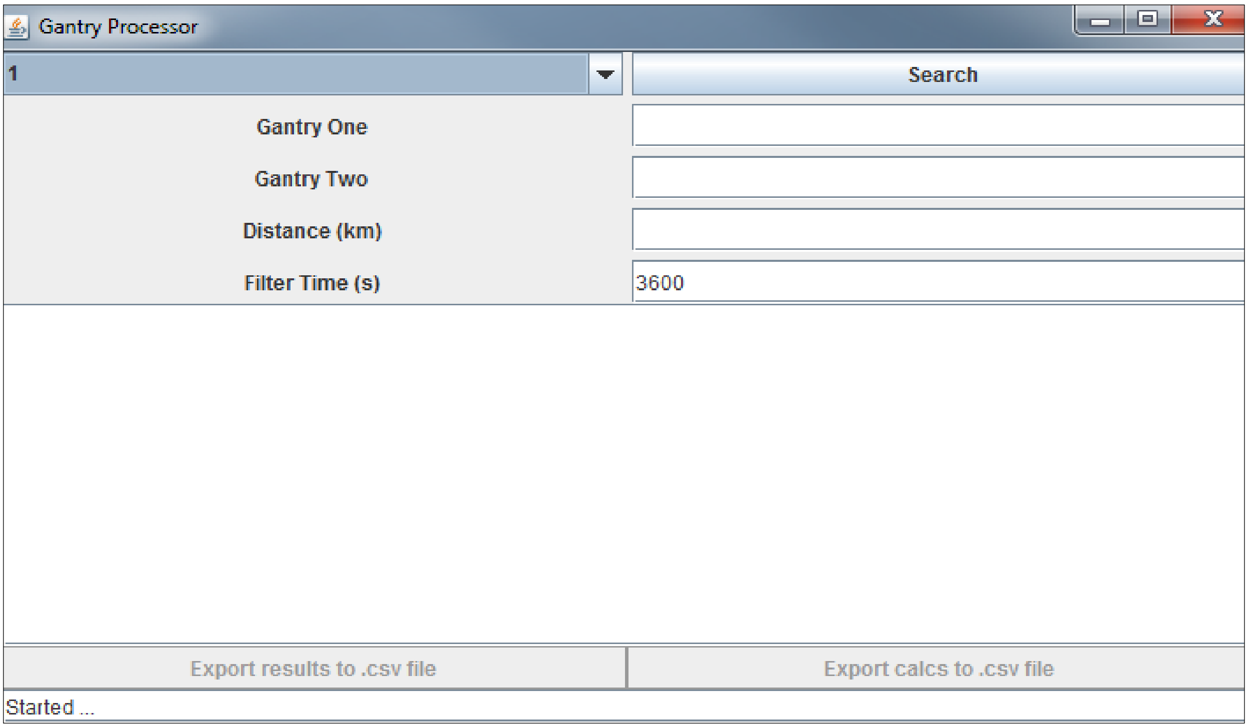


Figure 3.12: Graphical User Interface (GUI) for the data processing program

Subsequently, the results file, a Comma Separated Value (CSV) file with all the matched observations i.e. these are all the observations that were identified at the matching step (section 3.5.1.1), was created and available for export. A screenshot of the results file is shown in Figure 3.13.

ID	Gantry 1	Gantry 2	Date&Time - 1	Date&Time -2	Time (s)	Speed	Time Interval	
4501438	1010	1012	2015-02-02 05:00	2015-02-02 05:05	311	127.3312	05:00	05:15
4036138	1010	1012	2015-02-02 05:01	2015-02-02 05:06	342	115.7895	05:00	05:15
3973415	1010	1012	2015-02-02 05:01	2015-02-02 05:06	307	128.9902	05:00	05:15
1483413	1010	1012	2015-02-02 05:01	2015-02-02 05:08	401	98.75312	05:00	05:15
3373038	1010	1012	2015-02-02 05:01	2015-02-02 05:07	336	117.8571	05:00	05:15
1117329	1010	1012	2015-02-02 05:01	2015-02-02 05:07	365	108.4932	05:00	05:15
3456368	1010	1012	2015-02-02 05:01	2015-02-02 05:08	429	92.30769	05:00	05:15
366756	1010	1012	2015-02-02 05:01	2015-02-02 05:08	427	92.74005	05:00	05:15
2214102	1010	1012	2015-02-02 05:01	2015-02-02 05:06	305	129.8361	05:00	05:15
4899423	1010	1012	2015-02-02 05:01	2015-02-02 05:07	315	125.7143	05:00	05:15
4051709	1010	1012	2015-02-02 05:02	2015-02-02 05:11	570	69.47368	05:00	05:15
907932	1010	1012	2015-02-02 05:02	2015-02-02 05:07	316	125.3165	05:00	05:15
4555740	1010	1012	2015-02-02 05:02	2015-02-02 05:09	406	97.53695	05:00	05:15
3445432	1010	1012	2015-02-02 05:03	2015-02-02 05:10	457	86.65208	05:00	05:15

Figure 3.13: A screenshot of the results file showing the computation of travel times and speeds

A second file, a calculations CSV file containing the aggregated observations placed in their respective time interval, was created and available for export. A screenshot of the calculations file is shown in Figure 3.14.

Time interval		Speed	Number	Standard deviation	Standard error
05:00	05:15	113.5073	759	3.337248383	0.121134442
05:15	05:30	114.561	1600	2.950363806	0.073759095
05:30	05:45	114.3116	2565	3.167503228	0.062542233
05:45	06:00	114.5161	3857	3.174994418	0.051123214
06:00	06:15	112.2492	4915	3.211703721	0.045811416
06:15	06:30	110.6559	6241	3.063142641	0.038773957
06:30	06:45	109.4193	6910	2.736678921	0.032921895
06:45	07:00	104.3707	7519	3.060225189	0.035291762
07:00	07:15	93.97865	7497	4.105101756	0.047411115
07:15	07:30	73.78384	7764	5.757944859	0.065346862
07:30	07:45	59.08736	7475	5.518763817	0.06383167
07:45	08:00	56.60928	7107	5.90247603	0.070014997
08:00	08:15	56.34721	6722	5.241924869	0.063935422
08:15	08:30	57.28628	6438	4.86200758	0.060595468
08:30	08:45	57.27482	6688	3.969572444	0.048539534

Figure 3.14: A screen shot of the calculations file showing ORT processed data

When ORT data was processed up to the point shown in Table 3.9, the data was ready for comparison. TomTom speed estimates and number of observations were added to the corresponding time intervals. At the stage shown in Table 3.9, the data was ready for analysis. The sections that have been covered up to this stage were data collection, matching, filtering and aggregation. The data has been processed to a stage where the TomTom speed estimates can be compared to the reference speeds based on two numbers representing the aggregated speed for each time interval.

Table 3.9: Typical sample of the processed data

Time interval	TomTom speed	TomTom obs	ORT speed	ORT obs
06:00-06:15	80.2	10	74.6	2145
06:15-06:30	74.9	16	60.7	4216
06:30-06:45	74.0	12	75.2	7553
06:45-07:00	72.6	10	70.5	4656
07:00-07:15	45	12	40.9	5241

Note: Obs = observations

3.7 DATA ANALYSIS

This section discusses the mathematical and statistical analysis of the investigation. The quality measures investigated are accuracy, completeness, validity, coverage and accessibility. The method for determining each of the quality measures is presented in the sections that follow.

The quality measures were calculated at two levels, namely interval and freeway segment level. The quality measures at an individual 15-minute or 1-hour time interval level represent low-level analysis that revealed details on the periods that were problematic. A low-level analysis focuses on the individual intervals. A high-level analysis typically gives an indication of the quality measures for the freeway segment as opposed to individual time intervals.

3.7.1 ACCURACY

The first key information that a TomTom historical data user is interested in knowing is whether the mean speed estimates from the service provider are the same as the mean reference speeds. A statistical hypothesis test was conducted to determine whether there are significant differences between the mean TomTom speeds and the mean reference speeds.

Probe data users are interested in knowing the quantity of error in the data, if an error is there in the first place. Better still; a user would like to know what kind of an error it is i.e. the positive or negative bias in the data. The methods for determining the accuracy of TomTom historical data are discussed in this sections that follow.

3.7.1.1 HYPOTHESIS TESTING

A statistical hypothesis test was conducted for each test segment to get an insight on whether or not the mean TomTom speeds were significantly different from the mean reference speeds. For this test, the null hypothesis was that the mean TomTom speed (μ_{TOMTOM}) was not significantly different from the mean ORT speed (μ_{ORT}), at a 95 % confidence level.

$$\begin{cases} H0: \mu_{TOMTOM} = \mu_{ORT} \\ H1: \mu_{TOMTOM} \neq \mu_{ORT} \end{cases} \tag{3.3}$$

The test statistic used was expressed as:

$$t = \frac{(\mu_{TOMTOM} - \mu_{ORT})}{\sqrt{\frac{\sigma_{TOMTOM}}{N_{TOMTOM}}^2 + \frac{\sigma_{ORT}}{N_{ORT}}^2}} \tag{3.4}$$

Where:

μ_{TOMTOM} = mean TomTom speed in km/h;

μ_{ORT} = mean ORT speed in km/h;

σ_{TOMTOM} = standard deviation of the TomTom speeds in km/h;

σ_{ORT} = standard deviation of the ORT speeds in km/h;

N_{TOMTOM} = number of TomTom observations; and

N_{ORT} = number of ORT observations.

3.7.1.2 ACCURACY MEASUREMENT AT A TIME INTERVAL LEVEL

A method to give the reader a better perspective of the variability of the speeds was adapted from Battelle *et al.* (2004). Three error quantities were calculated for each 15-minute and 1-hour interval. In mathematical terms, the errors are described as follows:

$$E_1 = \frac{(S_{TOMTOM} - S_{ORT})}{S_{ORT}} \times 100 \quad (3.5)$$

$$E_2 = |S_{TOMTOM} - S_{ORT}| \quad (3.6)$$

$$E_3 = S_{TOMTOM} - S_{ORT} \quad (3.7)$$

Where:

S_{ORT} = ORT speed in a 15-minute or 1-hour interval; and

S_{TOMTOM} = TomTom speed estimate in a 15-minute or 1-hour interval.

3.7.1.3 ACCURACY MEASUREMENT ON A FREEWAY SEGMENT

In order to get an idea of the differences between the reference speeds and the TomTom speed estimates, the following error quantities were determined at the freeway segment level.

$$Signed\ Error\ (\%) = \frac{1}{N} \times \sum_{i=1}^{N} \left(\frac{S_{TOMTOM} - S_{ORT}}{S_{ORT}} \right) \times 100 = \frac{\sum_{i=1}^{N} E_1}{N} \quad (3.8)$$

$$AASE = \frac{1}{N} \times \sum_{i=1}^{N} |S_{TOMTOM} - S_{ORT}| = \frac{\sum_{i=1}^{N} E_2}{N} \quad (3.9)$$

$$SEB = \frac{1}{N} \times \sum_{i=1}^{N} (S_{TOMTOM} - S_{ORT}) = \frac{\sum_{i=1}^{N} E_3}{N} \quad (3.10)$$

Where:

S_{ORT} = ORT speed in a 15-minute or 1-hour interval;

S_{TOMTOM} = TomTom speed estimate in a 15-minute or 1-hour interval;

N = number of 15-minute or 1-hour intervals;

E_1 = signed error for a one 15-minute or 1-hour interval;

E_2 = AASE for one 15-minute or 1-hour interval; and

E_3 = SEB for one 15-minute or 1-hour interval.

It is important to note the difference between the error quantities, AASE and SEB. In the case of SEB, positive and negative errors can cancel out thus resulting in an average error of a less magnitude. This is not true for AASE since the error terms are positive and do not cancel out.

Note: The error quantities, E_1, E_2 and E_3, for individual 15-minute or 1-hour time interval are the *signed error*, *AASE* and *SEB* for N = 1, respectively. This was deliberate done so that the error quantities for an individual time interval could be aggregated and result in the error quantities for the freeway segments.

The criteria used to evaluate the accuracy measures at a freeway segment and individual interval level is shown in Table 3.10. It is necessary to get indication of the error quantities, E_1, E_2 and E_3 meeting the accuracy criteria. As explained above, the error quantities, E_1, E_2, E_3, are the *signed error*, *AASE* and *SEB* for N = 1, respectively. Therefore, the evaluation criteria for the freeway segments are similar to the criteria for the individual time intervals.

Table 3.10: Evaluation criteria for the accuracy measures

Error Type	**Maximum allowed error**
Signed error , E_1	±10 %
AASE, E_2	10 km/h
SEB, E_3	±7.5 km/h

The guidelines do not establish the level of accuracy that a traffic data provider must achieve hence the quality level must be defined by the user or purchaser of the data based on their intended application of the data (Turner *et al.*, 2011).

A ±10 % allowable signed error is recommended for planning and programming studies (Midwestern Consulting, 2008). Although Haghani *et al.* (2009) and Texas Transportation Institute (2012) used an allowable AASE of 10 mph (16 km/h), this limit was too wide. For this reason, this study reduced the allowable AASE limit to 10 km/h.

Unlike the AASE, the SEB allows positive and negative error to cancel out. For this reason, the allowable SEB should be typically less than the allowable AASE. An allowable SEB of ±7.5 km/h was selected. This is roughly in line with to the allowable SEB of ±5 mph (±8 km/h) used in the I-95

Corridor Coalition Project (Haghani *et al.*, 2009). Furthermore, Garber & Hoel (2010) recommended maximum allowable error of ±8 km/h, which is slightly greater than the ±7.5 km/h selected for this study.

3.7.1.4 COMPARISON OF THE 15-MINUTE AND 1-HOUR INTERVAL SPEEDS

A statistical hypothesis test was conducted for the combined and individual freeway segments to find out whether or not there were significant differences between the error quantities (signed error, AASE and SEB) derived from the 15-minute and 1-hour interval speeds. The null hypothesis was there were no significant differences between the error quantities derived from the 15-minute and 1-hour interval speeds, at a 95 % confidence level.

3.7.2 COMPLETENESS

Completeness is the degree to which data values are present in the attributes that require them (Cykana *et al.*, n.d.). Completeness was calculated as the number of time intervals that had usable data as a percentage of the total expected values. Data is "usable" if the TomTom sample size was equal to or greater than the minimum sample size. The minimum sample sizes were determined in section 3.4.4.3.

$$Percent\ Complete\ (\%) = \frac{n_{\text{AVAILABLE VALUES}}}{n_{\text{TOTAL EXPECTED}}} \times 100 \qquad (3.11)$$

Where:

$n_{\text{AVAILABLE VALUES}}$ = the number of rows with available values present; and

$n_{\text{TOTAL EXPECTED}}$ = the number of rows expected.

3.7.3 VALIDITY

Validity is the degree to which data values satisfy the specified acceptance requirements or fall within the respective domain of acceptable values. It is possible to calculate validity at a freeway segment level but this can be misleading because only six freeway segments were investigated. Validity was calculated at the interval level. Table 3.11 shows the validity criteria at the interval level.

$$Percent\ Valid\ (\%) = \frac{n_{\text{VALID}}}{n_{\text{TOTAL}}} \times 100 \qquad (3.12)$$

Where:

n_{VALID} = the number of rows with values meeting validity criteria; and

n_{TOTAL} = the total number of rows subjected to validity criteria.

Table 3.11: Validity criteria for 15-minute and 1-hour intervals

Data quality	**Evaluation criteria**
Very high quality	PV ≥ 85 %
High quality	75 % ≤ PV < 85 %
Moderate quality	50 % ≤ P < 75 %
Low quality	P < 50 %

Note: PV = Percent Valid

3.7.4 COVERAGE

Coverage was determined based on the total length of the segments that were investigated to the total length of all the freeway segments, expressed as a percentage. For this study, coverage was calculated as follows:

$$Percent\ Coverage\ (\%) = \frac{Length\ of\ freeway\ segments\ investigated}{Total\ length\ of\ all\ freeway\ segments} \times 100 \qquad (3.13)$$

3.7.5 ACCESSIBILITY

Accessibility is the relative ease with which data can be retrieved and manipulated by data consumers to meet their needs. For this study, accessibility focused more on the usability of the data. A qualitative assessment, in form of a description, was carried out to get an idea of the level of ease with which TomTom and reference data was collected, stored, processed and reduced.

3.8 LIMITATIONS AND ASSUMPTIONS

1. Turner *et al.* (1998) recommended freeway segments between 1.6 km and 4.8 km in length whilst Haghani *et al.* (2009) used segments greater than one mile (1.6 km). However, the freeway segments used in this study were considerably longer (8-12 km). The physical infrastructure of the ORT system is already in place in Gauteng. Therefore, it was not possible for the researcher to control the length of freeway segments used in this investigation.
2. Freeway segments that start on one freeway and end on another freeway were not considered in this study.
3. The ORT and TomTom link lengths were not equal. As a result, travel time analysis was not conducted. Only speed analysis was conducted.

4. The sample of the traffic that was considered in determining the travel times and speeds by the service provider is unknown. In addition, information about the individual vehicles in the sample is unknown. Only the aggregated traffic data for the sample was provided.
5. The standard deviation of the sample used by the service provider in deriving the speeds estimates was assumed to be equal to the standard deviation of the entire traffic stream.
6. The guidelines and standards recommended that data should be collected under good weather conditions and without any incidents that affect driver behaviour (Midwestern Consulting, 2008). In this study, the effects of weather conditions and accidents on the freeways were not investigated.
7. It was assumed that the ORT system captures the population of the traffic stream traversing past a gantry.

3.9 CONCLUSION

The research design and methodology chapter introduced a quantitative analysis as the choice of the research design and a field experiment as the research methodology. The various aspects of the study such as sampling, measurement errors and sample sizes were discussed. The procedure for processing and reducing data was discussed. Subsequently, a Java program for processing ORT data was presented. Lastly, the chapter described the analyses conducted to obtain the results and findings.

CHAPTER 4 : RESULTS

4.1 INTRODUCTION

This part of the thesis presents the results of the evaluation of the quality measures of TomTom historical data. A detailed discussion of the results follows in Chapter 5. Firstly, the speed profiles for the six freeway segments for both the 15-minute and 1-hour intervals are presented. Subsequently, the findings for the quality measures, namely accuracy, completeness, validity, coverage and accessibility are presented. To conclude the chapter, a traffic data scorecard is presented.

4.2 SPEED PROFILES

The sections that follow present the weekday speed profiles for the six freeway segments for the evaluation period, February 2015. The speed profiles show the speed comparisons between the TomTom speed estimates and the reference speeds aggregated over a 15-minute and 1-hour interval. The analysis period was between 05:00 and 20:00.

4.2.1 15-MINUTE SPEED PROFILES

The speed profiles at the 15-minute interval are presented in the sections that follow.

4.2.1.1 N1 BEN SCHOEMAN HIGHWAY

The speed profile for the N1 Ben Schoeman southbound section is shown in Figure 4.1.

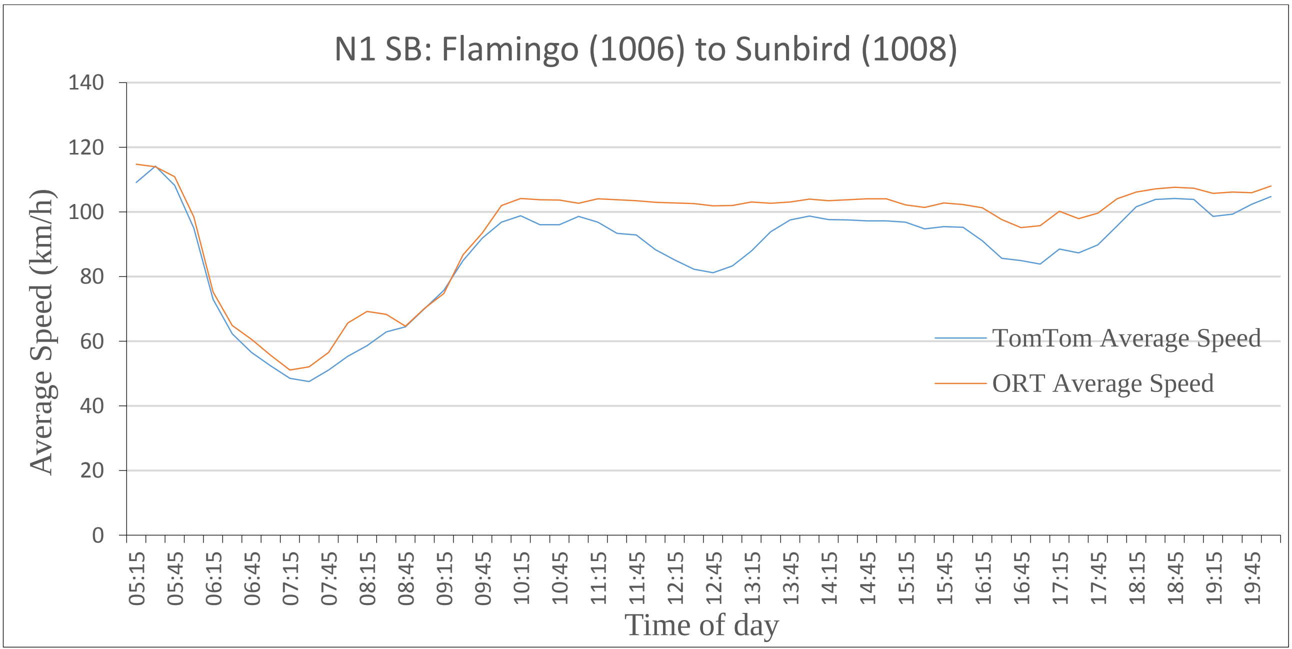


Figure 4.1: Speed profile between Flamingo (1006) and Sunbird (1008)

The N1 Ben Schoeman Highway southbound speed profile shows a reduction in speed during the morning peak period from 06:15 to 09:15, with speeds less than 80 km/h (Figure 4.1). For the rest of the analysis period, reference speeds just above 100 km/h were observed.

On the northbound section of the N1 Ben Schoeman Highway, no reduced speeds during the morning peak was observed but rather an evening peak between 16:15 and 17:45 (Figure 4.2). As for most of the analysis period, speeds around 105 km/h were observed, which were slightly higher than the speeds observed in the southbound direction. This is in line with the commuting patterns between Johannesburg and Pretoria.

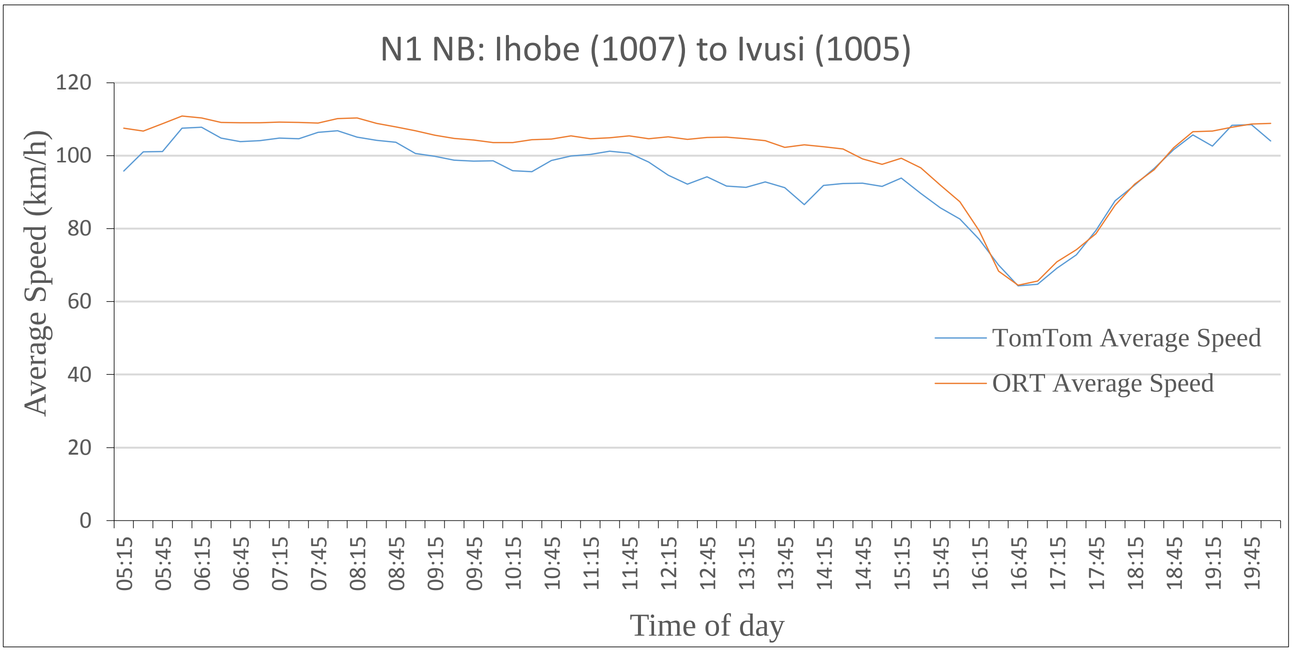


Figure 4.2: Speed profile between Ihobe (1007) and Ivusi (1005)

4.2.1.2 N1 WESTERN BYPASS HIGHWAY

Figure 4.3 shows the speed profile for the N1 Western Bypass southbound section. A reduced speed profile in the morning peak from 07:30 to 09:30 and evening peak from 15:45 to around 19:00 were observed.

The N1 Western Bypass northbound section shows one well-defined peak in the morning between 06:15 and 09:00 (Figure 4.4). During this morning peak, speeds as low as 25 km/h were recorded. For the off-peak periods, speeds around 110 km/h were observed.

As expected, the speed profiles observed for the N1 Western Bypass southbound and northbound section are different from each other because these two sections are independent and have different commuting patterns.

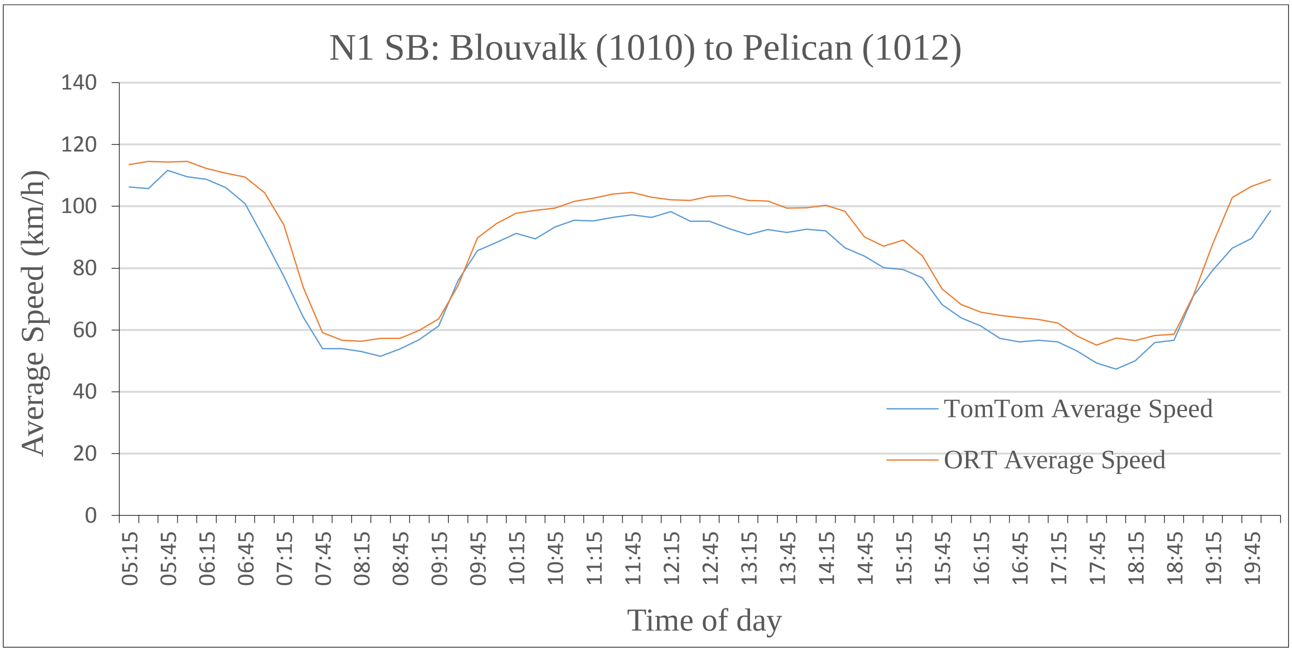


Figure 4.3: Speed profile between Blouvalk (1010) and Pelican (1012)

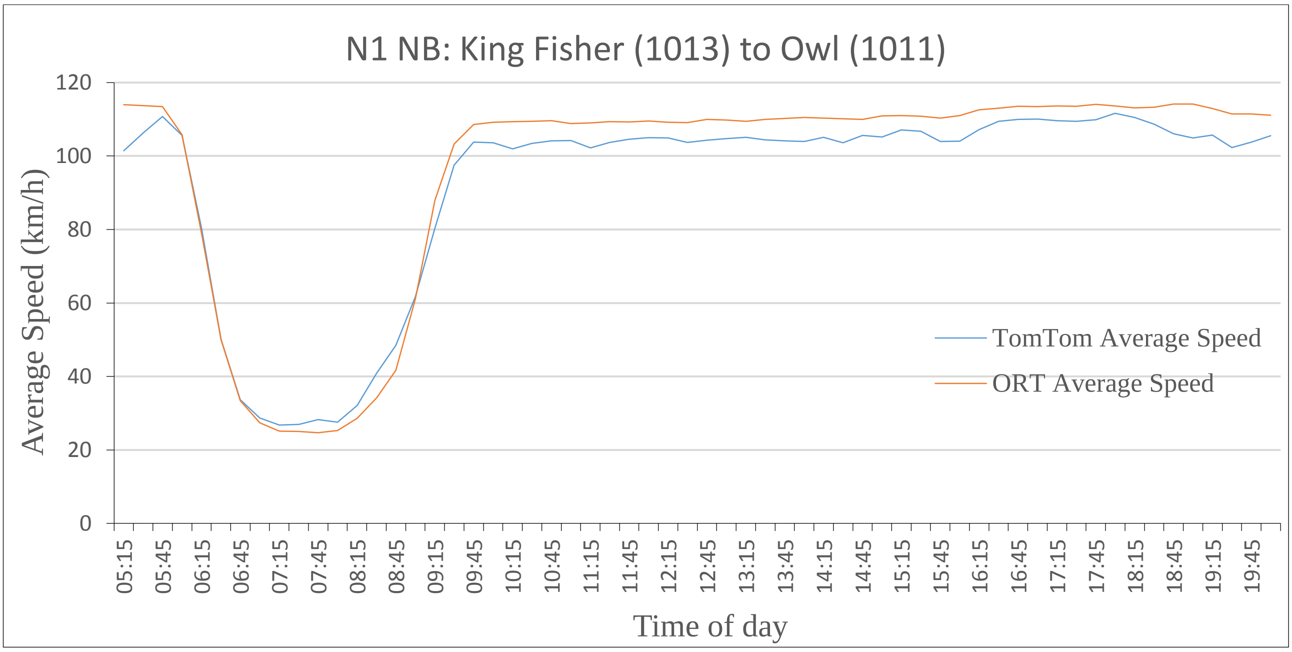


Figure 4.4: Speed profile between King Fisher (1013) and Owl (1011)

4.2.1.3 R21 ALBERTINA SISULU HIGHWAY

The speed patterns on the R21 Albertina Sisulu Highway sections were rather different from all four of the N1 sections because no changes in speeds during the morning, afternoon or evening peaks were observed. Speeds well above 100 km/h were observed throughout the analysis period for both sections.

For the southbound section, the speeds marginally decreased from 120 km/h in the morning to 107 km/h in the evening (Figure 4.5). However, on the northbound section, there was little variation in speeds throughout the analysis period (Figure 4.6). An average speed of about 115 km/h was observed.

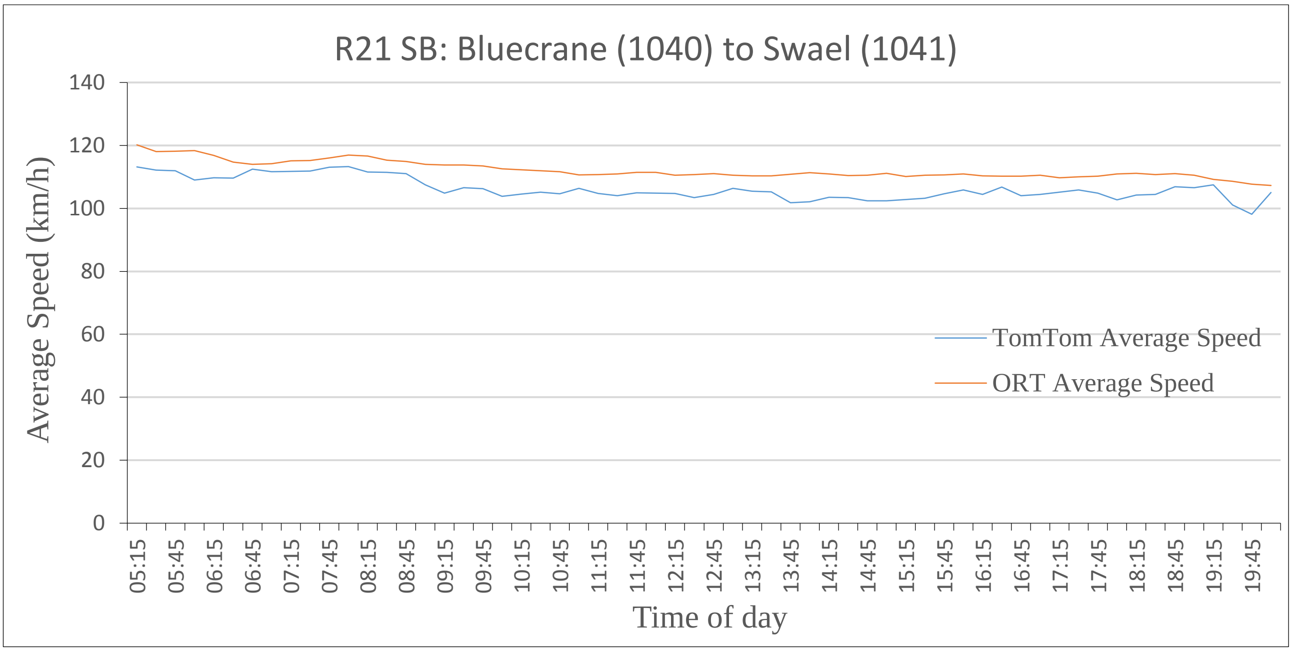


Figure 4.5: Speed profile between Bluecrane (1040) and Swael (1041)

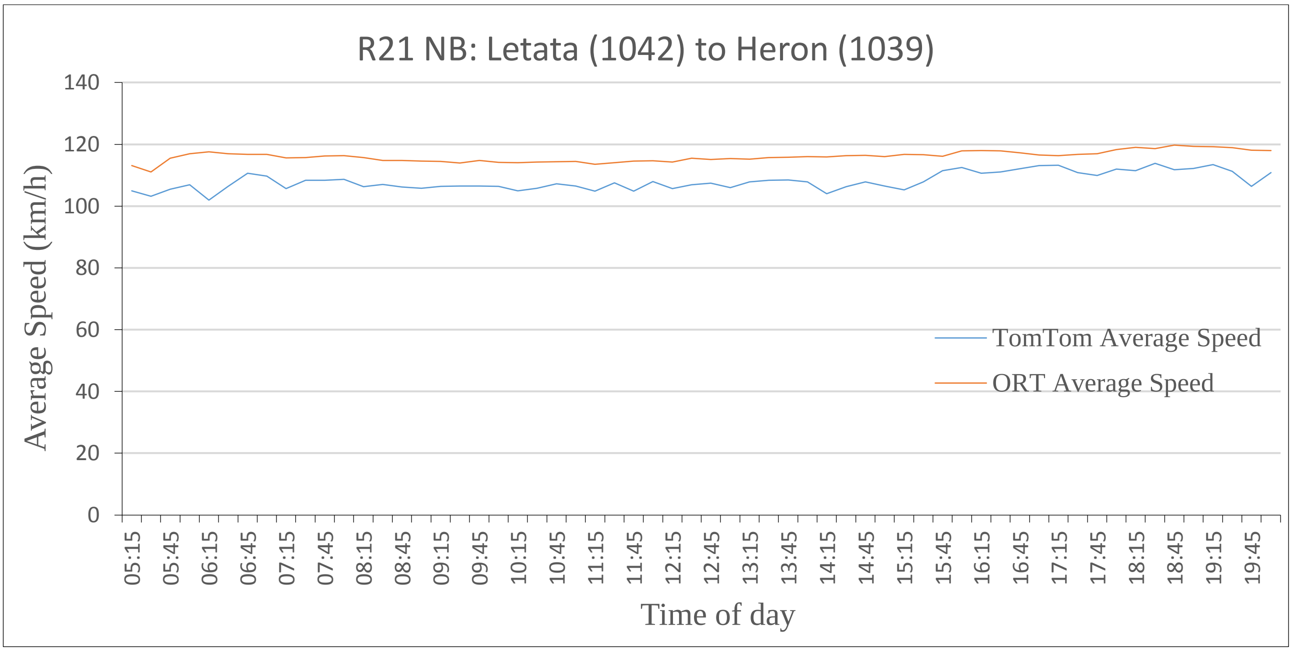


Figure 4.6: Speed profile between Letata (1042) and Heron (1039)

4.3 1-HOUR SPEED PROFILES

The speed profiles for the hourly intervals (Appendix A) were not much different from the speed profiles for the 15-minute interval. There was a slight change in the basic shape of the speed profiles from the 15-minute interval to the hourly interval. However, it is clear that the speed profiles for the 15-minute interval reveal more variation of the speeds throughout the analysis period compared to the hourly interval speed profiles. As a result, the hourly speed profiles were found to be considerably smoother compared to the speed profiles for the 15-minute interval.

4.4 TREND ANALYSIS

Figure 4.7 shows another graphical illustration of the comparison of TomTom speed estimates and the reference speeds for the combined freeway segments (15-minute interval). The trend between the two datasets is clearly depicted in Figure 4.7.

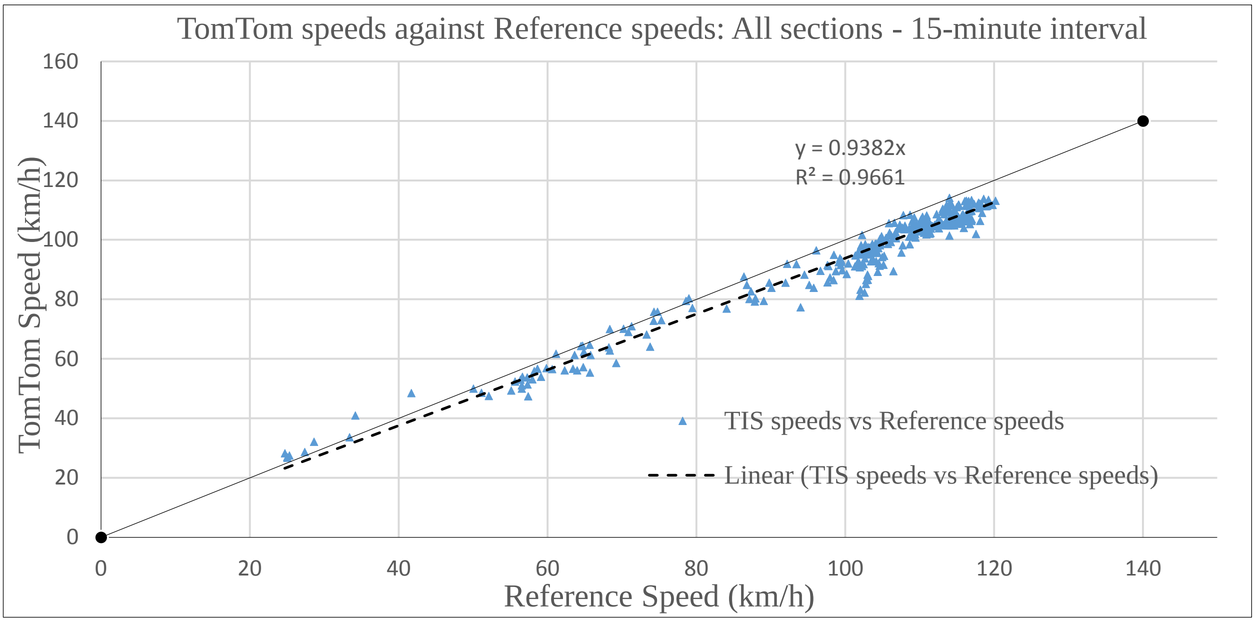


Figure 4.7: Comparison of TomTom and reference speeds (all segments - 15-minute interval)

A considerable proportion of the data points are below the line of correlation, which shows that TomTom speeds were lower than the reference speeds. This is in agreement with what was observed on the speed profiles (Figure 4.1, Figure 4.2, Figure 4.3, Figure 4.4, Figure 4.5 and Figure 4.6).

It was also observed that low speeds were closer to the line of correlation compared to the high speeds. This trend was also observed on the speed profiles of the N1 sections, where the differences between the TomTom speed estimates and reference speeds were less during peak times compared to off-peak times. The trend line's R^2 for the 15-minute interval was 0.966 (Figure 4.7). This demonstrates that there was good correlation between the TomTom speed estimates and reference speeds.

A comparison of the TomTom speed estimates and reference speeds for the 1-hour interval is shown in Figure 4.8. The trend observed for the 1-hour interval is similar to what was observed for the 15-minute interval, where the TomTom speed estimates were closer to the line of correlation at low speeds compared to high speeds. The trend line's R^2 for the 1-hour interval was slightly higher compared to that of the 15-minute interval.

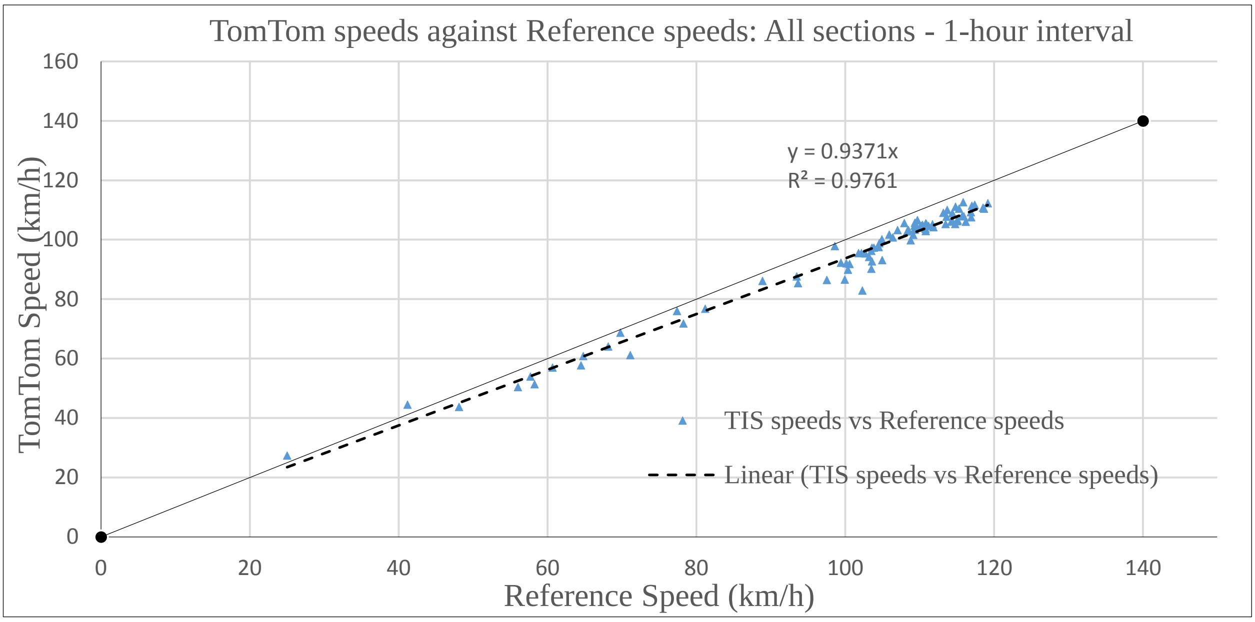


Figure 4.8: Comparison of TomTom and reference speeds (all segments - 1-hour interval)

4.5 ACCURACY

The results for the quality measure, accuracy, are presented in three parts. In the first part, the results of the hypothesis tests for the mean TomTom and reference speeds are presented. The accuracy findings at a freeway segment level follow the hypothesis test results. The results at the time interval level, both 15-minute and 1-hour intervals, are presented in the third part of this section.

4.5.1 HYPOTHESIS TEST

Table 4.1 shows the results of the statistical hypothesis tests that were conducted to find out whether or not the mean TomTom speeds were significantly different from the mean reference speeds (15-minute interval). For each freeway segment, the number of 15-minute intervals in the analysis period was 60 (15 hours x 4). The number of intervals for the combined freeway segments was 360 (15 hours x 4 x 6). For the 1-hour interval, the intervals were 15 (15 hours x 1) and 90 (15 hours x 6) for the individual freeway segment and combined freeway segments, respectively.

Table 4.1: Hypothesis test for TomTom and ORT mean speeds (15-minute interval)

Section Name	S_{TOMTOM}	σ_{TOMTOM}	S_{ORT}	σ_{ORT}	**p-value**
N1_SB_1006_1008	87.4	16.5	94.5	17.2	<0.01
N1_NB_1007_1005	95.1	10.9	100.4	11.9	<0.01
N1_SB_1010_1012	79.6	19.2	86.4	20.5	<0.01
N1_NB_1013_1011	92.1	27.4	96.0	30.4	<0.01
R21_SB_1040_1041	106.3	3.5	112.2	2.8	<0.01
R21_NB_1042_1039	108.1	2.8	116.1	1.7	<0.01
All	94.8	18.8	100.9	20.0	<0.01

The p-value is the calculated probability. For all the individual freeway segments and the combined freeway segments, p-values less than 0.01 were obtained. Therefore, the null hypothesis was rejected. It was concluded that there were statistically significant differences between the mean TomTom speeds and mean ORT speeds at a 95 % confidence interval. Similarly, for the 1-hour interval, there were significant differences between the mean TomTom speeds and mean reference speeds, for the combined and individual freeway segments.

4.5.2 ERROR QUANTITIES AT A FREEWAY SEGMENT LEVEL

The three error quantities considered for the freeway segments are *signed error*, *absolute average speed error (AASE)* and *speed error bias (SEB)*. The determination of these error quantities was discussed in section 3.7.1.3. Table 4.2 shows the *signed error*, *AASE* and *SEB*, for the combined freeway segments, for both the 15-minute and 1-hour intervals.

Table 4.2: Accuracy measures for the combined freeway segments

Error Type	**15-minute interval**	**1-hour interval**	**p-value**
Signed error	-5.8 %	-6.2 %	0.52
AASE	6.4 km/h	6.5 km/h	0.81
SEB	-6.2 km/h	-6.3 km/h	0.69

In addition, a statistical hypothesis test was conducted to find out whether or not there were significant differences between the error quantities derived from the 15-minute and 1-hour interval speeds. The *signed error* was well within the ±10 % allowable error, for both the 15-minute and 1-hour intervals. For the 15-minute interval, an error of -5.8 % was obtained. The *signed error* moderately increased to -6.2 % for the 1-hour interval.

The p-value is the calculated probability and alpha is the threshold value that is measured against the p-value. The p-value was greater than the alpha of 0.05 (Table 4.2). Therefore, there was no significant difference between the *signed error* for speeds aggregated at a 15-minute interval and the *signed error* for speeds aggregated at a 1-hour interval.

The *AASE* for the 15-minute and 1-hour intervals were 6.4 km/h and 6.5 km/h, respectively. Again, the *AASE* for the 1-hour interval was slightly greater than the *AASE* for the 15-minute interval. For both time intervals, the *AASE* were less than the allowable *AASE* of ±10 km/h. The p-value was greater than the alpha of 0.05 (Table 4.2). Therefore, there was no statistically significant difference between *AASE* for the 15-minute and 1-hour interval speeds. As per the definition, the *AASE* has an absolute value hence the *AASE* does not reveal the dataset with the higher speeds.

Similar to the trend that was observed for the *signed error* and *AASE*, there was a marginal increase in the *SEB* from the 15-minute interval to the 1-hour interval. The *SEB* increased by 0.1 km/h, from -6.2 km/h for the 15-minute interval to -6.3 km/h for the 1-hour interval. Furthermore, the *SEB* for the two time intervals were less than the maximum allowable *SEB* of ±7.5 km/h. Again, the p-value was greater than 0.05 (Table 4.2) hence there was no statistically significant difference between the *SEB* for the 15-minute and 1-hour interval speeds.

It is worthwhile to note that the *SEB* was said to have increased from -6.2 km/h to -6.3 km/h. The *SEB* was defined to measure the deviation of TomTom speed estimates from the reference speeds. The negative sign here simply shows that the TomTom speeds were lower than the reference speeds. For both the 15-minute and 1-hour intervals, the *signed error* and *SEB* were negative. This shows that the TomTom speeds were less than the reference speeds, which is in agreement with the trend observed on the speed profiles. Furthermore, the fact that there was a difference of 0.2 km/h in the absolute values of the *AASE* and *SEB*, at both the 15-minute and 1-hour intervals, also shows that the TomTom speeds were consistently lower than reference speeds.

Note: In order to get an insight on whether or not the three *error quantities* were dependent on the individual freeway segment, the *error quantities* at the individual segment level were determined. Table 4.3 shows the notation used in numbering the six freeway segments.

Table 4.3: Number notation for the freeway segments

Number	Section Name
1	N1_SB_1006_1008
2	N1_NB_1007_1005
3	N1_SB_1010_1012
4	N1_NB_1013_1011
5	R21_SB_1040_1041
6	R21_NB_1042_1039

The *signed errors* for the six freeway segments ranged between -2.4 % and – 8.5 % (Figure 4.9). Thus, the *signed errors* for all the six freeway segments were less than the ±10 % allowable signed error, for both the 15-minute and 1-hour intervals. The *signed errors* for the 15-minute interval on the N1 Ben Schoeman Highway sections were -7.4 % and -5.1 % in the southbound and northbound directions, respectively, whereas -8.0 % and -2.4 % were the *signed errors* on the N1 Western Bypass sections in the southbound and northbound directions, respectively. Furthermore, for the 15-minute interval, *signed errors* of -5.3 % and -6.9 % were obtained on the R21 Albertina Sisulu Highway sections in the southbound and northbound directions, respectively.

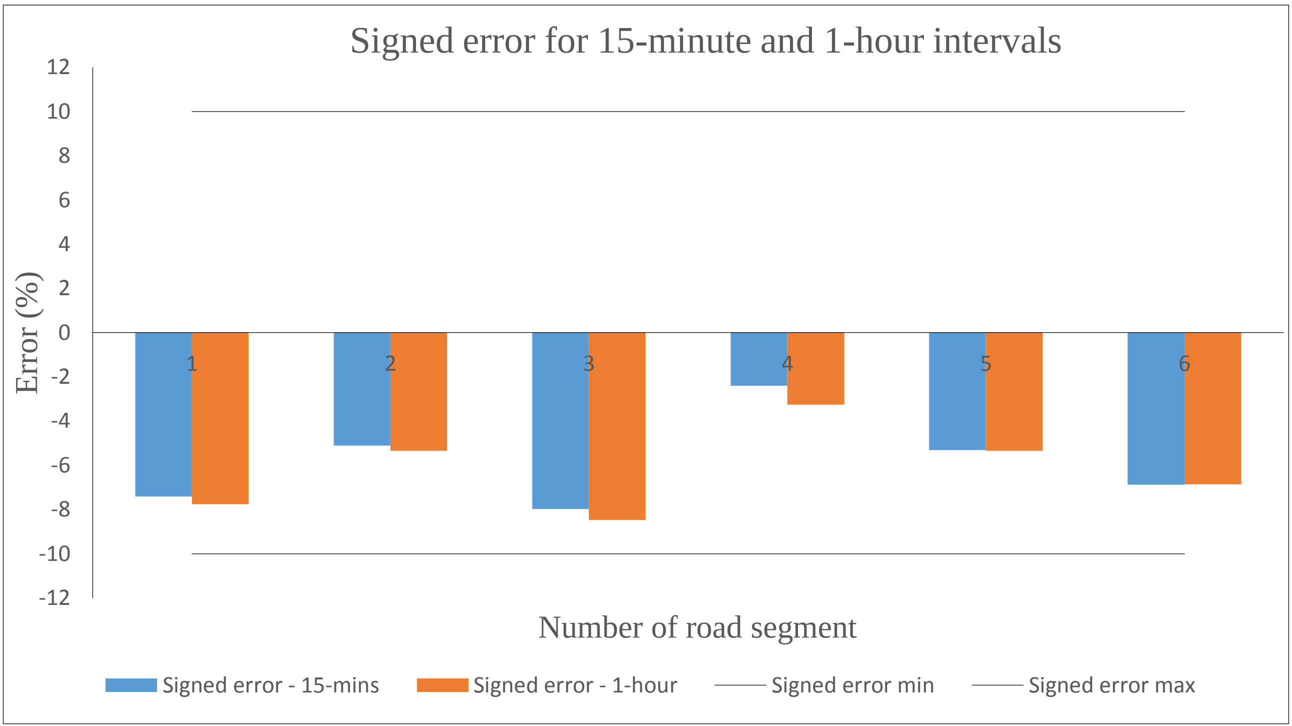


Figure 4.9: Signed error for the individual freeway segments

For each of the freeway segments, the *signed error* derived from a 1-hour interval speeds was slightly greater than the *signed error* derived from the 15-minute interval speeds, with the differences being more pronounced on the N1 sections compared to the R21 sections. The largest difference between the 15-minute interval *signed error* and 1-hour interval *signed error* was observed on the N1 Western Bypass section in the northbound direction.

A t-test (two sample assuming equal variances) was conducted for each segment to test whether or not the *signed error* for the 15-minute interval speeds was significantly different from the *signed error* for the 1-hour interval speeds. The p-values from the t-test were greater than 0.05 for all the six sections. It was concluded that there were no significant differences between the *signed errors* derived from the 15-minute and 1-hour intervals for all the six freeway segments.

The *AASE* for each of the six sections was less than the $AASE_{\text{allowable}}$ of ±10 km/h, for both the 15-minute and 1-hour intervals. The *AASE* for both intervals ranged from 4.9 km/h to 8.0 km/h (Figure 4.10). In addition, the t-test (two sample assuming equal variances) showed that there were no significant differences between the *AASE* for speeds derived from the 15-minute and 1-hour intervals for all the six freeway segments.

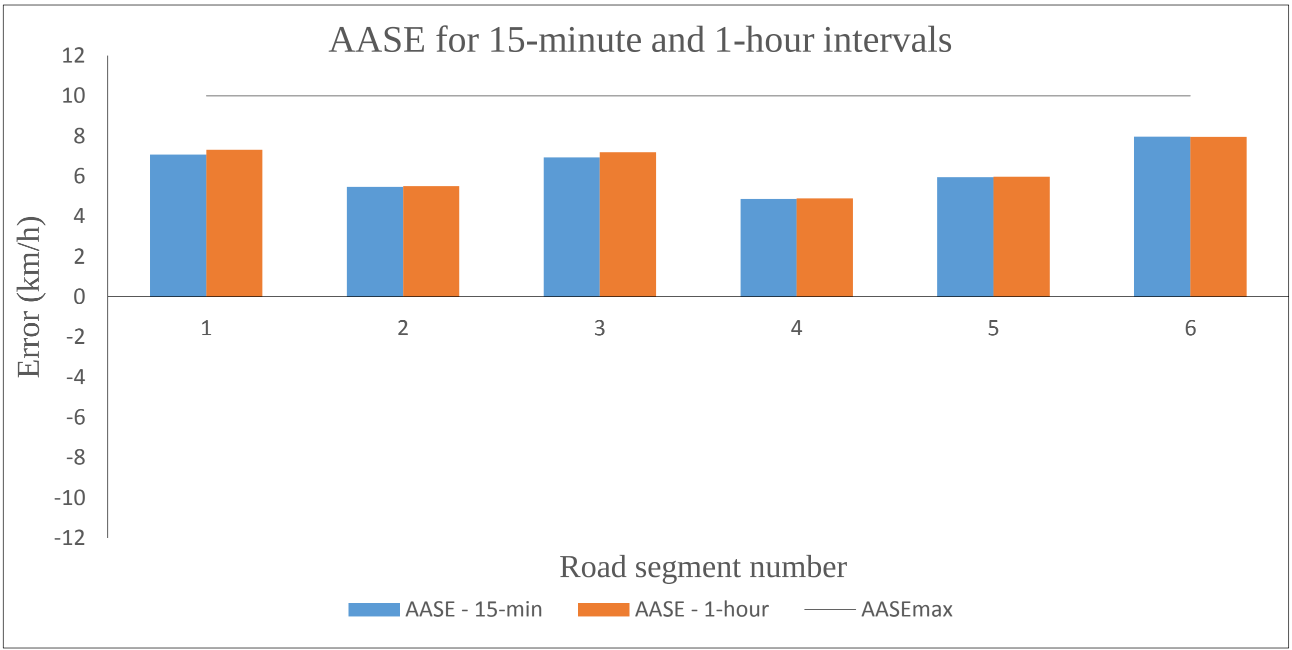


Figure 4.10: AASE for the individual freeway segments

The *SEB* was less than the $SEB_{\text{allowable}}$ of ±7.5 km/h in five of the six sections, for both the 15-minute and 1-hour intervals (Figure 4.11). The freeway segment that did not meet the specified criteria was the R21 northbound section between Letata (1042) and Heron (1039). The *SEB* for this section was

– 8.0 km/h for both the 15-minute and 1-hour intervals. The *SEB* for this section was 0.5 km/h more than the $SEB_{\text{allowable}}$ of ± 7.5 km/h.

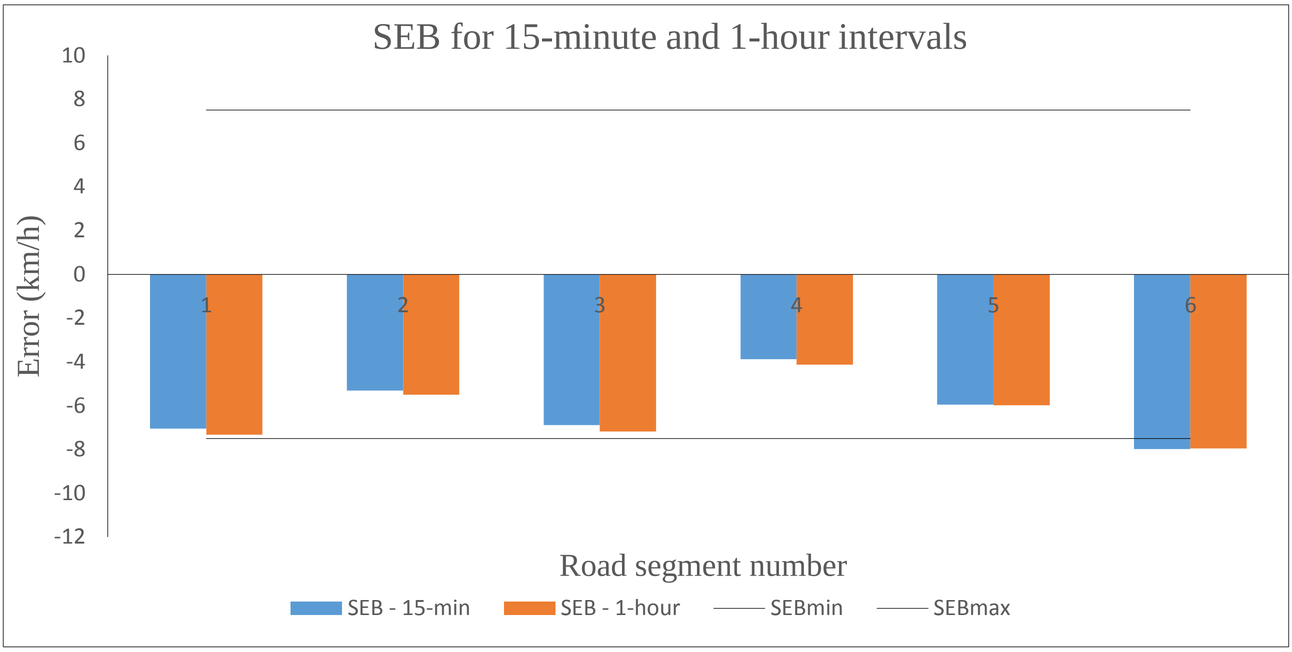


Figure 4.11: SEB for the individual freeway segments

A t-test was conducted to find out if the *SEB* for the 15-minute and 1-hour speeds were significantly different. Similar to the *signed error* and *AASE*, the p-values for all the six freeway segments were greater than 0.05. Therefore, there were no significant differences between the SEB for speeds derived from the 15-minute and 1-hour intervals for all the six freeway segments.

Interestingly, three trends were observed from *signed error*, *AASE* and *SEB* results (Figure 4.9, Figure 4.10 and Figure 4.11). Firstly, the fact that the *signed error* and *SEB* were negative shows that the TomTom speeds were consistently lower than the reference speeds for most of the analysis period. This was also observed on the speed profiles for all the six freeway segments.

The second trend was that the three error quantities for the 1-hour interval were marginally greater than errors for the 15-minute interval for all the six freeway segments, with more pronounced differences on the N1 sections compared to the R21 sections.

Thirdly, the three error quantities in the southbound direction were greater than errors in the northbound direction on the on all the N1 sections. Unlike the trend on the N1 sections, the errors on the northbound segment of the R21 were greater than the errors on the southbound section.

4.5.3 ERROR QUANTITIES AT AN INTERVAL LEVEL

The *signed error*, *AASE* and *SEB* for the combined freeway segments, presented in the previous section, met the specified criteria for accuracy. Because the errors at this level were aggregated, the errors quantities do not give much information on the individual 15-minute and 1-hour time intervals. For this reason, this section presents the findings for the error quantities at the interval level.

The three error quantities, E_1, E_2 and E_3, for the six the six freeway sections were calculated for both the 15-minute and 1-hour time intervals. The error quantities, E_1, E_2 and E_3, are the *signed error*, *AASE* and *SEB* for a single time interval, respectively. The method for determining these error quantities was described in section 3.7.1.2. Figure 4.12, Figure 4.13 and Figure 4.14 show the classification of E_1, E_2 and E_3 into different categories.

A large proportion (86.4 %) of the 360 15-minute intervals that were investigated resulted in an error, E_1, of less than 10 % (Figure 4.12). Another 10 % of the 15-minute intervals had E_1 between 10 % and 15 %. The number of the 15-minute time intervals with E_1 between 15 % and 20 % and greater than 20 % were significantly lower compared to the proportion of intervals in the first two categories. A similar trend was observed for the 1-hour interval, where 88.9 % of the intervals had E_1 less than 10 % and a further 10 % of the intervals had E_1 between 10 % and 15 %. The rest of the intervals were distributed across the last two categories (Figure 4.12).

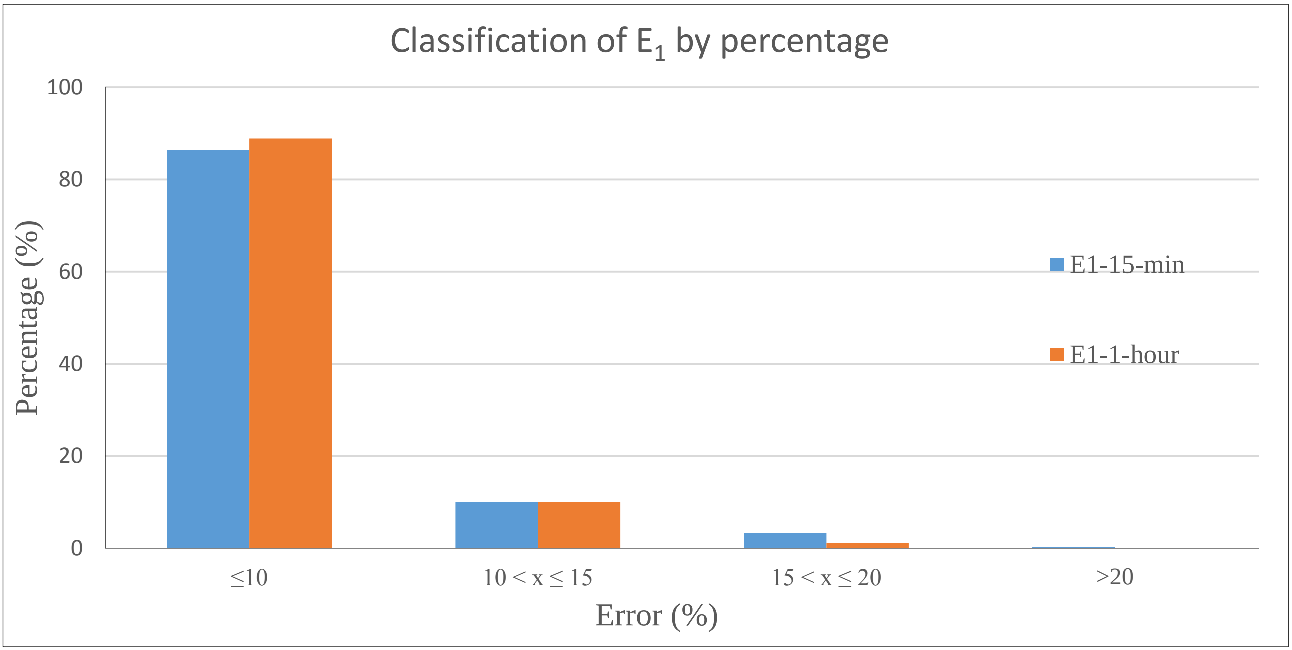


Figure 4.12: Classification of E_1 by percentage

About 87.8 % of 15-minute intervals resulted in E_2 of less than 10 km/h (Figure 4.13). A further 9.2 % of the intervals had E_2 between 10 km/h and 15 km/h whilst the remaining 3 % were distributed across the last two categories.

In the case of the hourly interval, over 90 % of the 1-hour intervals had E_2 less than 10 km/h and 7.8 % of the intervals resulted in E_2 between 10 km/h and 15 km/h. The trend observed for E_2 is noticeably similar to the trend observed for E_1, for both the 15-minute and 1-hour intervals. Over 85 % of the intervals for E_1 and E_2 belonged to the first category whilst the second category consisted of nearly 10 % of the intervals. A small proportion, less than 1 %, of the intervals belonged to the last category.

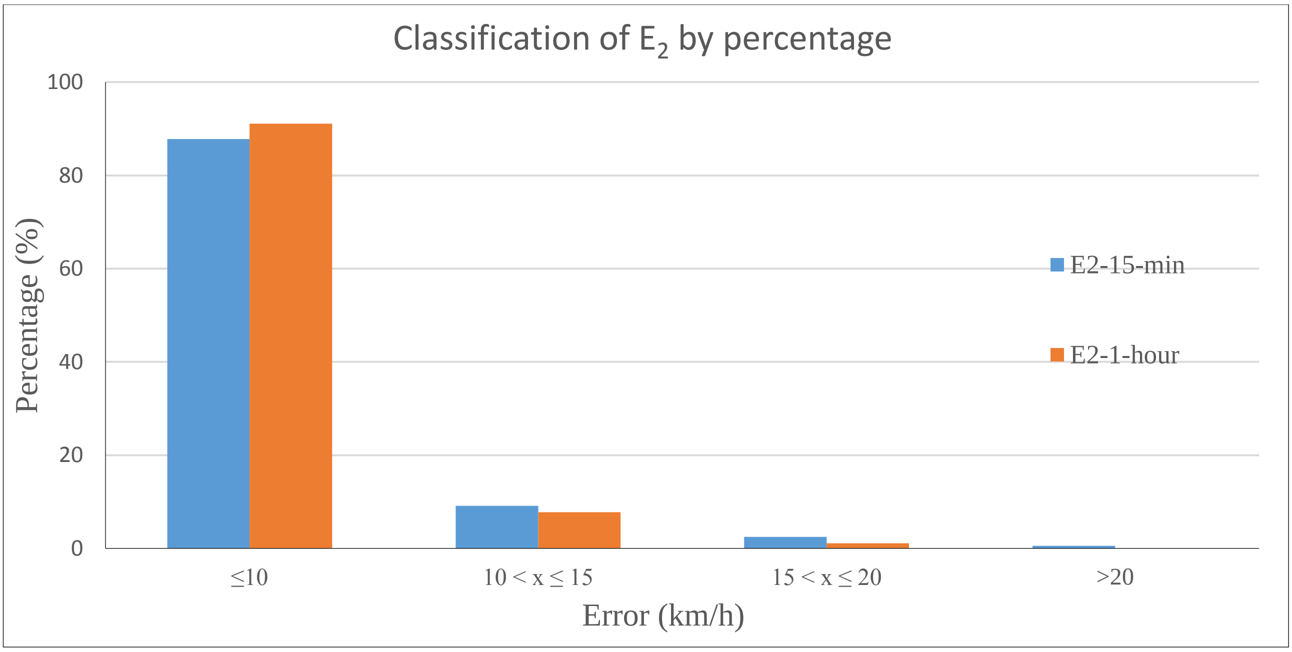


Figure 4.13: Classification of E_2 by percentage

The trend observed for E_3 is slightly different from the trends observed for E_1 and E_2 because the intervals were slightly distributed across the four categories (Figure 4.14). For both the 15-minute and 1-hour intervals, just less than 70 % of the intervals have E_3 less than 7.5 km/h. This is less than the intervals in the first category for E_1 and E_2.

The second category consisted of 18.3 % of the 15-minute intervals and 23.3 % of the 1-hour intervals, which is greater than a proportion of about 10 % that were observed for the second category for both E_1 and E_2 (Figure 4.14). The above 15 km/h category consisted of less than 4 % of the 15-minute and 1-hour intervals.

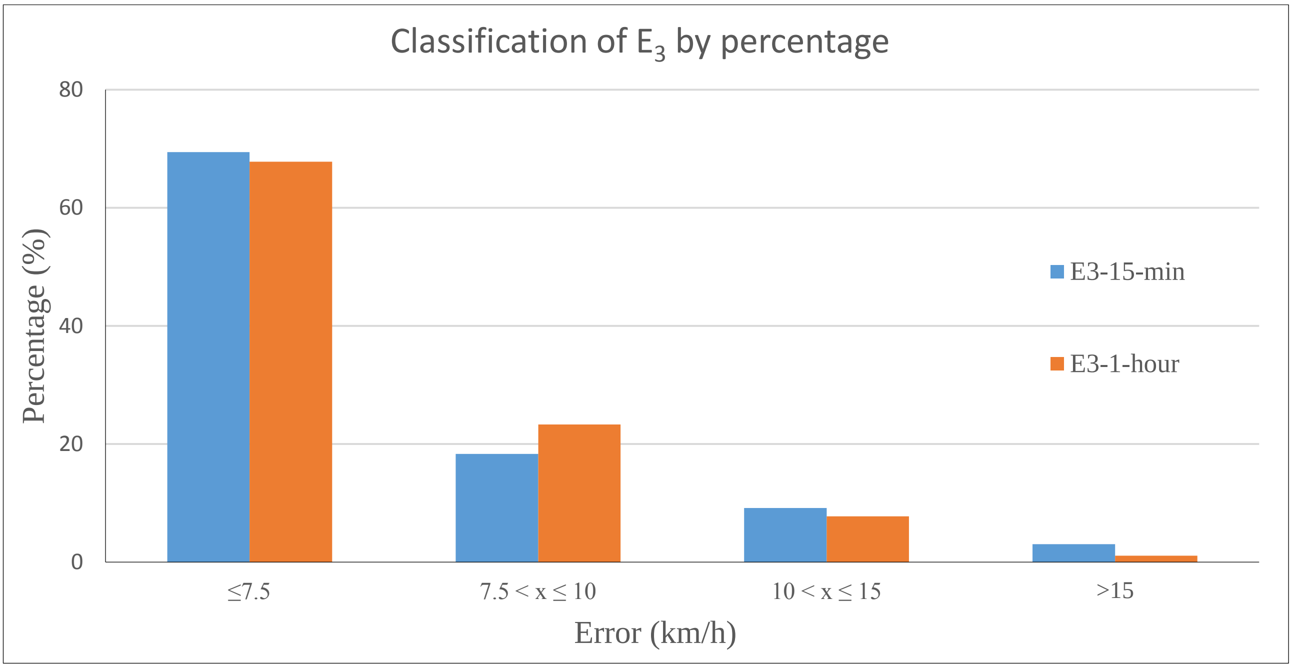


Figure 4.14: Classification of E_3 by percentage

4.5.4 CUMULATIVE AND FREQUENCY DISTRIBUTIONS

From the speed profiles, it was established that the TomTom speeds were generally lower than reference speeds. In trying to understand how TomTom speed estimates compare to the reference speeds, it is important to have an idea of the proportion of intervals with a bias. To address this issue, Figure 4.15 shows the frequency distribution curves of the speed bias for both the 15-minute and 1-hour intervals. Speed bias was the difference between TomTom speed estimates and reference speeds.

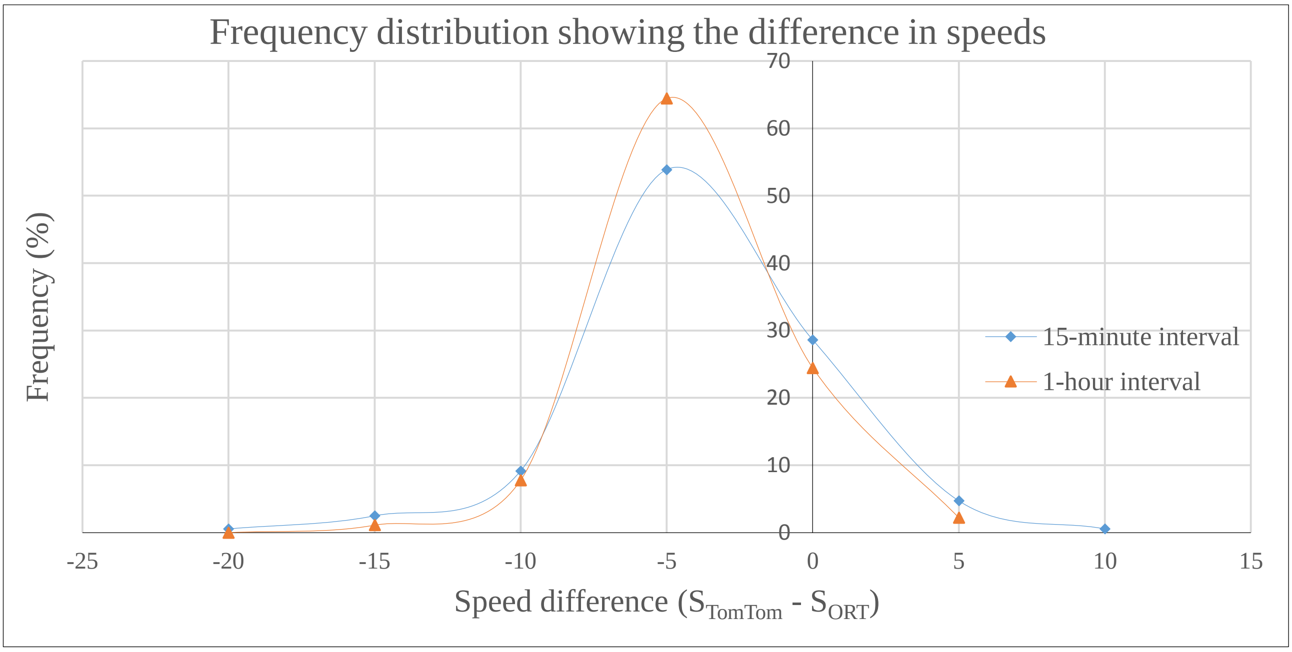


Figure 4.15: Frequency distribution of the differences in speeds

The differences between the two datasets ranged from -20 km/h to 10 km/h. In addition, the differences were negative which confirms that the TomTom speeds were lower than reference speeds. Although the frequency distribution curves are slightly different, both curves peak at the -10 km/h to -5 km/h speed bin, with 53.9 % of the 15-minute speeds and 64.4 % of the 1-hour interval speeds belonging to this speed bin.

The cumulative frequency distribution illustrating the difference between TomTom speeds and reference speeds is shown in Figure 4.16. In 94.7 % (341 of 360) of the 15-minute intervals, the TomTom speeds were lower than the corresponding reference speeds. TomTom speed estimates were lower than the corresponding reference speeds in 97.8 % (88 of 90) of the 1-hour intervals.

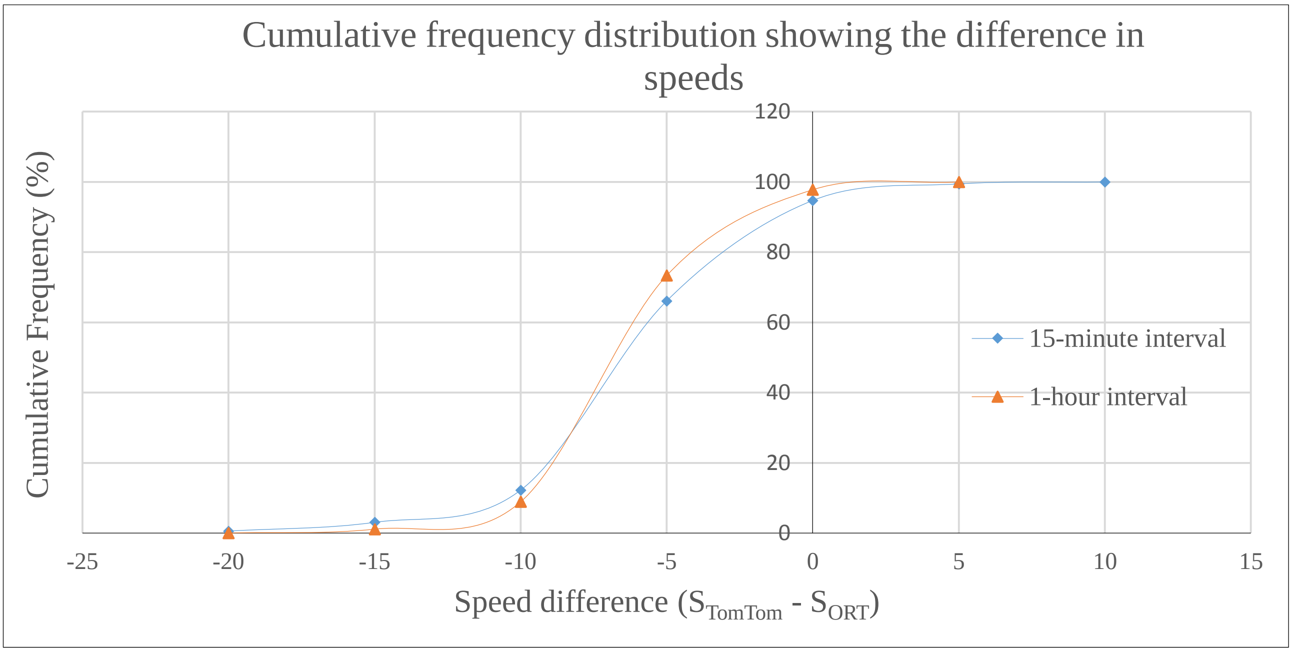


Figure 4.16: Cumulative frequency distribution of the differences in speeds

4.6 COMPLETENESS

Completeness measured the degree to which speed data was present. For all the 15-minute and 1-hour intervals, speeds in the particular intervals were present. Therefore, completeness was 100 %.

Table 4.4: Determination of completeness for 15-minute and 1-hour intervals

Time interval	Total expected	Available values	Percent complete
15-minute interval	360	360	100
1-hour interval	90	90	100

4.7 VALIDITY

Validity measured the degree to which the speed data satisfied the specified acceptance criteria. Three error quantities that were calculated were E_1, E_2 and E_3, as described in section 3.7.3. Table 4.5 and Table 4.6 show the percent valid calculation for the 15-minute and 1-hour intervals, respectively.

Table 4.5: Determination of validity for a 15-minute interval

Error	**Total**	**Valid**	**Percent valid**
E_1	360	311	86.4
E_2	360	316	87.8
E_3	360	250	69.4

Table 4.6: Determination of validity for a 1-hour time interval

Error	**Total**	**Valid**	**Percent valid**
E_1	90	80	88.9
E_2	90	82	91.1
E_3	90	61	67.8

For both the 15-minute and 1-hour intervals, well over 85 % of the intervals resulted in E_1 below the maximum allowable E_1, of ±10 %. A similar result was also obtained for E_2, where over 88 % of the errors, E_2, were less than the allowable error of 10 km/h for the two time intervals. The results obtained show that E_3 was the most challenging criteria to meet seeing that only less than 70 % of the errors were less than the allowable error of 7.5 km/h, for both the 15-minute and 1-hour time intervals.

4.8 COVERAGE

Coverage measured the extent to which the freeway segments investigated represent the population of the freeway segments. The total length of all the freeway segments was 301.12 km. The six freeway segments investigated had a total length of 59.54 km, resulting in a percent coverage of 19.8 %.

4.9 ACCESSIBILITY

Unlike the other quality measures, accessibility was measured by means of a qualitative assessment. For this study, accessibility focused on the usability of the data. Accessibility measured the relative ease with which data was retrieved, processed and reduced. TomTom data was already processed and no further manipulation was required. Therefore, the data was easily accessible.

The challenges of processing and reducing the ORT data were extensively outlined. ORT data was raw and no useful traffic information could be obtained from the raw data. It was not possible to process or manipulate the data using packages such as MS Excel. A Java program was developed to process and reduce the data (see section 3.6). Therefore, ORT data was not easily accessible.

4.10 SUMMARY

The results of the five quality measures investigated in this study are summarised in a traffic data scorecard. The quality measures investigated are accuracy, completeness, validity, coverage and accessibility. A traffic data scorecard for the data quality assessment is shown in Table 4.7.

Table 4.7: Data quality scorecard for Gauteng freeway segments, adapted from Battelle *et al.* (2004)

Data quality measures	**15-minute interval**	**1-hour interval**
Accuracy		
Signed error	-5.8 %	-6.2 %
AASE	6.4 km/h	6.5 km/h
SEB	-6.2 km/h	-6.3 km/h
Completeness		
Percent complete	100 %	100 %
Validity		
Percent valid		
Signed error	86.4 %	88.9 %
AASE	87.8 %	91.1 %
SEB	69.4 %	67.8 %
Coverage		
Percent coverage	19.8 %	19.8 %
Accessibility		
Qualitative assessment	TomTom data was already processed hence no further manipulation was required. ORT was raw and a Java program was developed to process the data.	

4.11 CONCLUSION

The results of the study were presented in this chapter. From the results, TomTom speed estimates were within the allowable errors. In addition, there were no significant differences between the error quantities for speeds derived from the 15-minute and 1-hour interval speeds. TomTom historical average speeds were consistent with the reference speeds, albeit, the service data provider generally underestimates traffic speeds. This resulted in the rejection of the null hypothesis (at a 95 % confidence level) that the TomTom mean speeds for each freeway segment were equal to the ORT mean speeds, at both the 15-minute and 1-hour intervals. Validity of the speed estimates did not only depend on the error quantity but also on the set threshold value of the error quantity.

CHAPTER 5 : DISCUSSION

5.1 INTRODUCTION

The results of this study were presented in Chapter 4 and this chapter discusses these results. Although the quality of TomTom historical data was consistent with the reference speeds, it was observed that the service provider generally underestimates traffic speeds. Possible reasons for the low speed estimates are explored and discussed. In order to assess the quality of TomTom historical data, the quality measures that were investigated were accuracy, completeness, validity, coverage and accessibility. A discussion on the results of these quality measures is also presented. To end the chapter, the limitations and outstanding issues are discussed.

5.2 SPEED PROFILES

The speed profiles, for both the 15-minute and 1-hour intervals, were interesting in a number of ways. As expected, there was a morning peak on the N1 Ben Schoeman southbound section (Figure 4.1) and an evening peak on the northbound section (Figure 4.2). This suggests that there was heavy traffic in the southbound direction from Pretoria to Johannesburg in the morning and heavy evening traffic in the northbound direction from Johannesburg to Pretoria. The speed profiles confirm the directional origin-destination pattern between Pretoria and Johannesburg.

A well-defined morning peak on the N1 Western Bypass northbound section (Figure 4.4) is due to the high trip productions and attractions in the Sandton, Bryanston and Randburg area. In addition, this segment also carries traffic travelling north to the N1 and N3 freeways. The morning and evening peaks on the N1 Western Bypass southbound section (Figure 4.3) are in agreement with the bi-directional travel pattern that is expected for a ring road. The N1 Western Bypass, together with N3 and N12, form the ring road around Johannesburg, as depicted on the Gauteng freeway network map in Chapter 3 (Figure 3.1).

The R21 is one of the only freeways in Gauteng that remains relatively uncongested, even during peak periods. Speeds over 100 km/h were experienced throughout the analysis period for both the R21 sections. This is because there is plenty of reserve capacity on the R21 compared to the N1 sections, which are typically operating close to capacity.

5.3 ACCURACY

A discussion on the results of the hypothesis tests, error quantities at both the freeway segment level and interval level is presented in the sections that follow.

5.3.1 HYPOTHESIS TESTS

The quality of TomTom historical average speeds was consistent with the reference speeds, albeit, the service data provider generally underestimates traffic speeds. This resulted in the rejection of the null hypothesis (at a 95 % confidence level) that the TomTom mean speeds were equal to the ORT mean speeds, at both the 15-minute and 1-hour intervals.

There were significant differences between the mean TomTom speeds and mean ORT speeds at both the 15-minute and 1-hour intervals. Research shows that speed is a distribution and with increasing variation in the speeds during a given time interval, a point estimate speed may not be sufficient in estimating the speed for that time interval (National Cooperative Highways Research Program, 2009).

The speeds in each 15-minute and 1-hour interval were aggregated to a single representative speed for that time interval. This resulted in the speeds from the two data sources being significantly different. A better method of resolving this issue is to view the speeds in a given time interval as a distribution, rather than a point estimate.

5.3.2 ERROR QUANTITIES AT A FREEWAY SEGMENT LEVEL

The results of the accuracy measures suggest that TomTom historical speeds were consistent with the reference speeds. Accuracy measures were investigated by means of three error quantities, namely *signed error, average absolute speed error (AASE)* and *speed error bias (SEB)*. The *error quantities* for the combined freeway segments were less than the respective allowable errors for speeds derived from the 15-minute and 1-hour time intervals.

The *signed error*s for the combined freeway segments, derived from 15-minute and 1-hour intervals, were -5.8 % and -6.2 %, respectively (Table 4.2). For both time intervals, the *signed errors* were less than the allowable error of ±10 %. The allowable *AASE* and *SEB* were 10 km/h and ± 7.5 km/h, respectively. The *AASE* for the combined freeway segments were 6.4 km/h and 6.5 km/h for the 15-minute and 1-hour intervals, respectively. In addition, *SEB* for the 15-minute interval speeds and 1-hour interval speeds were -6.2 km/h and -6.3 km/h, respectively. This demonstrates that the accuracy measures were within the allowable limits. Therefore, in terms of accuracy for the combined freeway segments, TomTom historical speed estimates were consistent with the reference speeds.

The findings for the *signed error, AASE* and *SEB* at an individual freeway segment level are shown in Figure 4.9, Figure 4.10 and Figure 4.11. For both the 15-minute and 1-hour intervals, the *signed error* and *AASE* were less than their respective allowable errors for all the six freeway segments. Somewhat surprisingly, the *SEB* for one of the six freeway segments was more than the allowable error of ±7.5 km/h. In fact, the *SEB* for the R21 northbound section between Letata (1042) and Heron (1039) was -8.0 km/h, for both the 15-minute and 1-hour intervals.

The Texas Transportation Institute evaluated the accuracy of TomTom historical speed data for 11 directional segments on the US 290 in Houston, Texas (Texas Transportation Institute, 2012). In the study, the average annual hourly speeds and travel times were evaluated for the weekdays in 2010. The findings of the study show that AASE less than 5 mph (8 km/h) were obtained in 10 of the 11 segments. These results are in agreement with the findings of this study. For the six freeway segments investigated, the maximum and minimum AASE for the 1-hour interval were 8.0 km/h and 4.9 km/h, respectively (Figure 4.10). The AASE for the 15-minute interval were of similar magnitude.

The I-95 Corridor Coalition Project evaluated probe data from Inrix on approximately 1500 miles of freeways and 1000 miles of arterials in six states of the USA (Hni *et al.*, 2009, 2013; Inrix, 2015). The two accuracy measures that were used to evaluate the probe data were AASE and SEB, with the allowable errors of 10 mph (16 km/h) and ± 5 mph (± 8 km/h), respectively. The study concluded that Inrix data across the system and by individual state generally satisfied the specified accuracy specifications i.e. AASE and SEB less than 10 mph and ± 5 mph, respectively. These results are comparable with the finding of this study.

Interestingly, there were no statistically significant differences between the *signed error*, *AASE* and *SEB* for the 15-minute and 1-hour interval speeds for the combined freeway segments (Table 4.2). Furthermore, no statistically significant differences were found between the three *error quantities* (*signed error*, *AASE* and *SEB*) for speeds derived from 15-minute and 1-hour intervals for the individual freeway segments.

The fact that there were no significant differences between the error quantities for the 15-minute and 1-hour intervals should not be surprising. TomTom probes in a 15-minute interval, say from 08:00-08:15, were also included in the 1-hour interval from 08:00-09:00. In a similar way, the probes for the 08:15-08:30, 08:30-08:45 and 08:45-09:00 intervals were also included in the 08:00-09:00 interval. In addition, there were a few additional probes which overlap between two or more 15-minute intervals that were also included in the 1-hour interval but not in the individual 15-minute intervals. As a result, the probes in the 1-hour interval were generally more than the probes in four 15-minute intervals in that hour. In addition, the reference observations in four 15-minute intervals equal the observations in the 1-hour interval (Table 3.5). For the 15-hour analysis period that was considered, the average speed for each freeway segment was the average of sixty 15-minute speeds and fifteen 1-hour speeds for the 15-minute and 1-hour intervals, respectively. For this reason, the average speeds derived from a 15-minute interval were not expected to be significantly different from the average speeds derived from a 1-hour interval.

5.3.3 ERROR QUANTITIES AT AN INTERVAL LEVEL

For the combined freeway segments, the three error quantities (E_1, E_2 and E_3) were determined at both the 15-minute and 1-hour interval level. For E_1 and E_2, over 85 % of the speeds resulted in errors within their respective allowable limits. This was expected because the *signed error* and *AASE* for the individual and combined freeway segments were less than the allowable errors.

Quite surprisingly, less than 70 % of the speeds resulted in errors within the allowable error of ± 7.5 km/h and yet, the SEB for the combined freeway segments was within the ± 7.5 km/h limit. It is possible that the threshold for E_3 was simply more challenging to satisfy compared to the thresholds for E_1 and E_2. The implication of the different error quantities and thresholds are discussed in the section that follows (section 5.4).

5.4 VALIDITY

Validity was determined at both the 15-minute and 1-hour intervals to understand how the TomTom speed estimates compared to the reference speeds. The errors, E_1, E_2 and E_3 are the signed error, AASE and SEB for one 15-minute or 1-hour interval. Validity is the degree to which data values satisfy the requirements of the specified criteria. The individual interval speeds that fell in the different error categories were shown in Chapter 4 (Figure 4.12, Figure 4.13, Figure 4.14).

More than 85 % of the speeds resulted in E_1 and E_2 less than their respective allowable errors. This was observed for both the 15-minute and 1-hour intervals. Hence, based on E_1 and E_2, TomTom speed estimates were of very high quality (Table 3.11). Considering the error quantity, E_3, slightly less than 70 % of the speeds resulted in an error within the allowable limit of ±7.5 km/h (Figure 4.14). This was observed for both the 15-minute and 1-hour intervals. Based on E_3, TomTom historical data was described as moderate quality data (Table 3.11).

To bring this into perspective, the results of E_1 and E_2 suggest that the TomTom historical speed data was of very high quality whereas E_3 defined the same data as moderate quality data. It is clear that validity is dependent on the selected measures and the set thresholds. Furthermore, the existing guidelines on evaluating the quality of travel time and speed data do not establish the thresholds for traffic data to be valid. The user defines these thresholds based on their intended application.

Therefore, data that is valid for one application might not be valid for another application. It is worthwhile to take caution when developing the validity criteria and thresholds to use since this can possibly affect the results (Turner *et al.*, 2011). In addition, Battelle *et al.* (2004) explained that the data falling outside the set thresholds reflect one of two things. Firstly, the data could be unacceptable for that intended use and secondly, the data ought to be used with caution.

5.5 OTHER QUALITY MEASURES

In this section, a brief discussion on the quality measures, coverage, completeness and accessibility, is presented. Not much information was obtained from these quality measures on their own. However, these quality measures help in reinforcing meaning and value of accuracy and validity, and in some cases, describe certain attributes of the data.

The 100 % *completeness* that was achieved for both the 15-minute and 1-hour intervals simply highlights that the sample sizes in the individual time intervals were more than the required minimum sample sizes. At that confidence level, the TomTom speeds were regarded as “reliable”.

In terms of *coverage*, 59.54 km of the 301.12 km were investigated, which represents about 20 % of the freeway segments. For the 20 % of the freeway segments that were investigated, there is confidence that the TomTom speed estimates were accurate and within the specified thresholds. Other freeway segments with the same characteristics are expected to have the errors of a similar magnitude.

Turning now to *accessibility*, TomTom historical data was easy to retrieve, manipulate and understand. Furthermore, the fact that the data is relatively cost-effective and does not require infrastructure-based devices for data collection makes it an attractive alternative for data consumers.

5.6 FACTORS AFFECTING ACCURACY

The two factors that varied for the different freeway segments were *length* and *sample size* of the TomTom data. In trying to understand the effect of these two variables on the accuracy measures, the relationship between the length of the freeway segment, sample size and the accuracy measures was investigated.

The relationship between the error quantities and the length of segment is investigated in Figure 5.1. In earlier sections, it was found that there were no significant differences between the error quantities for 15-minute and 1-hour interval speeds. With this in mind, only accuracy measures at the 15-minute interval are shown in Figure 5.1. There is no evidence to show a strong relationship between the length of segment and accuracy measures. Furthermore, there is weak to moderate correlation between the length of segment and the three errors quantities.

Figure 5.2 shows the variation of the accuracy measures with the sample size. Again, the errors shown in Figure 5.2 were derived from the 15-minute interval speeds. Similar to what was observed between the length of the segment and error size, there is not enough evidence to suggest that there is a strong relationship between the accuracy measures and sample size. Additionally, there is weak correlation between the three error quantities and the sample size of the probe data.

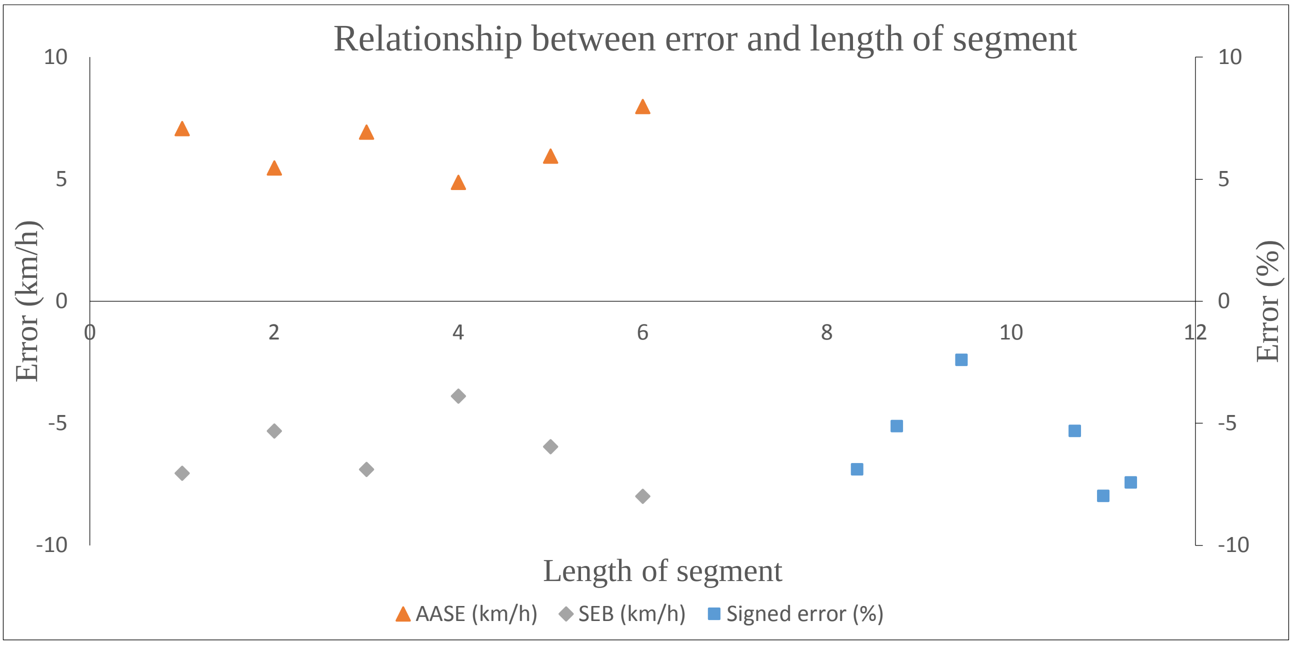


Figure 5.1: Relationship between error and length of segment

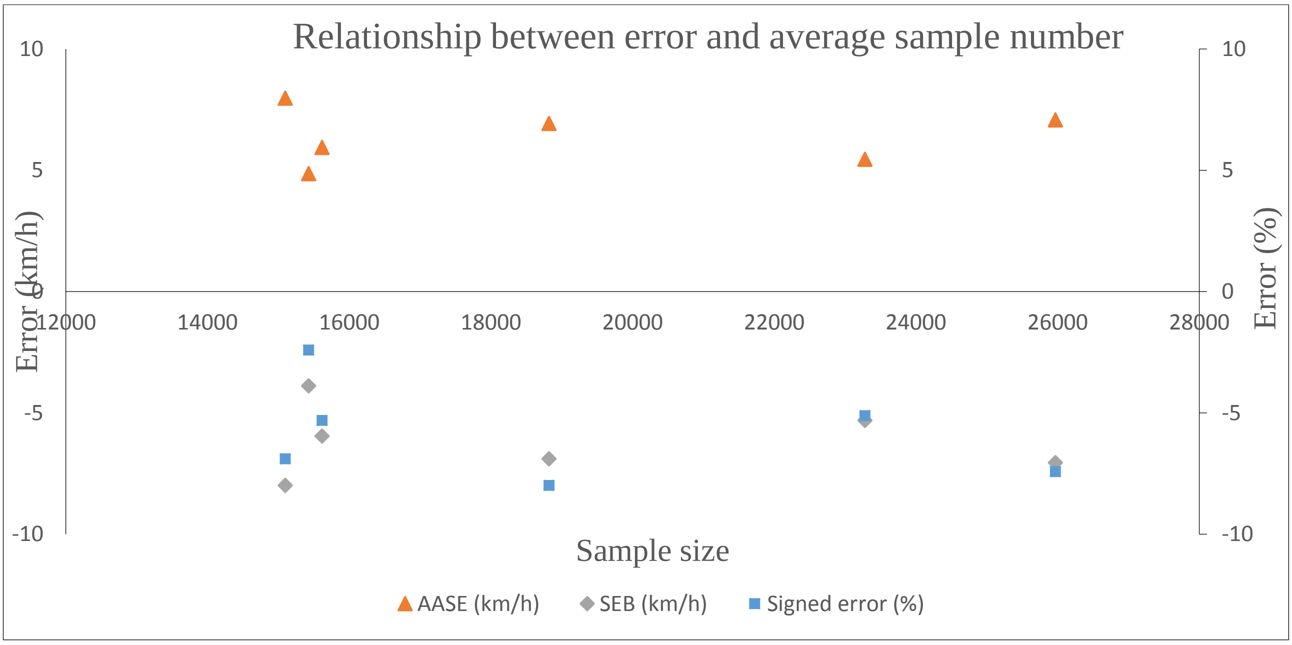


Figure 5.2: Relationship between size of error and average sample number

5.7 IMPLICATIONS OF THE UNEQUAL LINK LENGTHS ON SPEEDS AND ACCURACY

The ORT link lengths and TomTom link lengths were not the same on all the six segments (Table 3.2). For all the six freeway segments, the TomTom link lengths were greater than the ORT link lengths. The reason why the TomTom link lengths were longer than ORT link lengths is due to the network segmentation by the service provider, as explained in section 3.4.3.2.

Figure 5.3 illustrates TomTom sub-segmentation for the R21 northbound section between Letata (1042) and Heron (1039). The reference link length is 8.3 km whereas the TomTom link length is 9.7 km. The TomTom link consists of 13 sub-segments of different lengths.

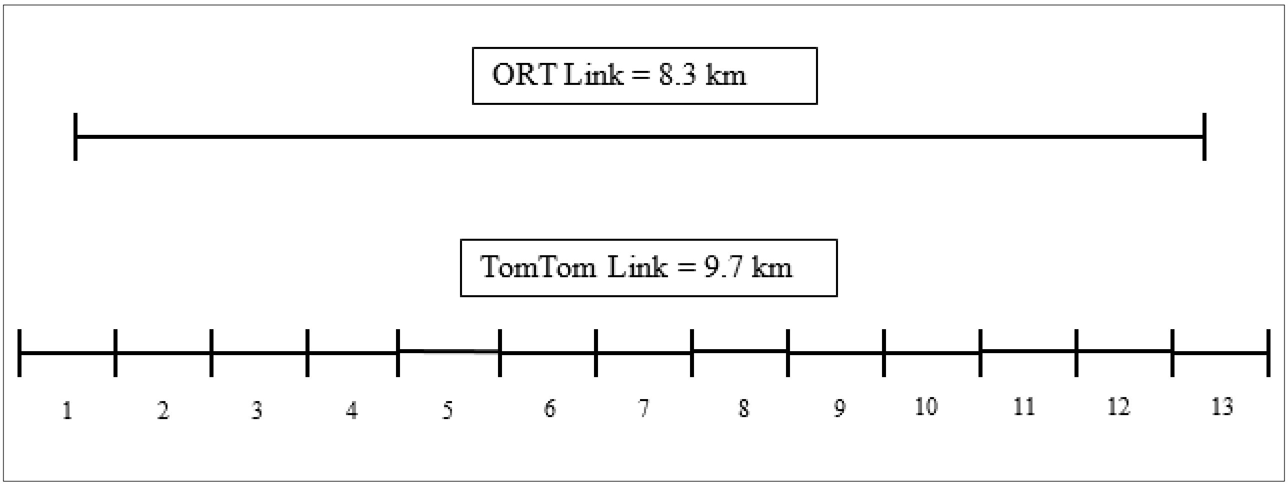


Figure 5.3: Illustration of TomTom and ORT links

To understand the implications of the different ORT and TomTom link lengths, Table 5.1 shows the different TomTom lengths from a combination of different sub-segments and the resultant average speeds for the 1-hour interval between 16:00 and 17:00. The differences in the average speeds shown in Table 5.1 are less than 0.7 km/h (0.6 %).

This suggests that differences between the ORT and TomTom link lengths only have a minor effect on the average speeds. Therefore, it is highly unlikely that the low TomTom speed estimates were due to unequal lengths of the service provider and reference links.

Table 5.1: Average speeds of the different combinations of TomTom sub-segments

Segment	**TomTom Link (km)**	**Average Speed (km/h)**	**% difference in speed**
Segment 1-13	9.7	111.8	-
Segment 2-13	8.2	112.1	0.3 %
Segment 1-12	8.8	112.0	0.2 %
Segment 2-12	7.3	112.4	0.6 %

5.8 EXPLANATION OF THE OBSERVED TRENDS

TomTom speed estimates were consistently lower than the reference speeds on all the six freeway segments throughout the analysis period. This trend was observed on all the six freeway segments, regardless of the time interval used in aggregating the speeds (Figure 4.1, Figure 4.2, Figure 4.3, Figure 4.4, Figure 4.5, Figure 4.6 and Appendix A).

The differences in the speeds varied with the traffic conditions on the road. It appears as if the differences between the TomTom and reference speeds in the peak periods were less compared to the differences in the off-peak periods. In addition, it might be the case that the differences between TomTom and reference speeds were smaller when the freeway segments were operating at low speeds and greater during high speed operating conditions. This was observed on all the N1 sections. Because there were no peaks on the R21 sections, the differences between the TomTom speed estimates and reference speeds on both sections were greater but constant throughout the analysis period.

It should be noted that the error quantities, *signed error* and *SEB*, do not reveal the variation of the differences in speeds since these were aggregated to obtain a single speed measurement to represent the freeway segments that were investigated. The speed profiles reveal that the differences between the two datasets were not simply a matter of a constant offset in the speeds.

It was observed that the N1 sections showed small differences between the TomTom speed estimates and the reference speeds during the peak periods, i.e. morning peak in Figure 4.1, evening peak in Figure 4.2, morning and evening peak in Figure 4.3 and morning peak in Figure 4.4. Notable differences in speeds from the two sources were observed during the off-peak periods for these sections. It appeared to be the trend. However, a peak period is defined in terms of traffic volume and not speed. In addition, there was weak correlation between the error quantities and the vehicles on the freeway segments.

The other characteristic that is also true and common for peak periods are the low speeds. Speeds are typically less than free-flow speeds during peak periods. The increase in traffic on the road results in the decrease in speeds. What was observed on the speed profiles was that the TomTom speeds were better estimates during periods characterised with low speeds. In other words, differences between the two datasets were less when the freeway segment was operating at low speeds. This is shown during certain times of the day on the N1 sections.

Furthermore, greater differences in the speeds were observed when the freeway segments were operating at high speeds. This was the case for the two R21 sections during the entire analysis period and the N1 sections during the off-peak periods. It is possible that the reason why the R21 northbound section between Letata (1042) and Heron (1039) had a *SEB* greater than the allowable limit was that the highest speeds were recorded on this freeway segment.

Another way to illustrate that differences between the TomTom and reference speeds varied with the operating speeds on the freeway segment was shown in Figure 4.7 and Figure 4.8. At lower speeds, the data points were closer to the line of correlation, whereas at high speeds, the data points were further away from the line of correlation.

5.9 REASONS WHY PROBE DATA ESTIMATES ARE BETTER AT LOW SPEEDS

Drivers have the flexibility to choose their desired speeds during off-peak and free-flow conditions. During these periods, high speed differentials occur, i.e. the difference between the highest and lowest speeds on the segment is high.

During the peak periods when the freeway is congested, the driver does not have the flexibility to travel at their desired speeds but have to adjust their speeds in relation to the speeds of other drivers. This is also experienced during incidences, crashes or lane closures. As a result, the speed differentials are low. This was proved in the Highway Capacity Manual (National Research Council, 2010).

The differences between the mean TomTom speeds and reference speeds are expected to be low during the time of day characterised with low speed differentials (peak hours). High differences between the two datasets are expected during the off-peak periods, which are typically characterised by high speed differentials.

It appears as if it is more challenging for the service provider to accurately estimate the traffic speeds during periods categorised with high speed differentials. It is surmised that the quality of the TomTom speed estimates decreases with an increase in the speed differentials on the freeway segment.

5.10 POSSIBLE EXPLANATIONS FOR THE LOW TOMTOM SPEED ESTIMATES

The outstanding issue is still to understand why the TomTom speed estimates were lower than the reference speeds. It is highly unlikely that the information obtained from the GPS devices is inaccurate because GPS is a mature technology that has been extensively tested. Rather, it is more likely that the sample that was used by the service provider in deriving the speed estimates was not a true representation of the traffic stream.

Although TomTom data is derived from many sources, a considerable portion of the data comes from GPS devices and commercial fleet services (Figure 3.3). Heavy vehicles typically travel at slower speeds compared to passenger cars (Hallmark, 2004). A sample that primarily consists of heavy vehicles is likely to result in average speed that is less than the reference speed.

Another small contributing factor has to do with GPS device users. Drivers who use GPS devices on the freeways are likely to be unfamiliar with the area. A GPS device is essential for navigation purposes. The Highway Capacity Manual introduced the concept of driver's familiarity with the facility (National Research Council, 2010). Drivers who are not familiar with a certain freeway facility are likely to drive at slower speeds compared to regular drivers.

In light of the above discussion, the reason why the service provider underestimates traffic speeds is possibly due to selection bias. Simply put, the aggregated TomTom speeds were derived from a sample consisting of mainly unfamiliar drivers and data from the commercial services (heavy vehicles). The speeds were derived from a biased sample that was not the true representation of the traffic population.

However, assuming that GPS users mainly consist of drivers who are not familiar with the facility should not be overstated. There is a renewed interest in the traffic conditions on the road during the course of a trip. A considerable proportion of drivers want to know if there are any incidences during the course of their journey to avoid delays, thus minimising trip times. As a result, a growing proportion of drivers use GPS devices for incident detection as opposed to navigation purposes. This highlights how erroneous it is to assume that drivers that use GPS devices are unfamiliar with the freeway facility.

5.11 LIMITATIONS

Some reference observations were filtered and discarded because they were deemed undesirable for reasons highlighted in the data-filtering plan (section 3.5.1.4). Perhaps, it is possible that eliminating these "unacceptable" observations distorted the "reference" speeds such that the speeds were no longer an accurate representation of the reference data. However, the data processing plan was based on the international guidelines for evaluating the traffic data quality, namely, Haghani *et al.* (2009), Turner *et al.* (2011) and Battelle *et al.* (2004). Therefore, the possibility of distorting the reference speeds by filtering out unacceptable observations is highly unlikely.

Furthermore, about 200 000 observations of the 70 million had missing IDs and it was not possible to use these observations in computing reference speeds. These observations were discarded. The discarded observations might have led to the differences in speeds. The observations with missing IDs were identified across all the gantries and not a specific area. It is likely that the 200 000 observations (which is about 0.28 % of the total observations) had no significantly influence on the speed differences that were reported.

The freeway segments that were investigated had interchanges, on-ramps or off-ramps at some point along its length. The effect of weaving, merging and diverging segments was not investigated. It was not possible to know what happens along the length of the freeway segment because vehicles were identified and re-identified at the two toll gantries that define a freeway segment.

The incidents that occurred on the freeways during the evaluation period were not taken into account. Incidents occur often on the freeways, with some of the incidents causing major delays, lane closures and in worst scenarios, a total standstill of traffic. Furthermore, none of the invalidating factors highlighted by Midwestern Consulting (2008), such as the weather conditions and absence of construction, were investigated. One must then question how the data service provider deals with these issues. With the processing of the ORT data, the data filtering was developed to address the low speeds regardless of the cause.

The major limitation associated with aggregating the speeds for the weekdays (Monday – Friday) is the fact that daily variations cannot be observed on the speed profiles. Studies for a single day need to be investigated to find out if similar quality measures would be achieved. The question that arises is "How accurate are the TomTom speed estimates on a given randomly chosen day?" The work that was done up to this point did not address this issue.

5.12 OUTSTANDING ISSUES

A few outstanding issues were identified. These issues were related to the limitations in the methodology, understanding if the proportion of heavy vehicles has an effect on the accuracy of the TomTom speed estimates and the correction of TomTom speeds so that the speed estimates are as close to the references speeds as possible. In addition, there is a need to investigate the how accuracy measures relate to the speeds at which the freeway segment is operating. The outstanding issues are summarised as follows:

i. The relationship between accuracy and speed.
ii. The effect of daily variation of speeds on accuracy.
iii. The effect of heavy vehicles on the accuracy of TomTom speed estimates.
iv. The use of percentile speeds as a way of correcting the TomTom average speed estimates.

5.13 CONCLUSION

The discussion of the results was presented in this chapter. Although the TomTom historical speeds were consistent with the reference speeds, the data service provider generally underestimates traffic speeds. The low speed estimates were possibly derived from a sample that is not a true representation of the traffic stream i.e. a sample consisting mainly heavy vehicles and drivers unfamiliar with the freeway facility. The length of the freeway segments and the average sample size did not affect the accuracy measures. The chapter concluded by highlighting a few outstanding issues that need to be addressed.

CHAPTER 6 : FURTHER ANALYSIS

6.1 INTRODUCTION

A further discussion and analysis of TomTom historical data is presented in this chapter. The chapter begins by comparing the results of this study and the findings from other data quality assessments. In addition, the outstanding issues identified in Chapter 5 are also discussed. The outstanding issues include the daily variation of speeds, correction of TomTom historical average speeds, effect of heavy vehicles on accuracy and the relationship between accuracy measures and traffic speeds.

6.2 COMPARISON WITH INTERNATIONAL RESEARCH

Other studies have also found similar results in terms of the trend in the differences between probe data estimates and reference speeds. Figure 6.1, Figure 6.2, Figure 6.3 and Figure 6.4 show the speed profiles for four sections on the US 290, a heavily congested freeway (Texas Transportation Institute, 2012). The speed profiles show how closely TomTom estimated the traffic speeds.

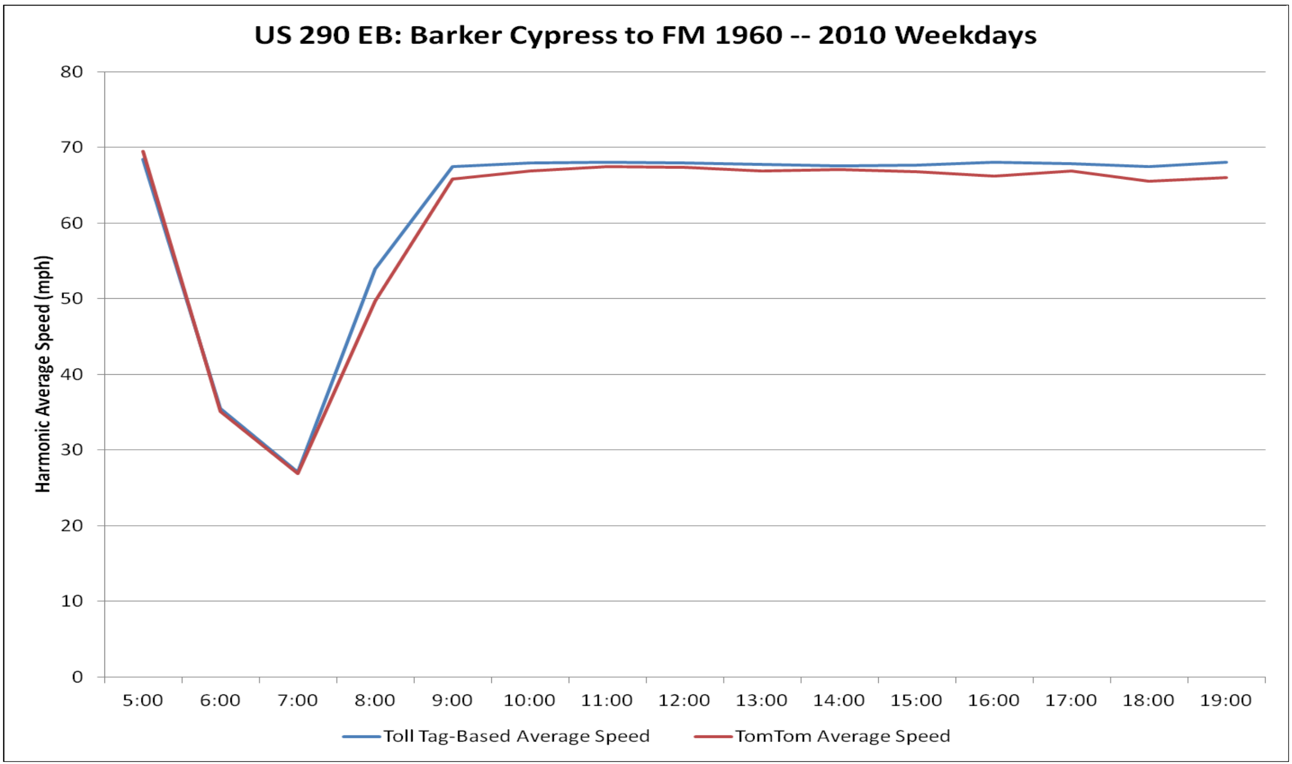


Figure 6.1: Speed profile on the US 290 Eastbound: Barker Cypress to FM 1960 (Texas Transportation Institute, 2012)

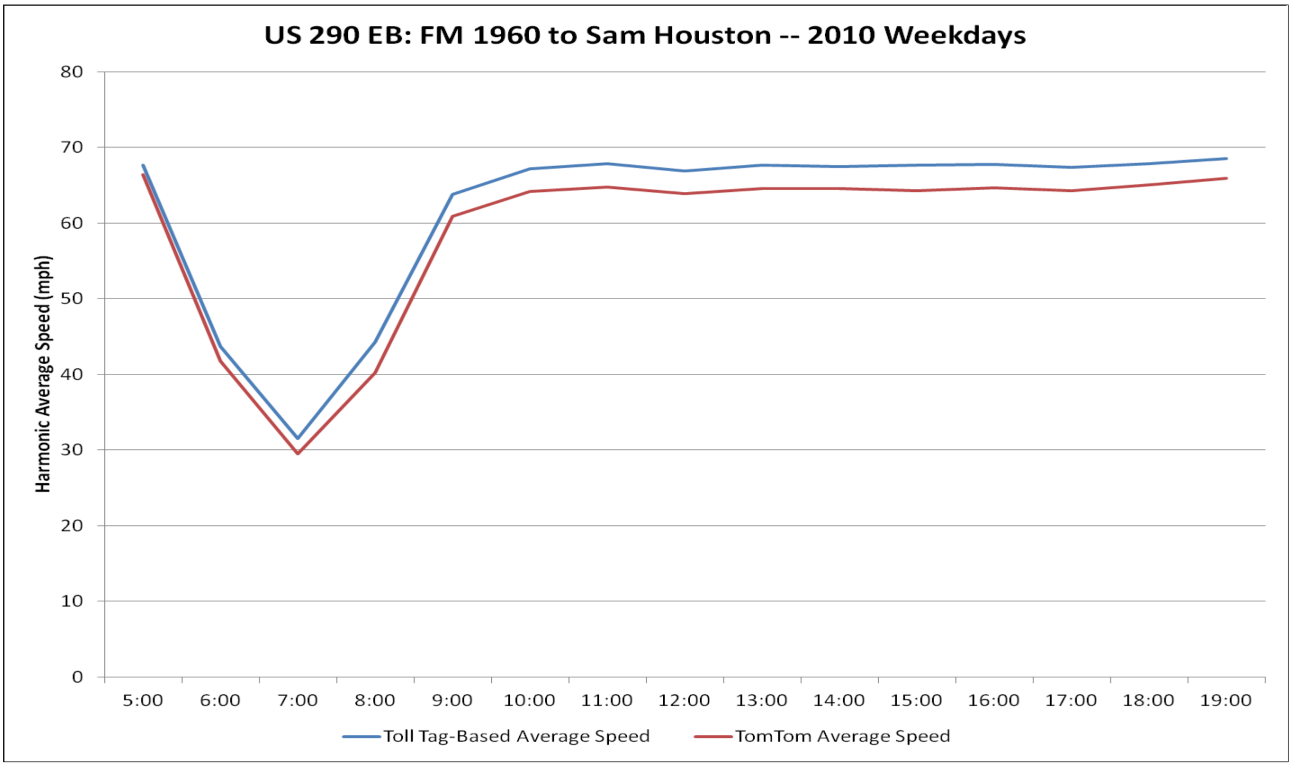


Figure 6.2: Speed profile on the US 290 Eastbound: FM 1960 to Sam Houston (Texas Transportation Institute, 2012)

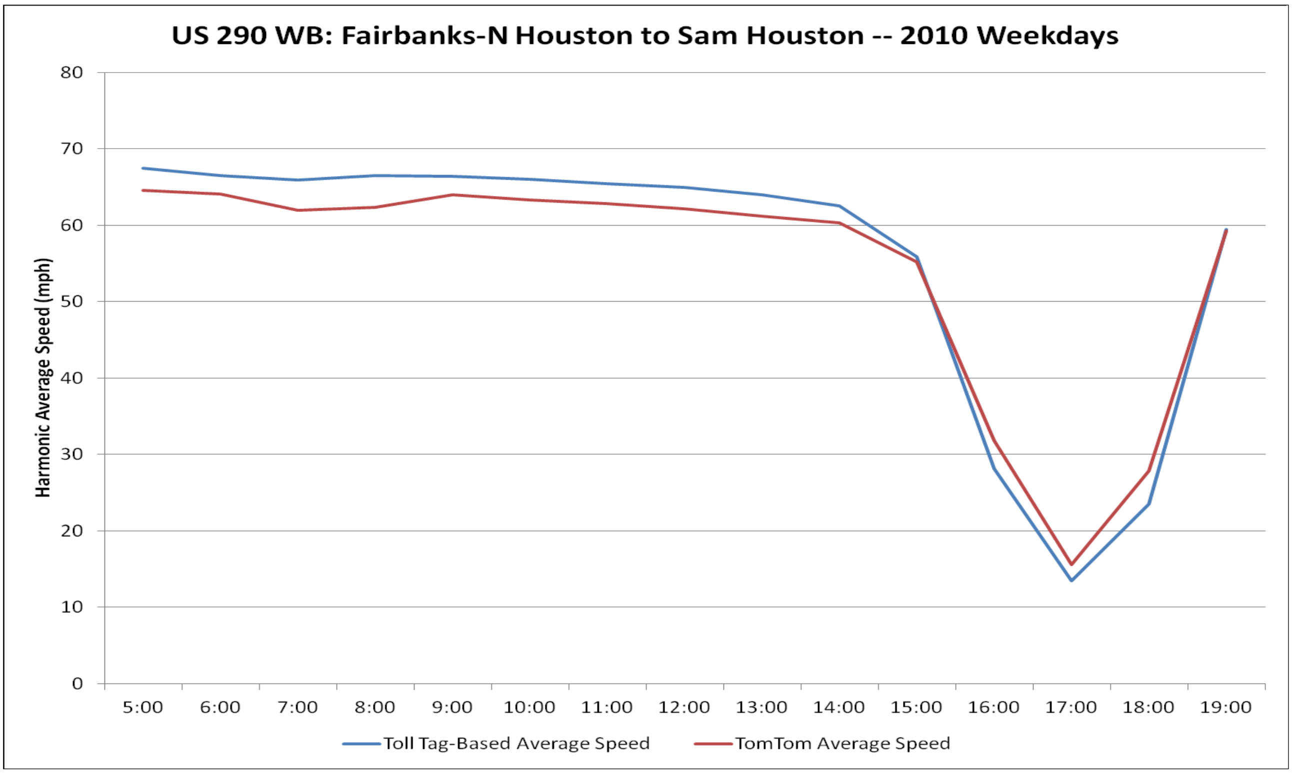


Figure 6.3: Speed profile on the US 290 Westbound: Fairbank-N Houston to Sam Houston (Texas Transportation Institute, 2012)

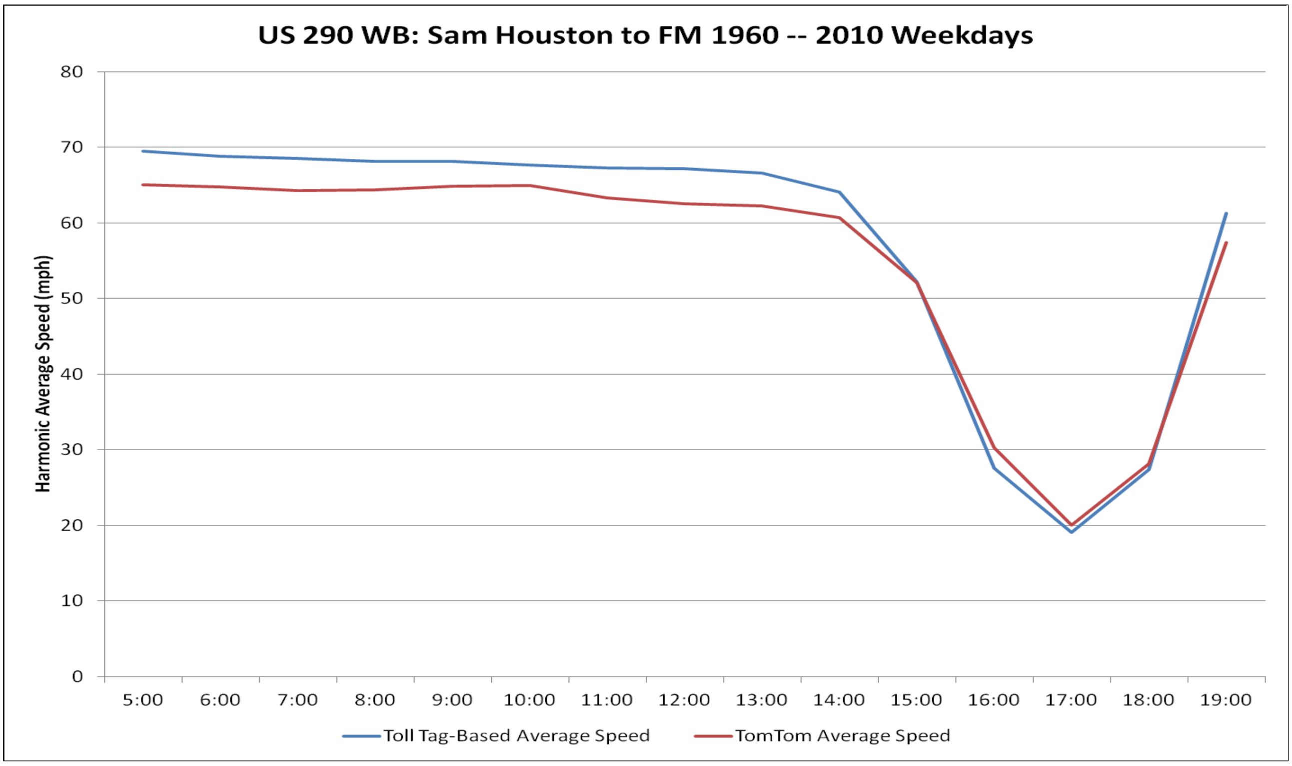


Figure 6.4: Speed profile on the US 290 Westbound: Sam Houston to FM 1960 (Texas Transportation Institute, 2012)

The trends observed on the US 290 speed profiles were similar to the trends observed on the six freeway segments investigated in this study (Figure 4.1, Figure 4.2, Figure 4.3, Figure 4.4, Figure 4.5, Figure 4.6 and Appendix A). The four speed profiles on the US 290 show that the TomTom speed profiles were below the reference speed profiles for most of the analysis period. This shows that TomTom speed estimates were lower than the reference speeds on these sections. As previously discussed in Chapter 5, the AASE that were obtained for this study were of similar magnitude as the AASE on the US 290 freeway segments.

Furthermore, the differences between the TomTom speeds and reference speeds were not constant throughout the analysis period. The US 290 speed profiles suggest that TomTom estimates were better when the speeds on the freeway segments are low, similar to what was observed on the speed profiles for this study. However, it should be noted that the findings from Haghani *et al.* (2009) (I-95 Corridor Coalition) showed that the accuracy of the Inrix probe data speeds increased with increasing speed, which is not supported by the findings of this study. It is possible that the accuracy of real-time data improves with an increase in operating speeds, as was the case for the I-95 Corridor Coalition, whilst the quality of historical probe data estimates declines with an increase in speeds, as the findings of this study suggest. Alternatively, it is possible that the quality of probe data is dependent on the service provider.

6.3 ACCURACY MEASURES IN SPEED BINS

The trends observed on the speed profiles suggest that TomTom historical speeds were better estimates at low speeds. To investigate this issue, the 15-minute and 1-hour speeds for the combined freeway segments were placed in speed bins before determining the accuracy measures for the different speed bins. The literature revealed that the speed bins commonly used for freeways were 0-30 mph (0-42.3 km/h), 30-45 mph (42.3-72.4 km/h), 45-60 mph (72.4-96.6 km/h) and 60+ mph (96.6 km/h) (Haghani *et al.*, 2009; Schneider IV *et al.*, 2010).

For South African traffic conditions, these speeds bins were not appropriate. The minimum and maximum speeds at the 15-minute interval were 24.7 km/h and 120.2 km/h, respectively. At the 1-hour interval, the minimum and maximum speeds were 25.0 km/h and 119.1 km/h, respectively. Only one 15-minute interval speed was greater than 120 km/h. In addition, six 15-minute and one 1-hour interval speeds were less than 30 km/h. With this in mind, three speed bins, namely 0-60 km/h, 60-90 km/h and 90+ km/h, were used. It is unlikely that the choice of the speed bins significantly affects the results of the accuracy measures. Table 6.1 and Table 6.2 show the accuracy measures for the 15-minute and 1-hour intervals, respectively. The number of observations in each speed bin and the accuracy measures for the combined speed bins are also shown in the tables.

Table 6.1: Error quantities for the different speed bins (15-minute interval)

Speed Bin (km/h)	No. of intervals	Signed error	AASE	SEB
0-60	26	-1.3	3.8	-1.6
60-90	38	-5.0	4.1	-3.6
90+	296	-6.4	6.9	-6.9
All	**360**	**-5.8**	**6.4**	**-6.2**

Table 6.2: Error quantities for the different speed bins (1-hour interval)

Speed Bin (km/h)	No. of intervals	Signed error	AASE	SEB
0-60	6	-3.2	4.4	-2.5
60-90	10	-6.2	4.4	-4.4
90+	74	-6.4	6.9	-6.9
All	**90**	**-6.2**	**6.5**	**-6.3**

As expected, the *signed error*, *AASE* and *SEB* for all the three speed bins were less than their respective allowable errors. It was interesting to note that the magnitude of the three error quantities increased as the speeds increased for both the 15-minute and 1-hour intervals (Table 6.1 and Table 6.2). This confirms that TomTom speed estimates were better at low speeds and the quality of TomTom speed estimates declined with increase in speeds. A detailed discussion on the possible reasons for the decline in the quality of TomTom speed estimates with increase in speeds was presented in Chapter 5, section 5.9.

6.4 DAILY VARIATION

In Chapter 4, the results for the 15-minute and 1-hour weekday speeds for February 2015 were presented. In other words, all the 15-minute and 1-hour speeds for February 2015 were aggregated to obtain an average speed for that 15-minute or 1-hour time interval. As previously mentioned, the major limitation associated with this method is the fact that the speeds were aggregated and the daily variation in speeds cannot be observed on the speed profiles.

In trying to understand how TomTom speed estimates compare to the reference speeds for a given day, 15-minute and 1-hour speeds for the N1 Western Bypass southbound section between Blouvalk (1010) and Pelican (1012) were computed for a randomly chosen day (Tuesday, 10 February 2015). Figure 6.5 and Figure 6.6 show the speed profiles for the N1 Western Bypass southbound section for the 15-minute and 1-hour intervals, respectively.

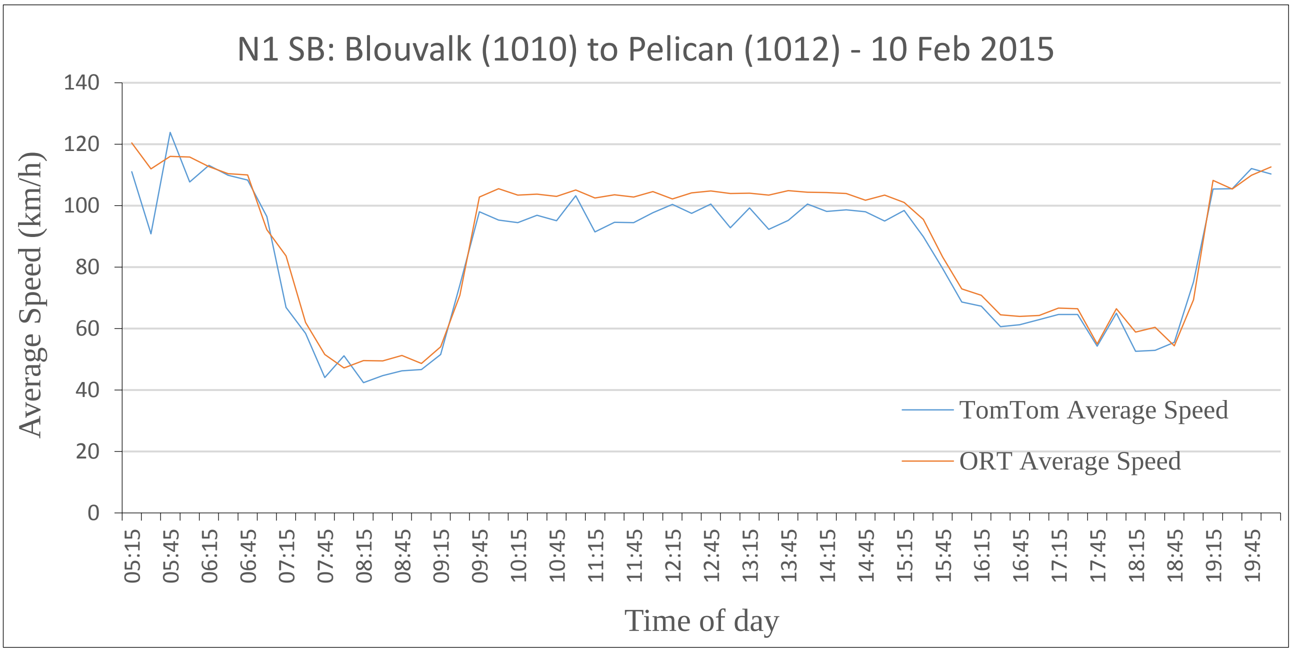


Figure 6.5: Speed profile between Blouvalk (1010) and Pelican (1012) for one day (15-minute interval)

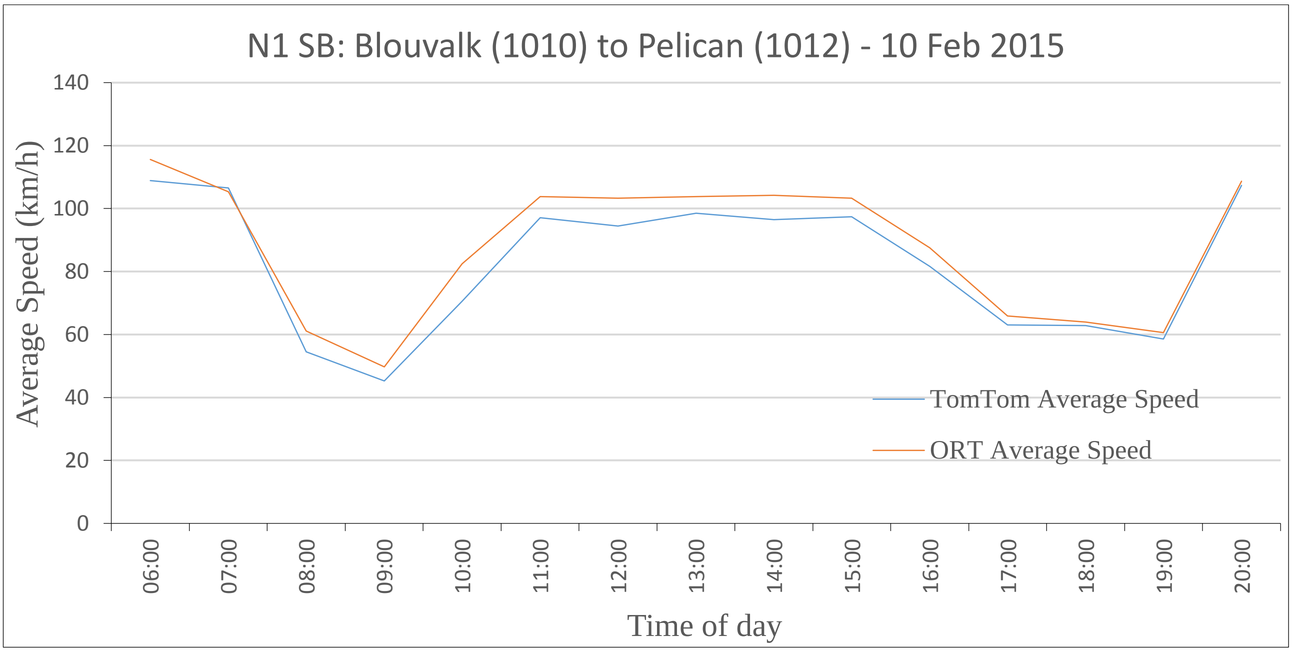


Figure 6.6: Speed profile between Blouvalk (1010) and Pelican (1012) for one day (1-hour interval)

The basic shape of Figure 6.5 is similar to the shape of the speed profile for the monthly speeds (Figure 4.3). A morning peak between 07:30 and 09:30 and evening peak between 16:30 and 18:45 were observed on the 10th of February 2015. During the off-peak periods, speeds just above 100 km/h were observed. It should also be noted that the speed profile for the 1-hour interval (Figure 6.6) for a single day was also similar to the speed profile for February 2015. Therefore, in terms of shape of the speed profiles, speeds, peak and off-peak periods, there was not much difference between the speed profiles for a single day and speeds aggregated over a one-month period.

The variation in the 15-minute speeds for the 10th of February 2015 (Figure 6.5) were more pronounced compared to the variation that was observed on the speed profile for the February 2015 (Figure 4.3). Furthermore, the variation in speeds on the 1-hour interval speed profile for a single day (Figure 6.6) was less pronounced compared to the variation observed on the 15-minute speed profile for a single day (Figure 6.5).

Turning to the accuracy measures for one-day and one-month periods, all the three error quantities were less than their respective allowable error. It was interesting to note that the accuracy measures for the 10th of February were less than the accuracy measures the month of February. This was observed at both the 15-minute and 1-hour intervals. In addition, the results of a t-test are shown in Table 6.3. The t-test was carried out to determine whether there were significant differences between the *signed error*, *AASE* and *SEB* for a single day (10 February 2015) and one month (February 2015). Surprisingly, there were statistically significant differences ($p<0.05$) for all the *error quantities* at 95 % confidence level for the 15-minute interval. However, there were no statistically significant differences at the 1-hour interval level ($p>0.05$).

Table 6.3: t-test results for 15-minute and 1-hour interval errors

Type of error	15-minute interval			1-hour interval		
	1-day	1-month	p-value	1-day	1-month	p-value
Signed error	-5.0	-8.0	<0.01	-6.0	-8.5	0.09
AASE	5.4	6.9	0.02	5.2	7.2	0.12
SEB	-4.4	-6.9	<0.01	-5.1	-7.2	0.11

Considerable variation was observed in the speeds for the N1 Western Bypass southbound section on the 10th of February 2015. The variation in speeds for one day was about three times the variation for the monthly speeds. This actually makes sense because when the speeds were aggregated and filtered, more data points that were further away from the mean speed were discarded for the one-month period compared to the one-day period. As a result, there was still considerable variation in the speeds for the one-day period.

For the daily variations, only one random day (10 February 2015) for one freeway segment (N1_SB_1010_1012) was investigated. For this reason, it was not possible to make conclusions on the quality measures for TomTom speed estimates for a single day. Further research on the issue of daily variations is required.

6.5 TOMTOM SPEED CORRECTION

The findings that were presented earlier in this chapter indicate that TomTom speed estimates were lower than the reference speeds. Although the differences were not constant throughout the analysis period, perhaps it might be the case that a correction is needed to shift the speeds so that the TomTom speed estimates are as close to the reference speeds as possible.

One way to make a correction is to consider using the percentile speeds instead of the average speeds. In Figure 4.15, it was observed that differences between the TomTom and reference speeds were negative, with the majority of speeds having a bias between 5 km/h and 10 km/h. For all the six freeway segments, the 40th, 45th, 50th and 55th percentile speeds were compared to the reference speeds, for both the 15-minute and 1-hour intervals. The percentile speeds that were closest to the reference speeds were noted. Figure 6.7 to Figure 6.12 show the speed profiles for the six freeway segments, describing the TomTom percentile speeds that were closest to the reference speeds.

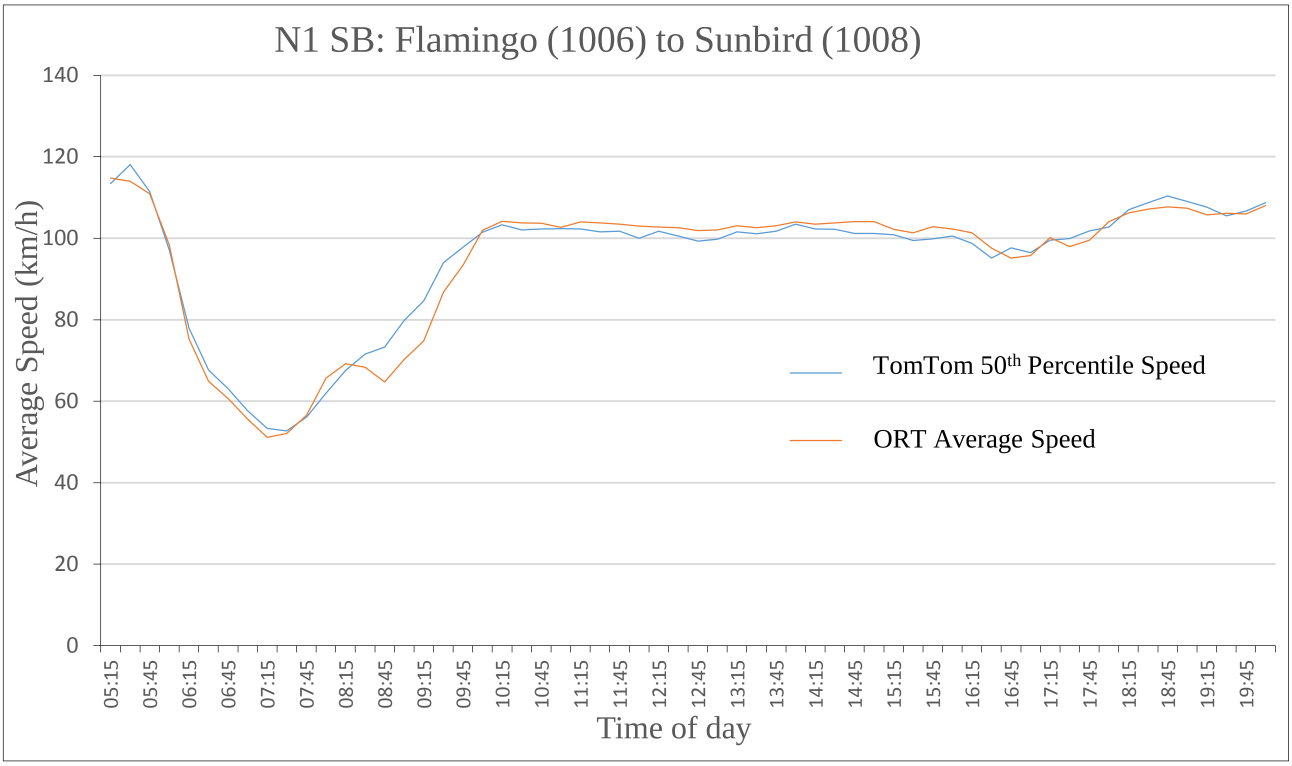


Figure 6.7: Corrected speed profile between Flamingo (1006) and Sunbird (1008)

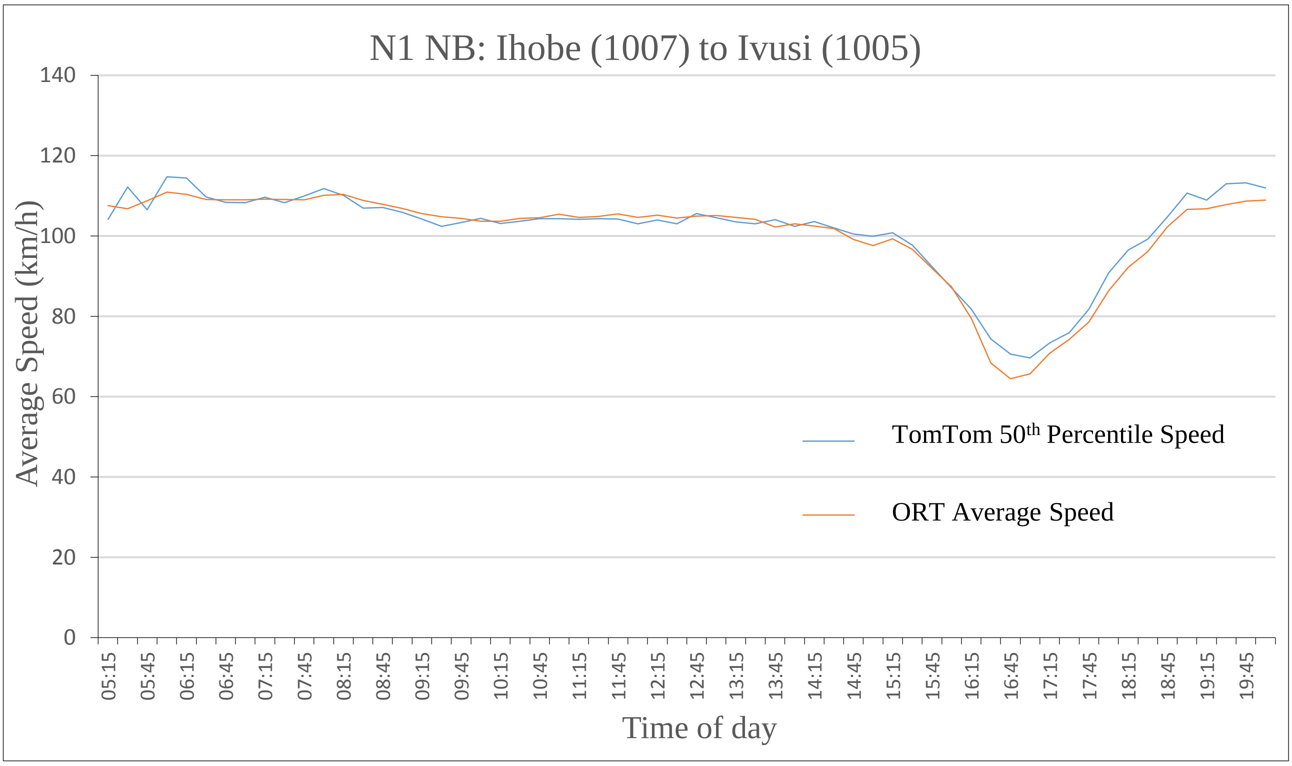


Figure 6.8: Corrected speed profile between Ihobe (1007) and Ivusi (1005)

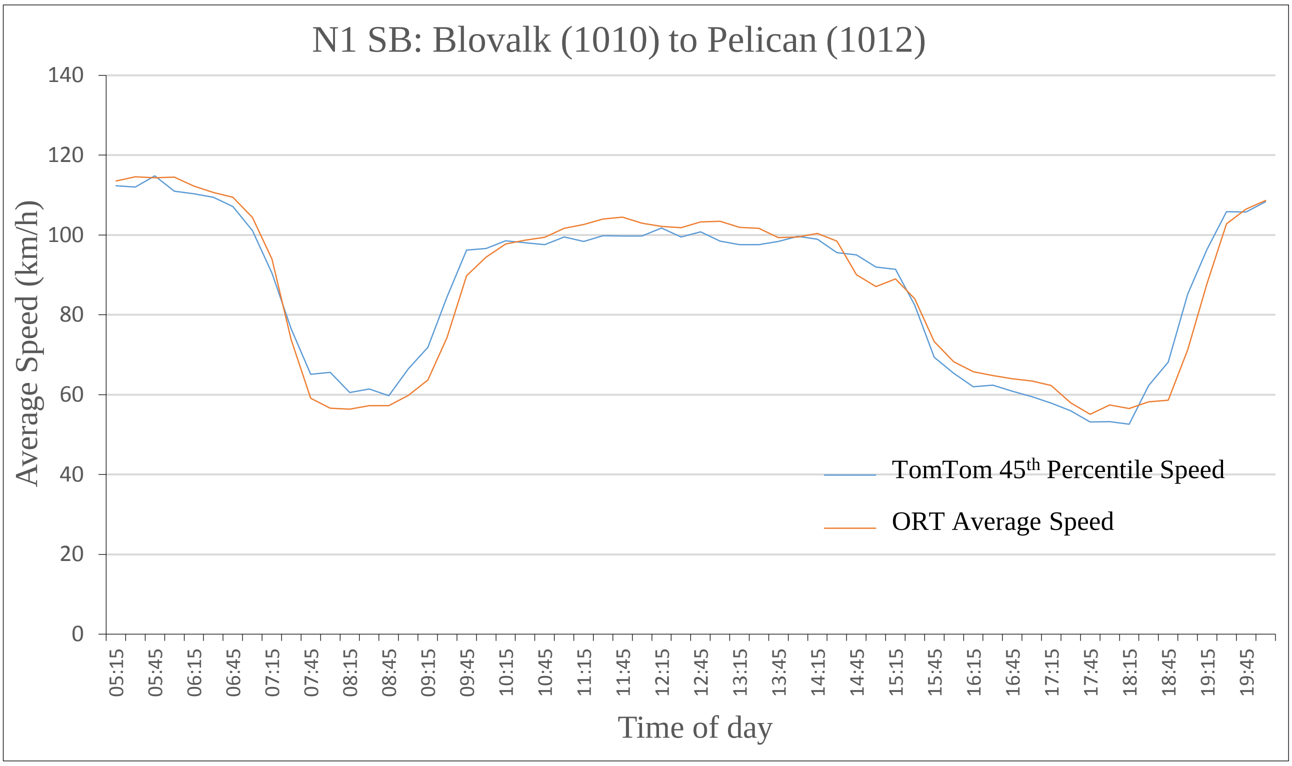


Figure 6.9: Corrected speed profile between Blouvalk (1010) and Pelican (1012)

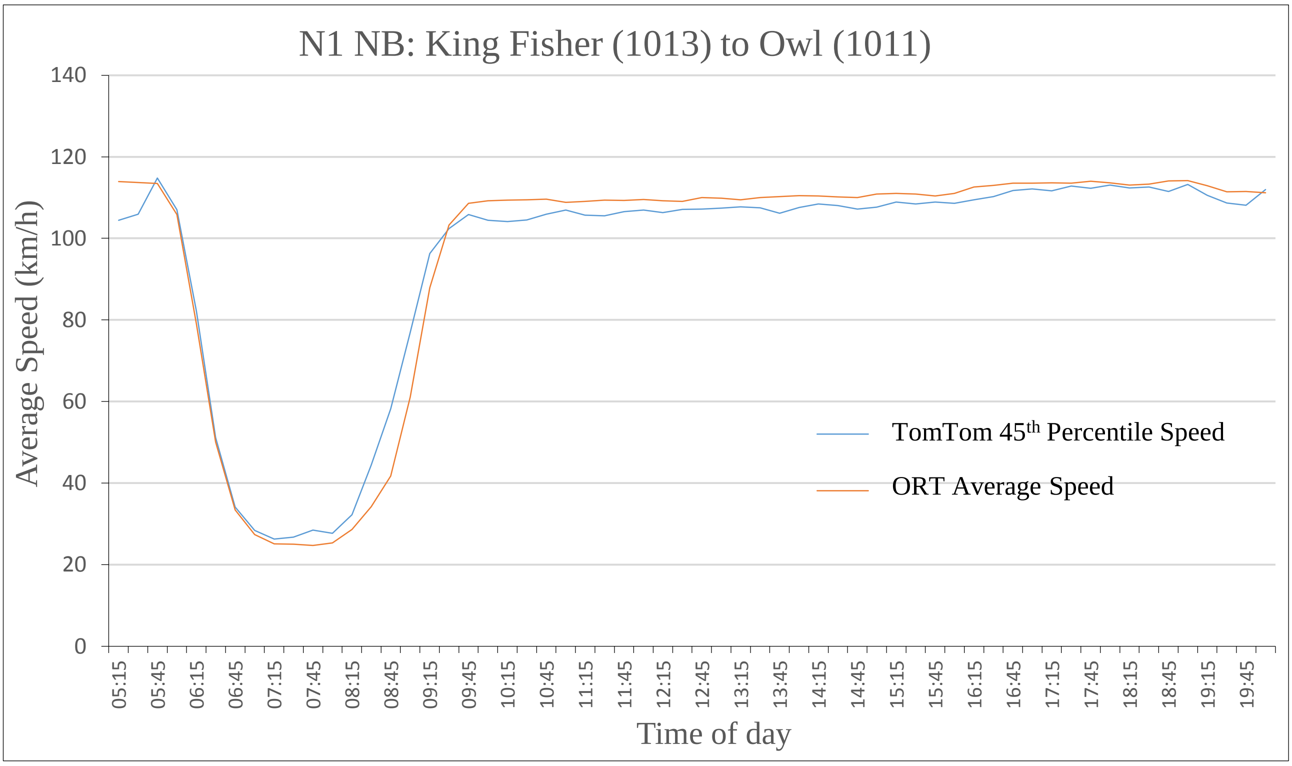


Figure 6.10: Corrected speed profile between King Fisher (1013) and Owl (1011)

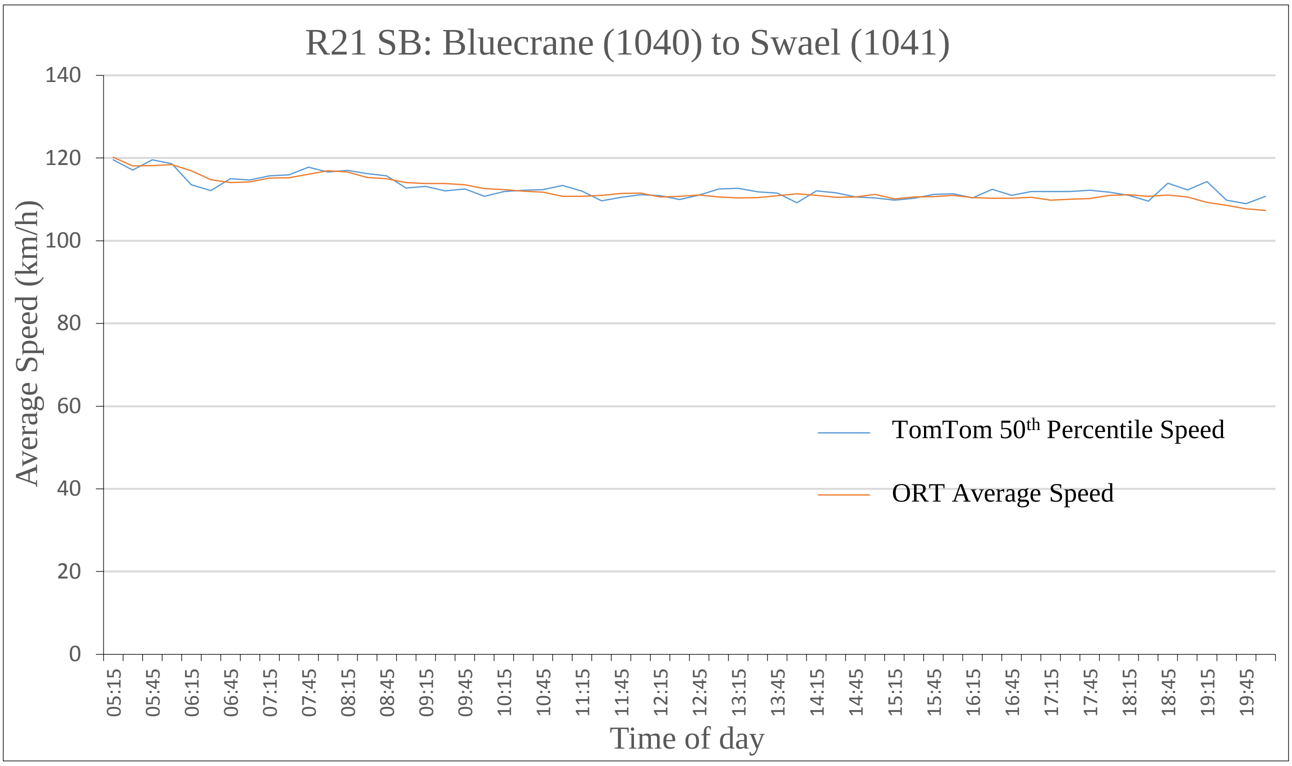


Figure 6.11: Corrected speed profile between Bluecrane (1040) and Swael (1041)

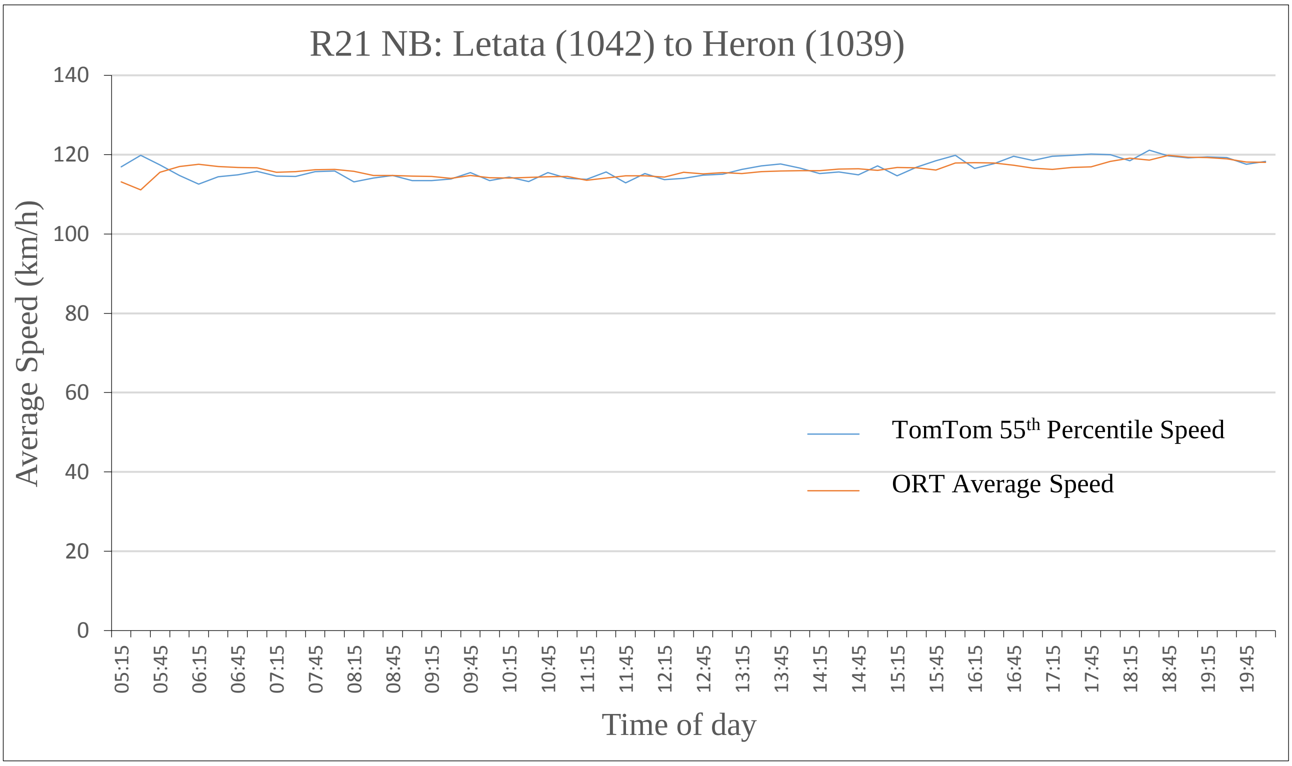


Figure 6.12: Corrected speed profile between Letata (1042) and Heron (1039)

For the N1 Ben Schoeman segments, the 50th percentile speeds were closest to the reference speeds (Figure 6.7 and Figure 6.8). The 45th percentile speeds were a close fit to the reference speeds for the N1 Western Bypass sections (Figure 6.9 and Figure 6.10). However, the R21 sections were different in that the 50th percentile speeds were a close fit to the reference speeds in the southbound direction (Figure 6.11) and the 55th percentile speeds were closest to the reference speeds in the northbound direction (Figure 6.12).

Table 6.4: Error quantities derived from average and percentile speeds (15-minute interval)

Road section	**Average Speed**			**Percentile speeds**			**Percentile**
	S.E	**AASE**	**SEB**	**S.E**	**AASE**	**SEB**	**Speed**
N1_SB_1006_1008	-7.4	7.1	-7.0	0.7	2.3	0.3	50th
N1_NB_1007_1005	-5.1	5.5	-5.3	1.1	1.9	1.0	50th
N1_SB_1010_1012	-8.0	6.9	-6.9	0.7	3.6	0.2	45th
N1_NB_1013_1011	-2.4	4.9	-3.9	1.0	3.2	-0.8	45th
R21_SB_1040_1041	-5.3	5.9	-5.9	0.4	1.2	0.4	50th
R21_NB_1042_1039	-6.9	8.0	-8.0	0.2	1.4	0.2	55th
All	-5.8	6.4	-6.2	0.7	2.3	0.2	-

Note: S.E is the signed error

Table 6.5: Error quantities derived from average and percentile speeds (1-hour interval)

Road section	**Average Speed**			**Percentile speeds**			**Percentile**
	S.E	**AASE**	**SEB**	**S.E**	**AASE**	**SEB**	**Speed**
N1_SB_1006_1008	-7.8	7.3	-7.3	-0.1	1.7	-0.2	50th
N1_NB_1007_1005	-5.3	5.5	-5.5	0.8	1.4	0.7	50th
N1_SB_1010_1012	-8.5	7.2	-7.2	0.3	2.9	0.0	45th
N1_NB_1013_1011	-3.3	4.9	-4.1	0.2	2.8	-1.0	45th
R21_SB_1040_1041	-5.3	6.0	-6.0	0.4	0.8	0.5	50th
R21_NB_1042_1039	-6.9	8.0	-8.0	0.1	0.9	0.1	55th
All	-6.2	6.5	-6.3	0.3	1.7	0.0	-

Note: S.E is the signed error

Table 6.4 and Table 6.5 show the accuracy measures derived from both the average speeds and the appropriate percentile speeds for the 15-minute and 1-hour intervals, respectively. The percentile speeds resulted in the three error quantities that were significantly lower than the errors derived from the average speeds. The signed error in only one of the six segments was greater than 1 % for the 15-minute interval, whereas for the 1-hour interval, signed errors less than 1 % were achieved. The AASE was reduced to less than 4 km/h for all the segments, at the 15-minute and 1-hour intervals. Interestingly, the SEB was reduced to less than 1 km/h for both the 15-minute and 1-hour intervals (Table 6.4 and Table 6.5). In all the cases, the accuracy measures were significantly reduced by using appropriate percentile speeds instead of the average speeds.

6.5.1 DISCUSSION ON THE CORRECTION OF TOMTOM AVERAGE SPEEDS USING THE PERCENTILES SPEEDS

In the previous section, it was shown that a correction was needed to shift the TomTom speed estimates so that the estimates are as close to the reference speeds as possible. One way to make a correction is by using a regression model to predict the reference speeds, using the existing benchmark and TomTom speeds. However, there is a major limitation associated with this approach. The regression model depends on the available data points. It is possible that the regression model will not be suitable for other freeway segments, particularly those that were not investigated in this study.

In trying to maximise the use of the data that is reported by TomTom, the percentile speeds were compared to the reference speeds. It was observed different percentile speeds were remarkably closer to the reference speeds than the average TomTom speeds (Figure 6.7 to Figure 6.12). Depending on the freeway segment, the 45^{th}, 50^{th} and 55^{th} percentile speeds were a close fit to the reference speeds. It was proposed that the reason for the low speed estimates from the service provider was mainly due to a biased sample that is not a true representation of the traffic stream. The motivation for using percentile speeds was to adjust the sample to be as close to the population as possible. By applying percentile speeds higher than the average speeds, it then shifted up the TomTom speed estimates. Better estimates were achieved by using an appropriate percentile speed as a correction.

It should be noted that in this particular study, it was possible to find the suitable percentile speeds that closely fit the reference speeds. In practise, however, there are no reference speeds to check which percentile speeds are appropriate. It appeared as if the appropriate percentile speeds depend on the section of the freeway. However, the R21 sections were approximated by the 50^{th} and 55^{th} percentile speeds in the southbound and northbound direction, respectively. This seems to suggest the freeway section does not determine the suitable percentile speeds.

6.5.2 THE EFFECT OF TRAFFIC COMPOSITION ON THE ACCURACY OF TOMTOM SPEED ESTIMATES

A discussion on why the different freeway segments were estimated by different percentile speeds is presented in this section. In particular, this section attempts to find the relationship between the traffic composition and the appropriate percentile speeds. The average reference speeds for the different freeway segments were presented in Table 4.1. The trend suggests that the percentile speeds that closely fit the reference speeds are related to the average reference speed. The 50th percentile speeds were suitable for the N1 Ben Schoeman sections with average speeds of 94.5 km/h and 100.4 km/h in the southbound and northbound directions, respectively.

The R21 northbound section had average reference speed of 116.1 km/h was closely matched by the 55th percentile speeds. The two anomalies were the N1 Western Bypass northbound section (average speed of 96.0 km/h) which was a close fit to the 45th percentile speeds and R21 Albertina Sisulu southbound section (average speed of 112.2 km/h) which closely matched the 50th percentile speeds. In light of the above discussion, the relationship between the average speed and the appropriate percentile speeds is unclear.

A more systematic approach would be to look at the traffic composition on the different freeway segments. The reason for the low TomTom speed estimates was due to possibly high composition of heavy vehicles in the sample used to derive the speed estimates. The aim here is to investigate the relationship between the traffic composition and the appropriate percentile speed to be used as a correction. Heavy vehicles are of particular interest in such an investigation. However, it should be noted that the issue of heavy vehicles and traffic composition in general, is more complex than what has been discussed above. Traffic composition on a freeway is not constant throughout the day. Because this issue is outside the scope of this project and the numerous considerations that have to be taken into account, further research is needed to address the issue of traffic composition and appropriate percentile speeds for correcting the TomTom speed estimates.

6.6 CONCLUSION

In this chapter, further analysis of the quality of TomTom historical data and possible correction measures were presented. It was shown that the quality of TomTom speed estimates declined with increase in speed. Although, the TomTom speed estimates were within the acceptable accuracy thresholds, it was found that by using percentile speeds instead of the average speeds, the accuracy of TomTom speed estimates was significantly improved. Thus, the accuracy of the speed estimates was enhanced by using an appropriate percentile speed as a correction.

CHAPTER 7 : CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

7.1 INTRODUCTION

The importance of traffic data in transportation engineering, planning and operations was discussed in Chapter 1. Although the conventional methods of data collection continue to be the main source of traffic data, the advancement in technology and vehicle tracking methods is impetus behind the emerging of alternative and innovative sources of traffic data, such as ITS data sources.

Vast amounts of data are obtained from ITS data sources. However, the data from the ITS sources raise new concerns about the data quality. This study set out to investigate the quality of TomTom historical average speeds on freeway segments in Gauteng, South Africa. This chapter highlights how the research question was answered, making reference to the research goals and objectives stated in Chapter 1.

7.2 SUMMARY OF FINDINGS

The first objective stated that:

Research objective i: To develop a suitable data processing plan to process and manipulate the reference data.

Raw ORT data was used as the reference data in this study. Before any meaningful information was obtained from ORT data, the data was processed and reduced by means of a computer program. The program used the matching, interval allocation, filtering and aggregation techniques discussed in the methodology chapter. Therefore, from the discussions in Chapter 3 (Research design and methodology), the first research objective was met.

Research Objective ii: To investigate the quality measures of TomTom historical average speeds.

The quality measures that were investigated were accuracy, validity, completeness, coverage and accessibility. In terms of the accuracy of the combined freeway segments for February 2015, the *signed error, AASE* and *SEB* were less than their respective allowable errors, for speeds derived from both the 15-minute and 1-hour intervals. Considering the individual freeway segments, TomTom speed estimates generally satisfied the specified accuracy thresholds. In addition, the accuracy measures for a single day were within the thresholds of the allowable errors. There were no statistically significant differences between the error quantities for speeds derived from 15-minute and 1-hour intervals.

The service data provider generally underestimated traffic speeds. This resulted in the rejection of the null hypothesis (at a 95 % confidence level) that the TomTom mean speeds were equal to the ORT mean speeds, at both the 15-minute and 1-hour intervals.

Considering the *signed error* and *AASE*, TomTom historical data was of very high quality. However, the same data was of moderate quality based on the *SEB*. This demonstrates how validity was dependent on the selected measures and set thresholds. In as far as the other quality measures are concerned, it was established that they help in reinforcing and explaining the results of accuracy and validity. In addition, TomTom historical data was easy to use and understand as it was already processed. In light of the aforementioned remarks and findings, it is evident that the second research objective was achieved.

Research Objective iii: To identify ways of explaining and reducing the differences between TomTom speed estimates and reference speeds.

Although the accuracy measures of TomTom speed estimates were within the specified accuracy thresholds, it was observed that TomTom generally underestimates traffic speeds on freeway segments. However, the differences between TomTom speeds estimates and reference speeds were not constant throughout the analysis period. TomTom speeds were closer to the reference speeds during the peak periods compared to off-peak periods. In addition, the quality of TomTom speed estimates declined with increase in speeds. In other words, TomTom speeds estimates were better when the freeway segments were operating at low speeds.

A biased sample that was not a true representation of the traffic population was attributed as a possible reason for the low TomTom speed estimates. Heavy vehicles generally travel at lower speeds compared to passenger cars. The issue of unfamiliar drivers using GPS devices was identified as a small contributing factor. Drivers unfamiliar with the freeway facility tend to drive slower than those familiar with the area. It is likely that unfamiliar drivers also use GPS devices for navigation purposes. A sample consisting mainly of heavy vehicles and drivers unfamiliar with the freeway facilities who also use GPS devices is likely to result in speed estimates lower than the reference speeds.

To address the data quality deficiencies and improve on the quality of TomTom historical data, different percentile speeds were compared to the reference speeds. The motivation being that perhaps percentile speeds higher than the average speeds will result in speeds that are closer to the reference speeds. It was found that using certain percentile speeds instead of average speeds significantly reduces the magnitude of the error quantities. It appears as though the appropriate percentile speeds were related to the average speed and traffic composition on the freeway. Therefore, the accuracy of the speed estimates can be improved by using an appropriate percentile speed as a correction.

7.3 GENERAL CONCLUSIONS

The research question was stated as:

Research question: Is the quality of TomTom historical average speeds on freeway segments consistent with reference speeds?

The discussion in section 7.2 demonstrates how the research question was answered. It was shown that although TomTom generally underestimates traffic speeds, the quality of TomTom historical average speeds was in line with the reference speeds.

In this study, five quality measures were investigated and it was evident that some data consumers value certain data quality measures more than others do. It might be useful and perhaps desirable to consider composite data quality measures as long as there is understanding on the meaning and implications of the composite data quality measures.

Currently, probe data is used to supplement traffic data from traditional data sources. There is potential for probe data to be a key source of traffic data in mainstream applications of transportation engineering and operations. The need for alternative data sources presents an opportunity to unlock the wider use probe data. In addition, the use of probe data promotes sustainable and cost-effective transportation systems and contributes in addressing the critical issues facing developing countries such as South Africa.

7.4 RECOMMENDATIONS FOR FUTURE RESEARCH

This section discusses the recommendations for future research. It was established that the TomTom speed estimates were generally lower than reference speeds. The reason for the low speed estimates was possibly due to a biased sample, which was not a true representation of the traffic stream. Based on the work presented in this study, the following areas are recommended for future studies:

i. Investigate and develop a model for predicting the appropriate percentile speeds to be used for correcting TomTom speed estimates. The model should take into consideration the type of facility, average speed and the distribution of the traffic stream, in terms of heavy vehicles, passenger cars and other type of vehicles.
ii. Assess the quality measures of probe data from other service providers.

LIST OF REFERENCES

Aliari, Y. & Haghani, A. 2012. Bluetooth sensor data and ground truth testing of reported travel times. *Transportation Research Record: Journal of the Transportation Research Board*. 2308(18):167–172.

Ayala, D., Lin, J., Wolfson, O., Rishe, N. & Tanizaki, M. 2010. Communication reduction for floating car data-based traffic information systems. In *2nd International Conference on Advanced Geographic Information Systems, Applications, and Services, GEOProcessing 2010*. 44–51.

Bailey, R.A. 2008. *Design of comparative experiments*. Cambridge University Press.

Battelle, Cambridge Systematics Inc & Texas Transportation Institute. 2004. *Traffic data quality measurement*. Ohio. [Online], Available: http://ntl.bts.gov/lib/jpodocs/repts_te/14058.htm [2015, September 20].

Breitenberger, S., Gruber, B., Neuherz, M. & Kates, R. 2004. Traffic information potential and necessary penetration rates. *Traffic Engineering & Control*. 45(11):396–401.

Cetin, M., List, G.F. & Zhou, Y. 2005. Factors affecting the minimum number of probes required for reliable estimation of travel time. *Transportation Research Record: Journal of the Transportation Research Board*. 1917(518):1–18.

Chen, M. & Chien, S. 2000. Determining the number of probe vehicles for freeway travel time estimation by microscopic simulation. *Transportation Research Record: Journal of the Transportation Research Board*. 1719:61–68.

Chen, M. & Chien, S. 2001. Dynamic freeway travel-time prediction with probe vehicle data: Link based versus path based. *Transportation Research Record: Journal of the Transportation Research Board*. 1768:157–161.

Cykana, P., Paul, A. & Stern, M. n.d. *DOD Guidelines on data quality management*. Virginia: Defence Information Systems Agency. [Online], Available: http://mitiq.mit.edu/ICIQ/Documents/IQ Conference 1996/Papers/DODGuidelinesonDataQualityManagement.pdf [2016, June 20].

Dannehy, M. & Krootjes, P. 2013. *How to use TomTom's Traffic Stats portal to access historic traffic data*. [Online], Available: https://vimeo.com/63090411 [2015, August 07].

English, L.P. 1999. *Seven deadly misconceptions about information quality*. Brentwood, Tennessee: Information Impact International. [Online], Available: https://www.fhwa.dot.gov/policyinformation/hpms/dataquality.cfm [2016, February 09].

Ferman, M.A. & Blumenfeld, D.E. 2006. *Penetration requirements for real-time traffic information from probe vehicles*. Michigan.

Ferman, M.A., Blumenfeld, D.E. & Dai, X. 2005. An analytical evaluation of a real-time traffic information system using probe vehicles. *Journal of Intelligent Transportation Systems*. 9(1):23–34.

Fontaine, M.D., Yakkala, A.P. & Smith, B.L. 2007. *Probe sampling strategies for traffic monitoring systems based on wireless location technology*. Virginia. [Online], Available: http://www.virginiadot.org/vtrc/main/online_reports/pdf/07-cr12.pdf [2016, March 14].

Friesen, M.R. & McLeod, R.D. 2015. Bluetooth in Intelligent Transportation Systems: A survey. *International Journal of Intelligent Transportation Systems Research*. 13(3):143–153.

Garber, N.J. & Hoel, L.A. 2010. *Traffic and highway engineering*. 4th ed. Cengage Learning.

Green, M., Fontaine, M. & Smith, B. 2004. Investigation of dynamic probe sample requirements for traffic condition monitoring. *Transportation Research Record: Journal of the Transportation Research Board*. 1870:55–61.

Haas, R., Carter, M., Perry, E., Trombly, J., Bedsole, E. & Margiotta, R. 2009. *iFlorida model deployment: Final evaluation report*. Viriginia. [Online], Available: http://ops.fhwa.dot.gov/publications/fhwahop08050/iFlorida.pdf [2016, March 14].

Haghani, A., Hamedi, M. & Sababadi, K.F. 2009. *I-95 Corridor Coalition Vehicle Probe Project : Validation of Inrix data July-September 2008 Final Report*. [Online], Available: http://www.i95coalition.org/wp-content/uploads/2015/02/I-95-CC-Final-Report-Jan-28-2009.pdf [2015, May 27].

Haghani, A., Hamedi, M., Sadabadi, K., Young, S. & Tarnoff, P. 2010. Data collection of freeway travel time ground truth with Bluetooth sensors. *Transportation Research Record: Journal of the Transportation Research Board*. 2160(2160):60–68.

Haghani, A., Hamedi, M. & Parvan, K. 2013. *I-95 Corridor Coalition Vehicle Probe Project : Validation of Inrix data*. [Online], Available: http://i95coalition.org/wp-content/uploads/2015/02/I-95-CC-Valid-Report-Aug-2013-data-May-2013-GA.pdf?dd650d [2015, May 27].

Hallmark, S.L. 2004. *Role of passenger vehicle and heavy truck speed differences in emissions modeling: Tech transfer summary*. Ames. [Online], Available: http://www.intrans.iastate.edu/publications/_documents/t2summaries/emission_model.pdf [2016, August 21].

Hellinga, B. & Fu, L. 1999. Assessing expected accuracy of probe vehicle travel time reports. *Journal of Transportation Engineering*. (November/December):524–530.

Hellinga, B.R. & Fu, L. 2002. Reducing bias in probe-based arterial link travel time estimates. *Transportation Research Part C: Emerging Technologies*. 10(4):257–273.

Hoogenboom, M. 1999. The grand draw. *Traffic Technology International*. (October/November):66–68.

Inrix. 2006. *Frost and Sullivan study recognises Inrix as the leading provider of real-time information*. [Online], Available: http://inrix.com/press/2583/ [2016, June 21].

Inrix. 2015. *I-95 Corridor Coalition Vehicle Probe Project: Data validation summary*. [Online], Available: http://inrix.com/wp-content/uploads/2015/09/INRIX-I-95-VPP-Data-Summary-Validation-2015.pdf [2016, September 07].

Kandarpa, R., Sangillo, J., Burgess, L. & Toppen, A. 2010. *Real-time traveller information market assessment white paper*. Washington DC. [Online], Available: http://ntl.bts.gov/lib/32000/32900/32927/rtti_wp.pdf [2016, April 09].

Kennedy, J., Cantrell, C.R., Varney, M.D., Czyzewski, Z. & Smith, B.D. V. 2004. Highway travel time analysis using license plate image capture techniques. *Proc. SPIE Industrial and Highway Technology*. 5272:294–303.

Kranthi, S., Pranathi, K. & Srisaila, A. 2011. Automatic number plate recognition. *International Journal of Advancements in Technology*. 2(3):408–422.

Li, Y. 2008. Short-term prediction of motorway travel time using ANPR and loop data. *Journal of Forecasting*. 27(6):507–517.

Liu, H.X. & Ma, W. 2009. A virtual vehicle probe model for time-dependent travel time estimation on signalized arterials. *Transportation Research Part C: Emerging Technologies*. 17(1):11–26.

Long Cheu, R., Xie, C. & Lee, D.-H. 2002. Probe vehicle population and sample size for arterial speed estimation. *Computer-Aided Civil and Infrastructure Engineering*. 17(1):53–60.

Luke, R. & Heyns, G. 2013a. Public transport policy and performance: The results of a South African public opinion poll. *Journal of Transport and Supply Chain Management*. 7(1):1–8.

Luke, R. & Heyns, G. 2013b. *State of transport opinion poll: South Africa*. Johannesburg. [Online], Available: https://www.uj.ac.za/newandevents/Pages/UJ-Transport-Opinion-Poll-20152016-highlights-public-transport-concerns-.aspx [2015, October 24].

Luke, R. & Heyns, G. 2016. South African public opinion on the state of urban transport : An appraisal of the achievement of policy objectives. In WIT Press *22nd International Conference on Urban Transport and the Environment (UT 2016)*. 1–12.

Mahmassani, H.S. 2014. Connected systems, smart cities and urban mobility: Realising the potential of user-centred intelligence. In Instabul *1st National Congress and Exhibition on Intelligent Transport Systems for Highways*. 1–63.

Majangaza, S. 2015. *Move to reduce e-toll fees, improve collection*. [Online], Available: http://www.dispatchlive.co.za/news/2015/05/22/move-to-reduce-e-toll-fees-improve-collection/ [2016, July 29].

Midwestern Consulting. 2008. *SEMCOG Travel time data collection*. Michigan.

Miles, J.C. & Chen, K. 2004. *The Intelligent Transport Systems handbook: Recommendations from the World Road Association (PIARC)*. 2nd ed. AIPCR - PIARC.

Montgomery, D.C. & Runger, G.C. 2007. *Applied Statistics and Probability for Engineers*. 4th ed. Wiley.

Mouton, J. 2012. *How to succeed in your Master's and Doctoral Studies: A South African guide and resource book*. 16th ed. Pretoria: Van Schaik Publishers.

National Cooperative Highways Research Program. 2009. *Travel time data quality assessment: State of the practice*. [Online], Available: http://www.pooledfund.org/Document/Download/1990 [2015, April 24].

National Research Council. 2010. *Highway Capacity Manual 2010*. 5th ed. Washington DC: Transportation Research Board.

North Amercian Traffic Working Group. 2010. *Traffic information benchmarking guidelines*. [Online], Available: http://natwg.itsa.wikispaces.net/file/view/NATWG Traffic Information Benchmarking Guildlines TIBG version 1.0 RELEASED 4-22-10.pdf/136649945/NATWG Traffic Information Benchmarking Guildlines TIBG version 1.0 RELEASED 4-22-10.pdf [2016, July 29].

Quiroga, C. & Bullock, D. 1998. Determination of sample sizes for travel time studies. *ITE Journal on the Web*. 68(8):92–98.

Ressler, S., Thomas, C. & Dannehy, M. 2013. *Sustainable communities through intelligent information: Harnessing the power of road data*. [Online], Available: http://www.slideshare.net/govloop/sustainable-communities-through-intelligent-information-harnessing-the-power-of-road-data [2015, October 03].

Robinson, A. 2014. The use of Automatic Number Plate Recognition (ANPR) data for transport model development. In Adelaide: Australian Institute of Traffic Planning and Management *Australian Institute of Traffic Planning and Management (AITPM) National Conference*. 1–15.

Robinson, A. 2016. *Using new "big" data to improve strategic traffic models*. Stellenbosch.

SANRAL. 2016. *Electronic toll collection in South Africa*. [Online], Available: http://www.sanral.co.za/e-toll/index.php?gort_xml+template~~menu/main~~L1_2065~~ENG~~~~~~ [2016, August 26].

Sanwal, K.K. & Walrand, J. 1995. *Vehicle as probes*. (California PATH Working Paper). [Online], Available: http://escholarship.org/uc/item/3gh0890x [2016, May 17].

Schneider IV, W.H., Turner, S., Roth, J. & Wikander, J. 2010. *Statistical validation of speeds and travel times provided by a data service vendor*. Akron. [Online], Available: http://ntl.bts.gov/lib/55000/55800/55855/FHWA-OH-2010-2.PDF [2016, March 04].

Shafer, J. 2002. *Measures of data quality and reability*. Department of Defence. [Online], Available: http://slideplayer.com/slide/11250472/ [2015, November 22].

Srinivasan, K. & Jovanis, P. 1996. Determination of number of probe vehicles required for reliable travel time measurement in urban network. *Transportation Research Record*. 1537:15–22.

Strong, D.M., Lee, Y.W. & Wang, R.Y. 1997. 10 potholes in the road to information quality. *Computer*. 30(8):38–46.

Texas Transportation Institute. 2012. *Evaluation of TomTom historical average speed data*. Texas. [Online], Available: http://www.tomtom.com/lib/doc/licensing/TTI-Eval-Summary-of-TomTom.pdf [2015, May 23].

TomTom. 2016. *TomTom Traffic Stats*. [Online], Available: https://trafficstats.tomtom.com/resources/session/new [2016, June 05].

Turner, S. 2002. *Defining and measuring traffic data quality*. [Online], Available: http://ntl.bts.gov/lib/jpodocs/repts_te/13767.html [2016, May 09].

Turner, S. & Holdener, D.J. 1995. Probe vehicle sample sizes for real-time information: The Houston experience. In Seattle, Washingtom *Proceeding of the Sixth International Conference on Vehicle Navigation and Information Systems*. 3–9.

Turner, S., Richardson, J., Fontaine, M. & Smith, B. 2011. *Guidelines for evaluating the accuracy of travel time and speed data*. [Online], Available:

https://www.michigan.gov/documents/mdot/mdot_its_accuracy_eval_guidelines_368478_7.pdf [2015, June 25].

Turner, S.M., Eisele, W.L., Benz, R.J. & Douglas, J. 1998. *Travel time data collection handbook*. [Online], Available: https://www.fhwa.dot.gov/ohim/tvtw/natmec/00020.pdf [2015, June 25].

Wright, J. & Dahlgren, J. 2001. *Using probe vehicles equipped with toll tag as probes for providing travel times*. (California PATH Working Paper). Berkeley. [Online], Available: http://www.its.berkeley.edu/sites/default/files/publications/UCB/2001/PWP/UCB-ITS-PWP-2001-13.pdf [2016, January 21].

Wunnava, S. V, Yen, K., Babij, T., Zavaleta, R., Romero, R. & Archilla, C. 2007. *Travel time estimation using cell phones for highways and roadways*. Miami. [Online], Available: http://www.dot.state.fl.us/research-center/Completed_Proj/Summary_TE/FDOT_BD015_12_rpt.pdf [2016, May 18].

Ygnace, J.L., Drane, C., Yim, Y.B. & de Lacvivier, R. 2000. *Travel time estimation on the San Francisco Bay area network using cellular phones as probes*. [Online], Available: http://escholarship.org/uc/item/8xn8m01v [2016, June 02].

Young, S. 2008. *Bluetooth traffic monitoring technology: Concept of operations and deployment guidelines*. [Online], Available: http://www.catt.umd.edu/sites/default/files/documents/UMD-BT-Brochure_REV3.pdf [2016, February 09].

Young, S. & Haghani, A. 2010. Bluetooth traffic monitoring: Technical atrributes and applications. In Maryland *North American Travel Monitoring Exposition and Conference (NAMTEC)*. 1–35.

Zhang, X., Hamedi, M. & Haghani, A. 2015. Arterial travel time validation and augmentation with two independent data sources. *Transportation Research Record: Journal of the Transportation Research Board*. 2526:79–89.

APPENDIX A : 1-HOUR INTERVAL SPEED PROFILES

Speed profiles for February 2015 (1-hour interval)

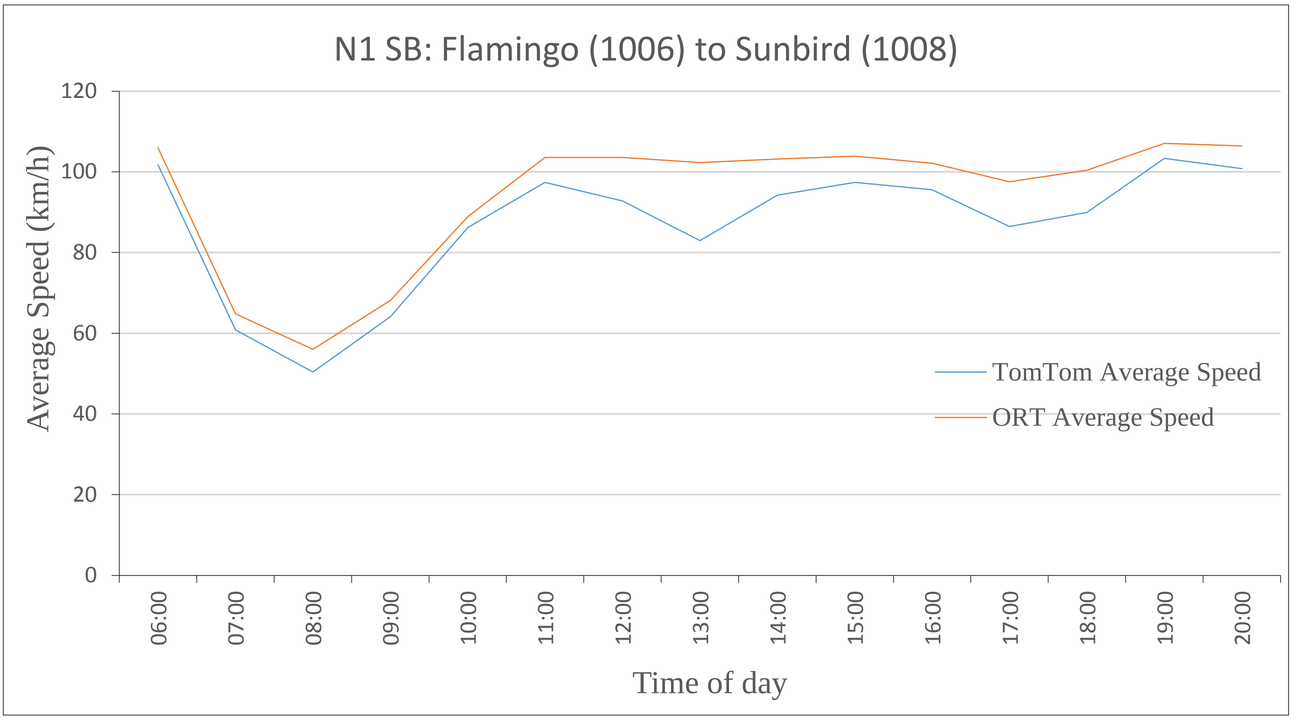


Figure A.1: Hourly speed profile between Flamingo (1006) and Sunbird (1008)

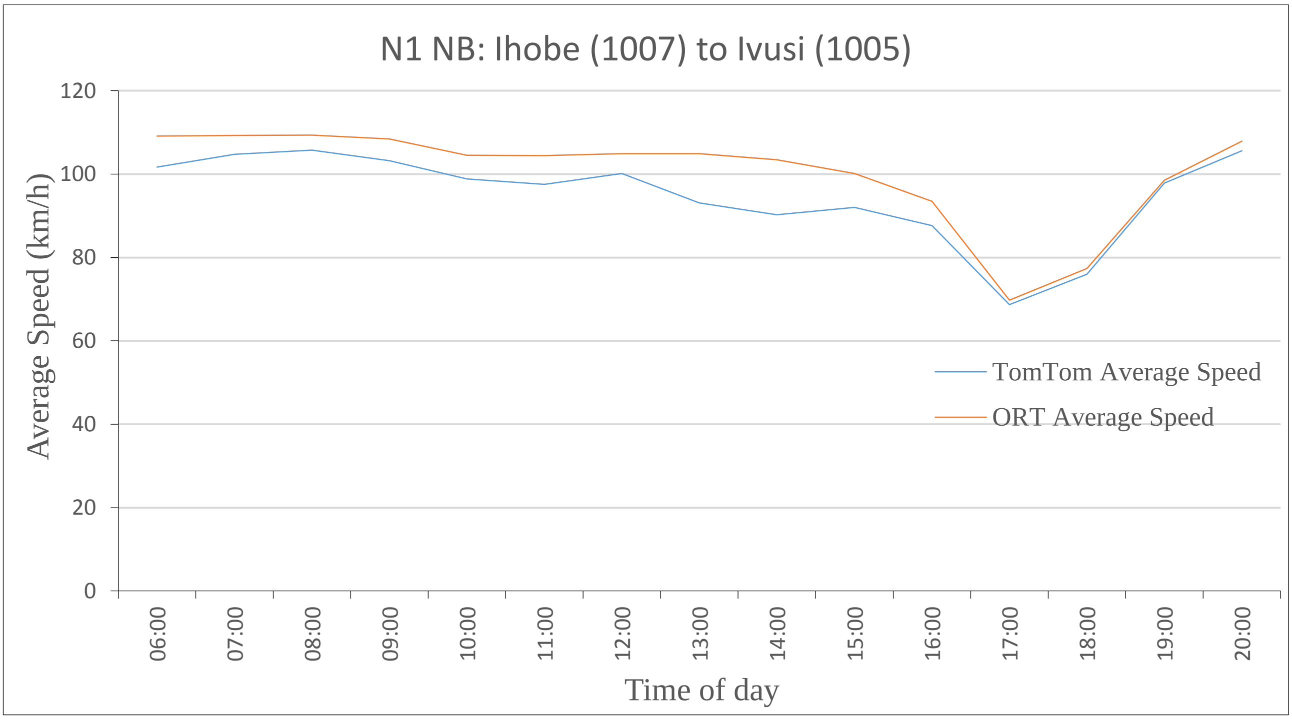


Figure A.2: Hourly speed profile between Ihobe (1007) and Ivusi (1005)

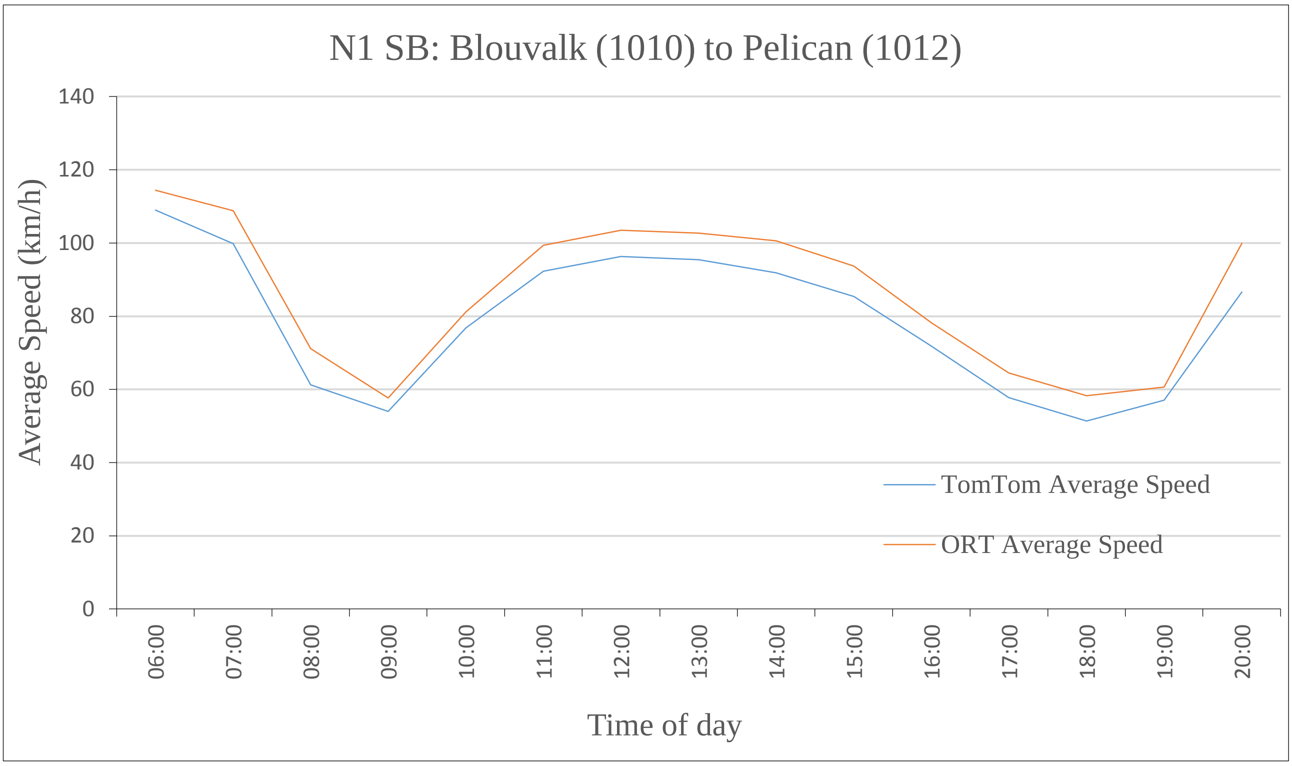


Figure A.3: Hourly speed profile between Blouvalk (1010) and Pelican (1012)

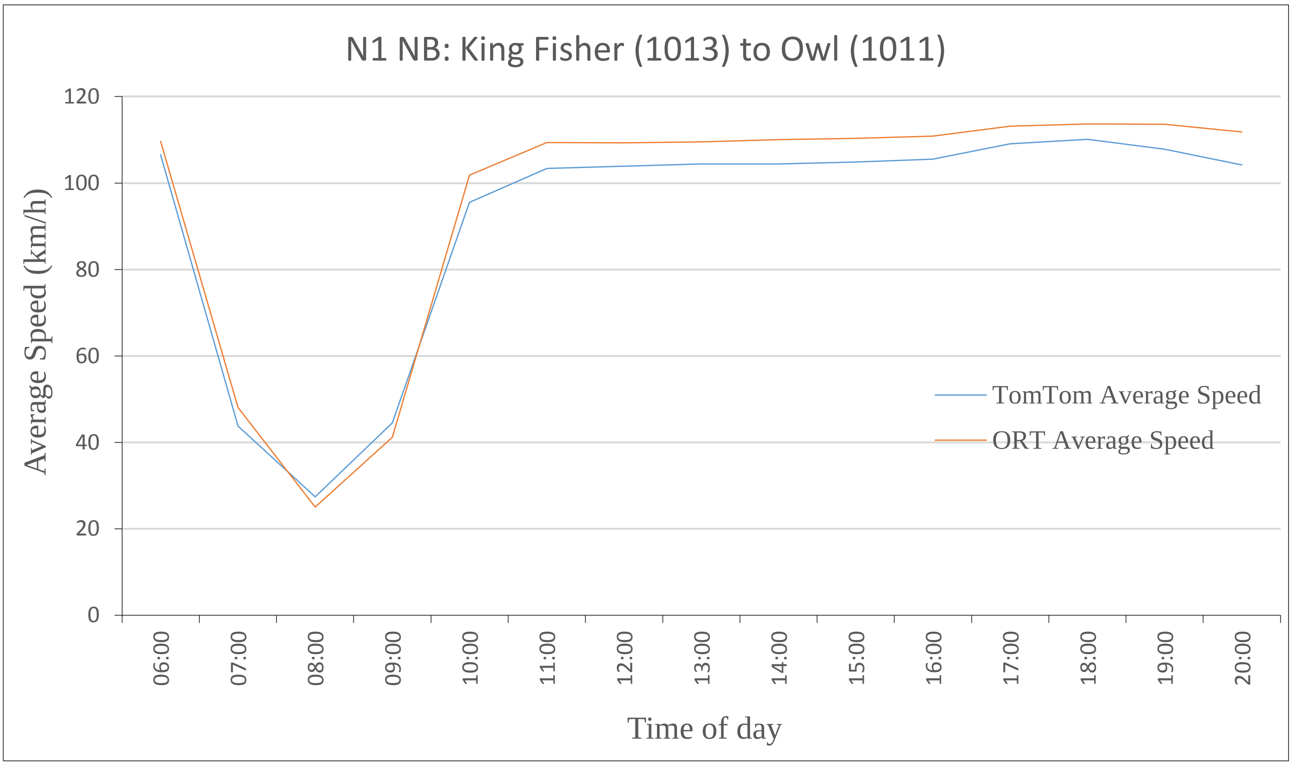


Figure A.4: Hourly speed profile between King Fisher (1013) and Owl (1011)

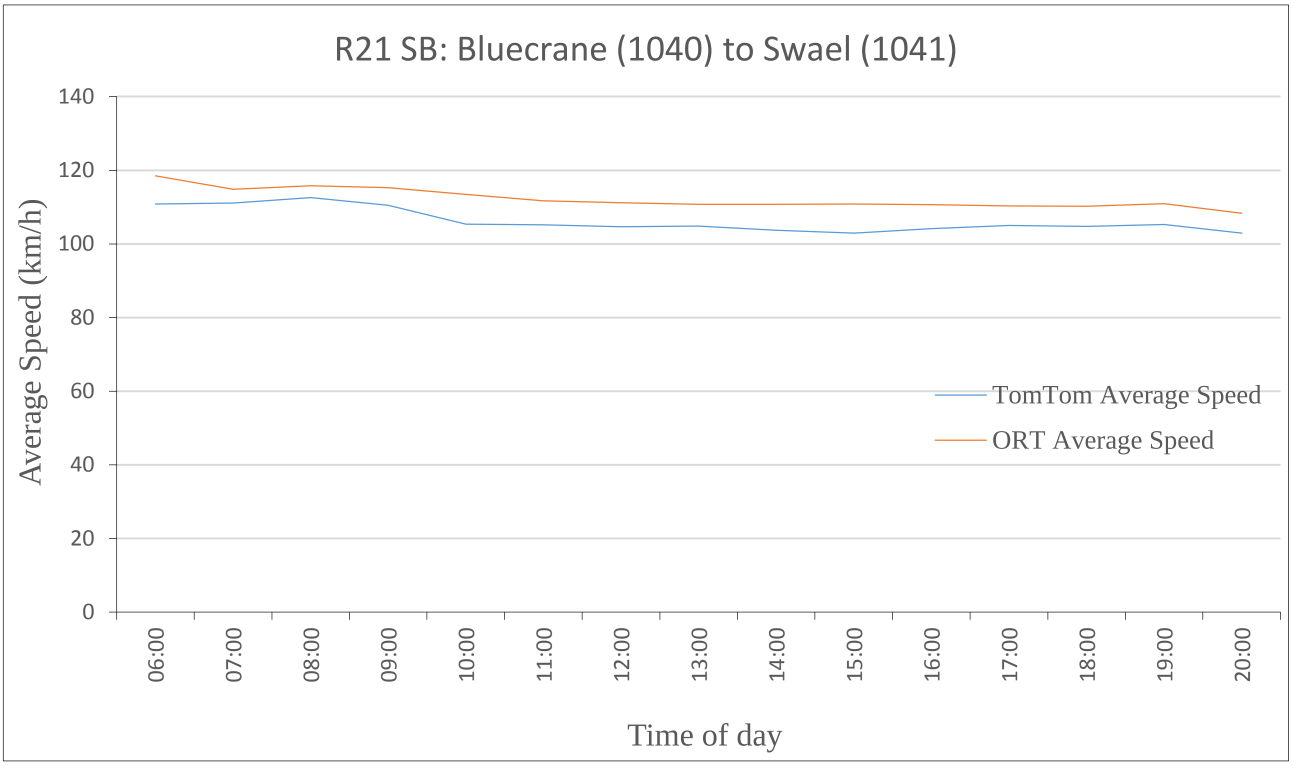


Figure A.5: Hourly speed profile between Bluecrane (1040) and Swael (1041)

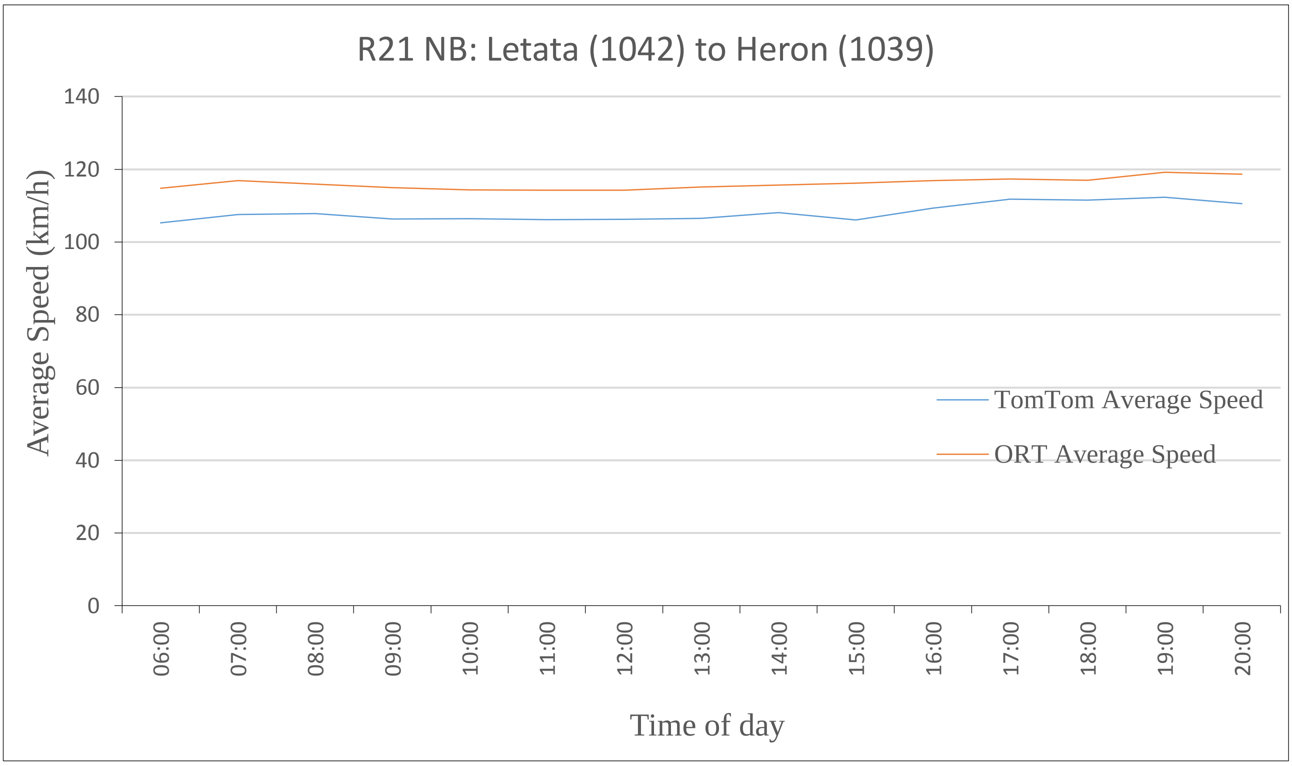


Figure A.6: Hourly speed profile between Letata (1042) and Heron (1039)

APPENDIX B : PERCENTILE SPEED PROFILES

Corrected speed profiles for February 2015 (1-hour interval)

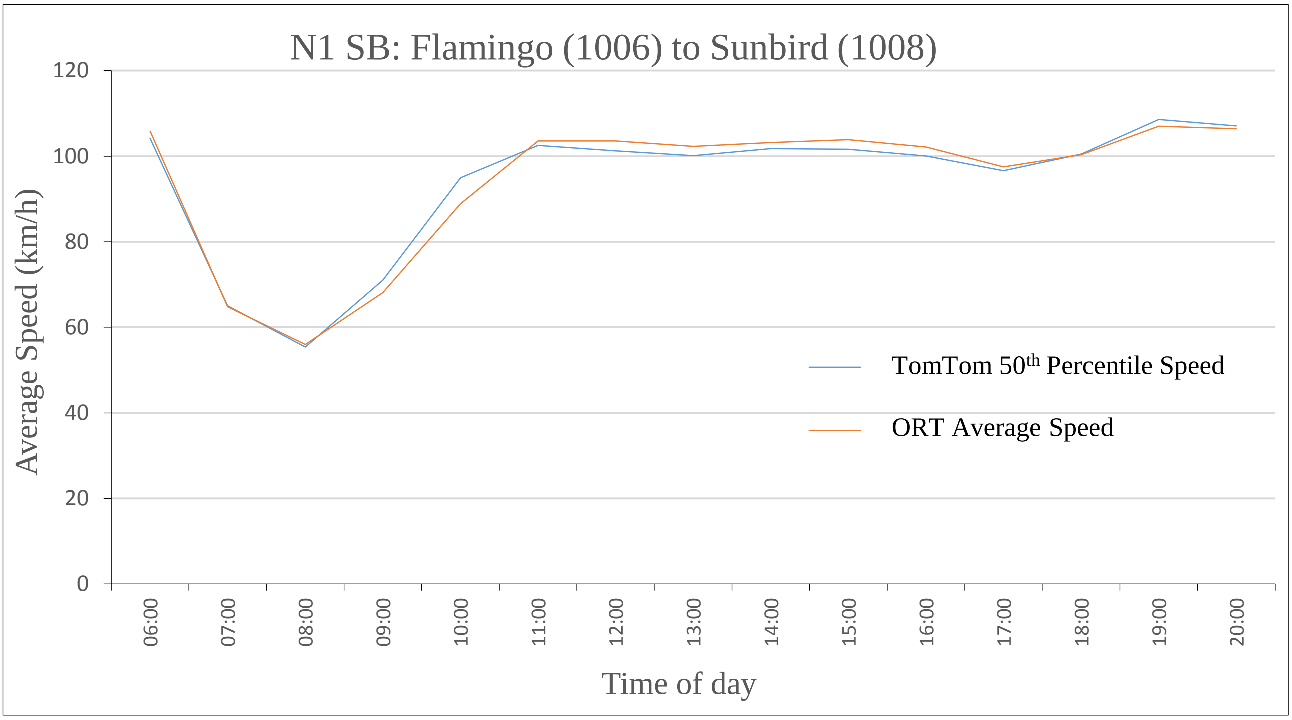


Figure B.1**:** Corrected hourly speed profile between Flamingo (1006) and Sunbird (1008)

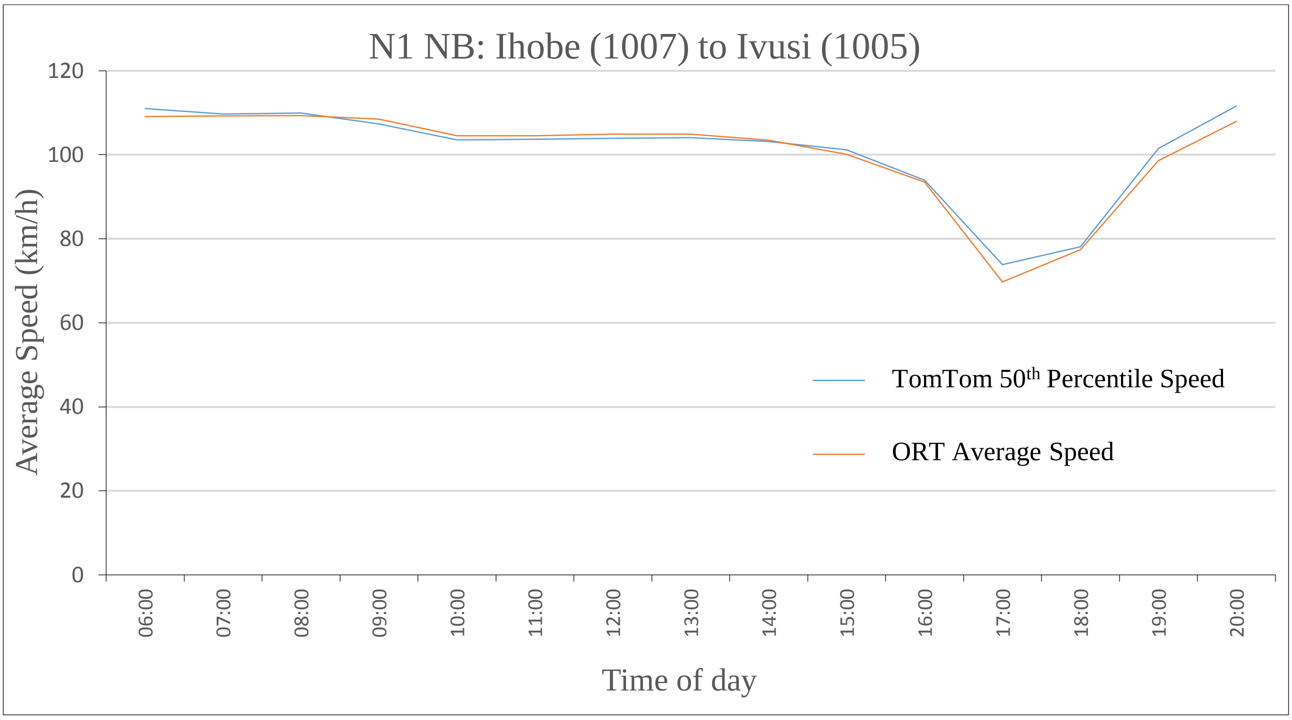


Figure B.2: Corrected hourly speed profile between Ihobe (1007) and Ivusi (1005)

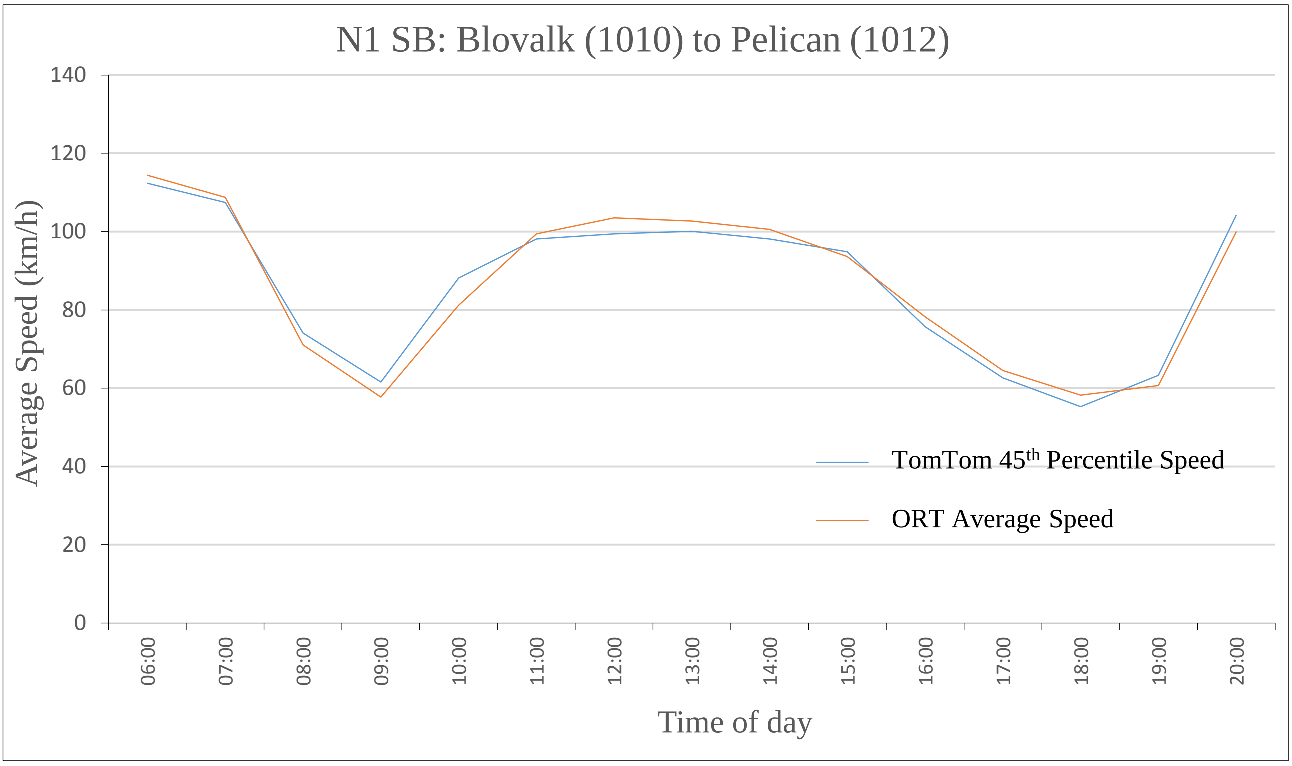


Figure B.3: Corrected hourly speed profile between Blouvalk (1010) and Pelican (1012)

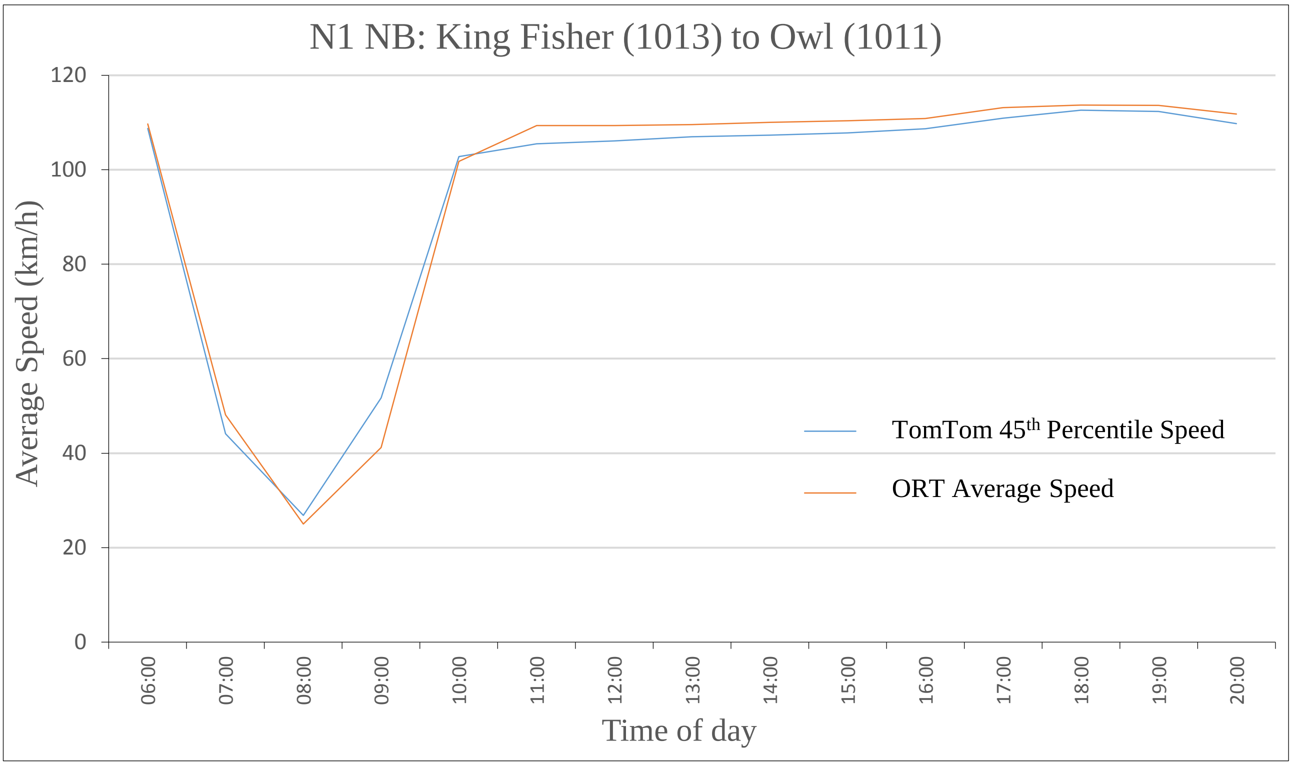


Figure B.4: Corrected hourly speed profile between King Fisher (1013) and Owl (1011)

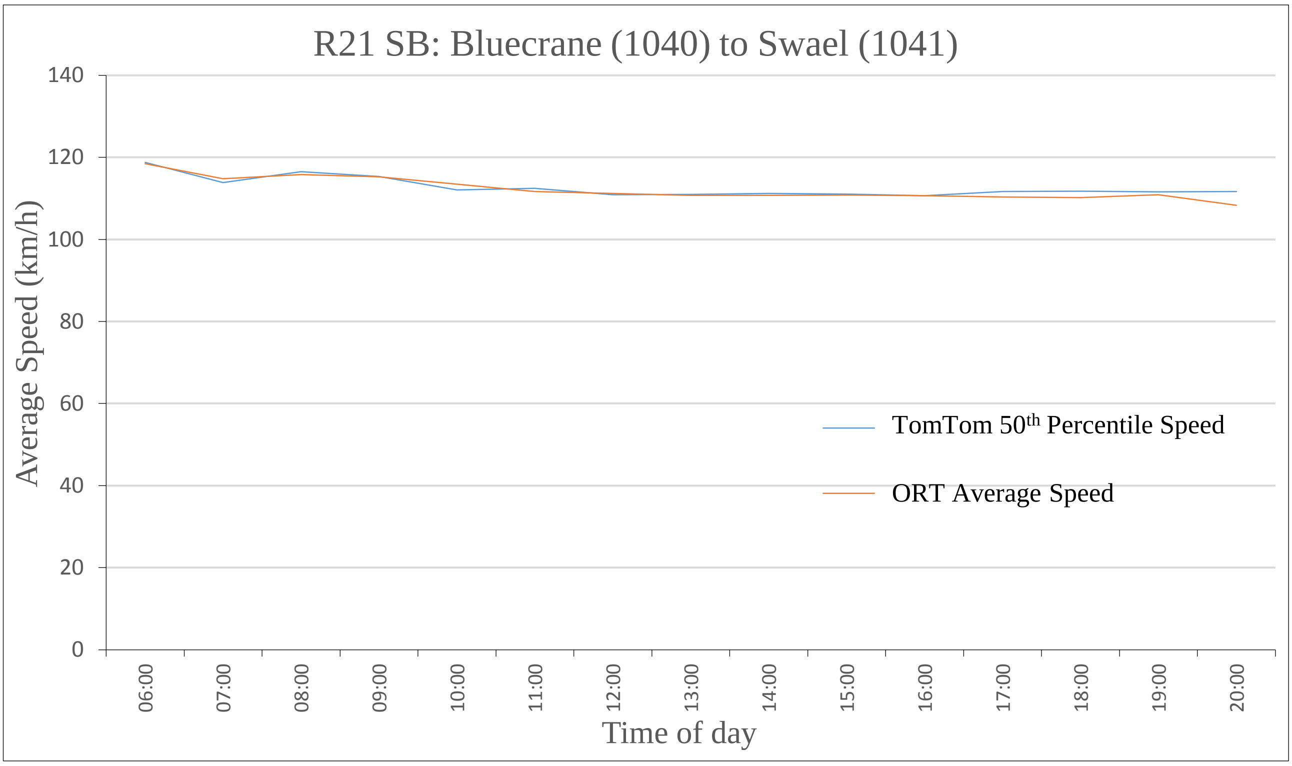


Figure B.5: Corrected hourly speed profile between Bluecrane (1040) and Swael (1041)

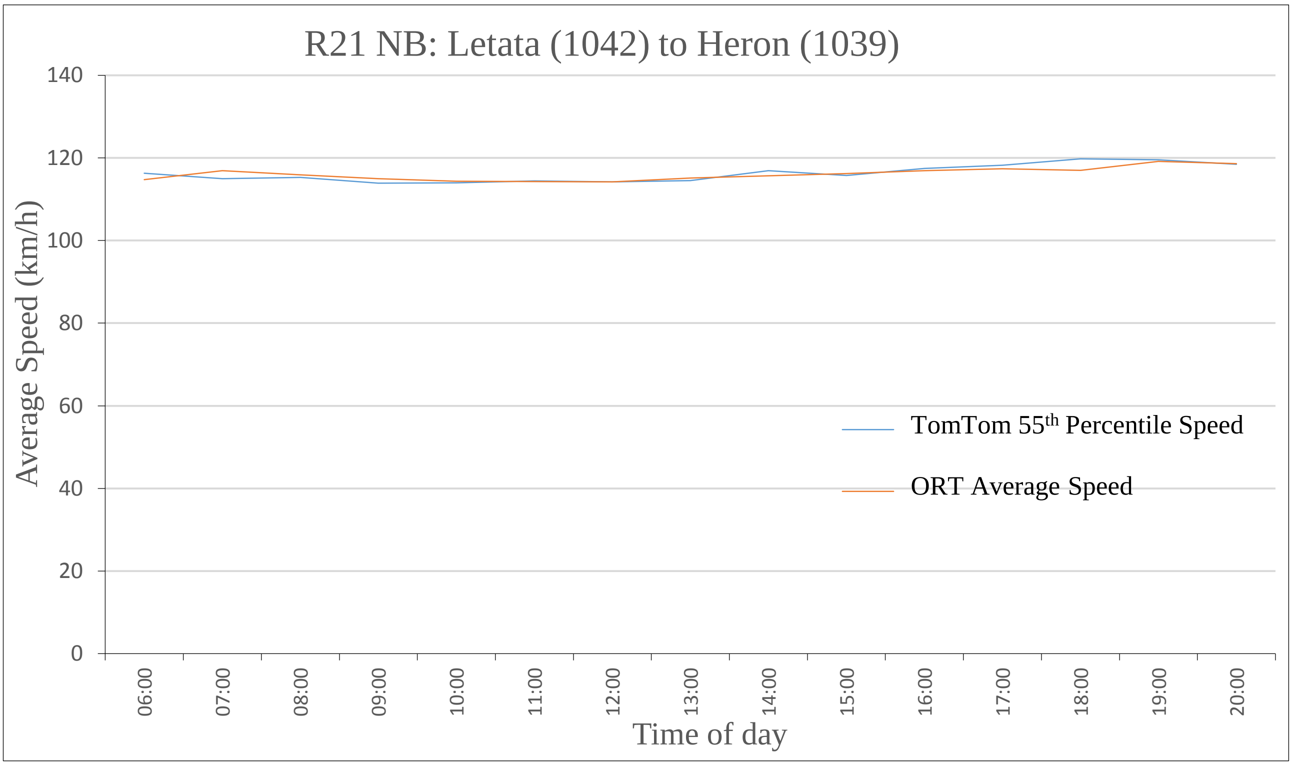


Figure B.6: Corrected hourly speed profile between Letata (1042) and Heron (1039)

APPENDIX C : CONCEPTUAL DESIGN OF THE DATA PROCESSING PROGRAM

1. Build a database with ORT observations for February 2015 from the 2.5 GB text file (Figure 3.11).
2. Create a Graphical User Interface (GUI), which allow users to enter the first gantry, second gantry and the length of the freeway segment (Figure 3.12).
3. Using the anonymised ID, identify the traffic that traverses the entire freeway section i.e. vehicles that cross the first and second gantry. Determine the travel times and speeds from the matched observations (*Matching step* – section 3.5.1.1).
4. Generate a results file with all the observations from Step 3 and allow the user to export the file (Figure 3.14).
5. Allocate the trips into 15-minute and 1-hour intervals using the *Method for allocating trips to a 15-minute interval* (section 3.5.1.2) and *Method for allocating trips to a 1-hour interval* (section 3.5.1.3).
6. Filter outliers (*Data Filtering plan* – section 3.5.1.4):
    a. Discard observations outside the analysis period (*Filter 1* – section 3.5.1.4.1).
    b. Discard illogical observations and observations with unreasonably large travel times (*Filter 2* – section 3.5.1.4.2)
    c. Discard observations with speeds that are considerably different from the average speeds observed in that particular period (*Filter 3* – section 3.5.1.4.3).
7. Calculate the space mean speed and standard deviation in each 15-minute and 1-hour interval (*Data aggregation* – section 3.5.1.5.2).
8. Generate calculations file containing the speeds placed in their respective 15-minute and 1-hour intervals (Figure 3.14).