`

ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF CIVIL AND ENVIRONMENTAL ENGINEERING

TRAVEL TIME AND RELIABLITY MODELING ON URBAN ROAD SEGMENTS (THE CASE STUDY OF ADDIS ABABA, ETHIOPIA)

By: Mekdes Ewnetu Gezahegn

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of
The Requirements for the Degree of
Master of Science
In
Road and Transport Engineering
Research Advisor: Bikila Teklu Wodajo (PhD)

June, 2019
Addis Ababa, Ethiopia

ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF CIVIL AND ENVIRONMENTAL ENGINEERING

TRAVEL TIME AND RELIABLITY MODELING ON URBAN ROAD SEGMENTS (THE CASE STUDY OF ADDIS ABABA, ETHIOPIA)

By: Mekdes Ewnetu Gezahegn

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of
The Requirements for the Degree of
Master of Science
In
Road and Transport Engineering

Approved by Board of Examiners:

Dr. Bikila Teklu	___________________	__________________
Advisor	Signature	Date
_________________	_________________	_________________
Internal Examiner	Signature	Date
__________________	__________________	__________________
External Examiner	Signature	Date
_________________	_________________	_________________
Chair Person	Signature	Date

UNDERTAKING

I, the undersigned, certify that this research work titled “***Travel time and reliability modeling at urban road segments(The Case study of Addis Ababa, Ethiopia***”) is my original work performed under the supervision of my research advisor Dr. Bikila Teklu and has not been presented elsewhere for assessment and for a degree in any other university. All sources of materials used for this thesis have also been duly acknowledged.

Signature: ______________

Undertaker’s Full Name: Mekdes Ewnetu Gezahegn
Place of Undertaking: Addis Ababa, Ethiopia.
Date: June, 2019

ABSTRACT

Nowadays, traffic congestion, delay and unreliability are terms that are most associated with present day travel. Transport users expend their precious time on long traffic queues and at the same time, fuel emission of vehicles increases in congested traffic conditions. Because of this and other problems, late arrival to work places and appointments for social or business activities have become common. And no study has developed a model for estimating travel time and reliability models on arterial roads. In this study a more general travel time and reliability model was developed. To study travel time probability distribution and reliability of a road segments in Addis Ababa city seven segments are selected. Seven segments were selected for the major modeling study. Manual method of data collection was used for collecting vehicle travel time and traffic volume data at peak hour for three days.

Travel time data for two vehicle types were fitted to a probability distribution for each segment. Reliability of a road segment was analyzed using travel time reliability measures. Multiple linear regression models were developed to predict travel time reliability of a road segments. From the regression analysis all variables were found statistically significant to predict travel time reliability of a road segment. Traffic volume, segment grade, % of heavy vehicles, number of intersection legs, signal cycle length and segment length are negatively correlated with travel time reliability of a road segment while average segment width and average speed are correlated positively. An interaction effect in regression analysis was also developed to study the effects of interaction in regression model. A developed regression model can also be used to identifying these factors is a key contribution to HCM and AASHTO where studies are being conducted to include reliability factors into roadway design features. In addition to this correlation among the influencing factors, this is necessary for traffic operations, traffic design, and long-range transportation planning and it explains the contributing impacts of other factors on travel time reliability improvement on a corridor by modifying one factor.

Keywords: *Travel Time Reliability, Travel time reliability measurements, Regression model probability distribution*

ACKNOWLEDGMENTS

Firstly, I praise God the Almighty for granting me this opportunity and allowing me to complete this thesis. My initial and final thanks always go to Him.

Secondly, I would like to express thank and deep gratitude to my Advisor, ***Dr. Bikila Teklu***, who helped me for the study to reach the final stage without any restriction entirely with his valuable guidance, encouragement, respected hospitality, kind and energetic approach.

Thirdly, I would like to thank ***Mr. Raeed Ali*** for his genuine guidance and valuable comments so that I can achieve my goal in this research.

Fourthly, it's my pleasure to express my gratitude to my beloved husband ***Dr.Legesse Demissie*** for his unconditional love and great support on this research and throughout the whole period of my graduate study.

Next, I would like to thank my friends for helping me during data collection and data extraction stage and for their valuable comments and suggestions through the whole period of my research work**.**

Moreover, I must express my very profound gratitude to my loved parents for their trust in me, their endless support, love encouragement, and their sacrifice throughout their life for me to reach to this stage. Especially my gratitude goes to my beloved sister ***Habtam Ewnetu (Habina) and Getaw Ayay.***

Finally, I would like to thank Ethiopian Roads Authority (ERA) for giving me this scholarship.

Dedication

This research is dedicated to my wonderful Parents Ato Ewnetu Gezahegn and W/ro Ageritu Ayalew with appreciation for the support, help, prayer, and encouragement that I have received during the years that I have devoted for the master’s degree.

Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

- FHWA---Federal Highway Administration
- HCM---Highway capacity manual
- BTI---Buffer Time Index
- BT--Buffer time
- PTI--Planning Time Index
- MLES---Maximum likelihood estimates
- OLS--Ordinary list square
- DOT--Department of transportation
- RTT--Reliable travel time
- H0--Null hypothesis
- H1--Alternate hypothesis
- Sig---Significance
- VIF--Variance Inflation Factor
- MLR---Multiple Linear Regression

CHAPTER ONE: INTRODUCTION

1.1. General Background

Traffic Congestion is an ever-growing chronic problem in the transportation system soon after the invention and mass production of automobiles. All major cities both in developed and developing countries are facing the problem due to increasing travel demand which follows economic and population growth. Traffic congestion directly affects commuters with an increased travel time, excessive delay in a queue, increased fuel cost, delay for important appointment and job, loss in productive hours; and it indirectly affects the living standard and the environment as well. Congestion, delay and unreliability are terms that are most associated with present-day travel. Both individuals and companies suffer from economic losses due to lost time as well as additional vehicle operating and detour costs.

Travel time has been a critical measure and important measure used to evaluate an effectiveness of the road. For this reason, travel time reliability measure is important to travelers in order for them to plan their trips effectively as well as shippers for them to plan and select routes appropriately ***(Sathya Prabha R and Mathew, 2013).***

According to***(Chen et al., 2014)*** travel time reliability is "an important measure of service quality for travelers". Transport authorities design and maintain road networks and transport systems in order to provide reliable travels to road users. But it has been a challenge to them, and therefore it is vital to identify and assess the less reliable routes in a road network.

Development of suitable metrics for travel time reliability and then ranking the different links according to the performance will be important to transport planners in order to priorities the improvements to road corridors and to implement the mitigation strategies. However, the travel time experienced by a traveler making a trip on an arterial segment is not just the result of his or her own travel choices (destination, mode, route, speed), but also the choices of many other travelers, not necessarily only those travel in the same segment. Moreover, a substantial component of driver behavior may not be classified as rational choice behavior, but rather a product of the different characteristics of individual drivers; for example attention level, driving style, risk assessment, and their vehicles, such as acceleration and deceleration capabilities ***(Van Lint,J.,2004).***

Finally, travel time reliability on an arterial segment is also determined by processes completely beyond the control of individual or groups of drivers or even the organization responsible for the road facility such as weather, calamities, incidents and accidents, traffic patterns, seasonal patterns and so on. Therefore, the travel time reliability on arterial networks is usually not only a function of traffic flow, driver behavior, traffic composition, link capacity and speed limit, but also involves numerous other factors such as signal timing, roadway and intersection geometries, adjacent land use and development, median type, signalized intersection spacing, and conflicting traffic from cross streets. In this research, the linear regression model seeks to deduce the general relationships among factors that influence travel time reliability on urban arterials and develop probably distribution models for selected types of vehicle in order to understand the nature of distribution.

1.2. Statement of the problem

Nowadays, traffic congestion, delay and unreliability are terms that are most associated with present day travel. Transport users expend their precious time on long traffic queues and at the same time, fuel emission of vehicles increases in congested traffic conditions. In addition, waiting time for the limited public transportation, both vehicle owners and public transport users, are forced to delay within the congested traffic lane especially at peak hour. Because of this and other problems, late arrival to work places and appointments for social or business activities have become common. And no study has developed a model for estimating travel time and reliability models on arterial roads in Addis Ababa city. With this in mind, the research will focus on development of travel time and reliability models in Addis Ababa city and investigate different factors that influence travel time reliability on urban road is significant.

1.3. Research questions

By the end of the research, the following questions will be answered.

- What are the factors which affects vehicle travel time reliability on urban road segment of Addis Ababa city?
- What reliability measures used to assess travel time reliability of a road segments?
- Which probability distribution is more suitable for the distribution of vehicles travel time?
- Which mathematical model is vital for modeling travel time reliability?
- What should be done to prevent those factors?

1.4. Objectives

1.4.1. General Objective

The general objective of this research is to describe and recognize the nature of vehicles travel time at urban road segment and to develop applicable travel time reliability models for urban road segment of Addis Ababa city.

1.4.2. Specific objectives

- To investigate the factors affecting travel time reliability along a section of urban road in the case study area.
- To assess travel time reliability of selected road segments using travel time reliability measurements.
- To fit probability distributions to vehicles travel time data for two vehicle types and estimate the parameters of the fitted probability distribution.
- To develop a mathematical model for vehicle travel time reliability at urban road segments in Addis Ababa city.
- To provide measures that help to improve the travel time reliability.

1.5. Scope of the study and Limitations

1.5.1. Scope of the Study

The scope of the research covers vehicle travel time reliability modeling (regression and distribution model) and calculation of reliability measurements for urban arterial roads from a study of some selected multi-lane urban road segment in Addis Ababa city. A validation study will be made at one randomly selected urban road segment of Addis Ababa city.

1.5.2. Limitations of the Study

This research doesn’t focused factors that affect travel time reliably of urban roads in non-recurrent conditions (Traffic incident, Work zone, and Weather conditions) and pedestrian activities. Due to the shortage of time and budget, the selected urban road segment for this study are only seven this may affect the statistical result of my study.

1.6. Thesis Organization

This study is organized in five chapters. Chapter one gives a brief overview of the general background of the study, statement of the problem, the general and specific objectives of the study, research question, scope and limitation of the research. Chapter two deals with review of relevant literatures. Chapter three describes research methods, materials and procedures including description of the study area, data collection, extraction and analysis methodology, sample size determination and study design. Detailed data analysis and discussion of results is presented in chapter four. Conclusion and recommendation are drawn in the last chapter.

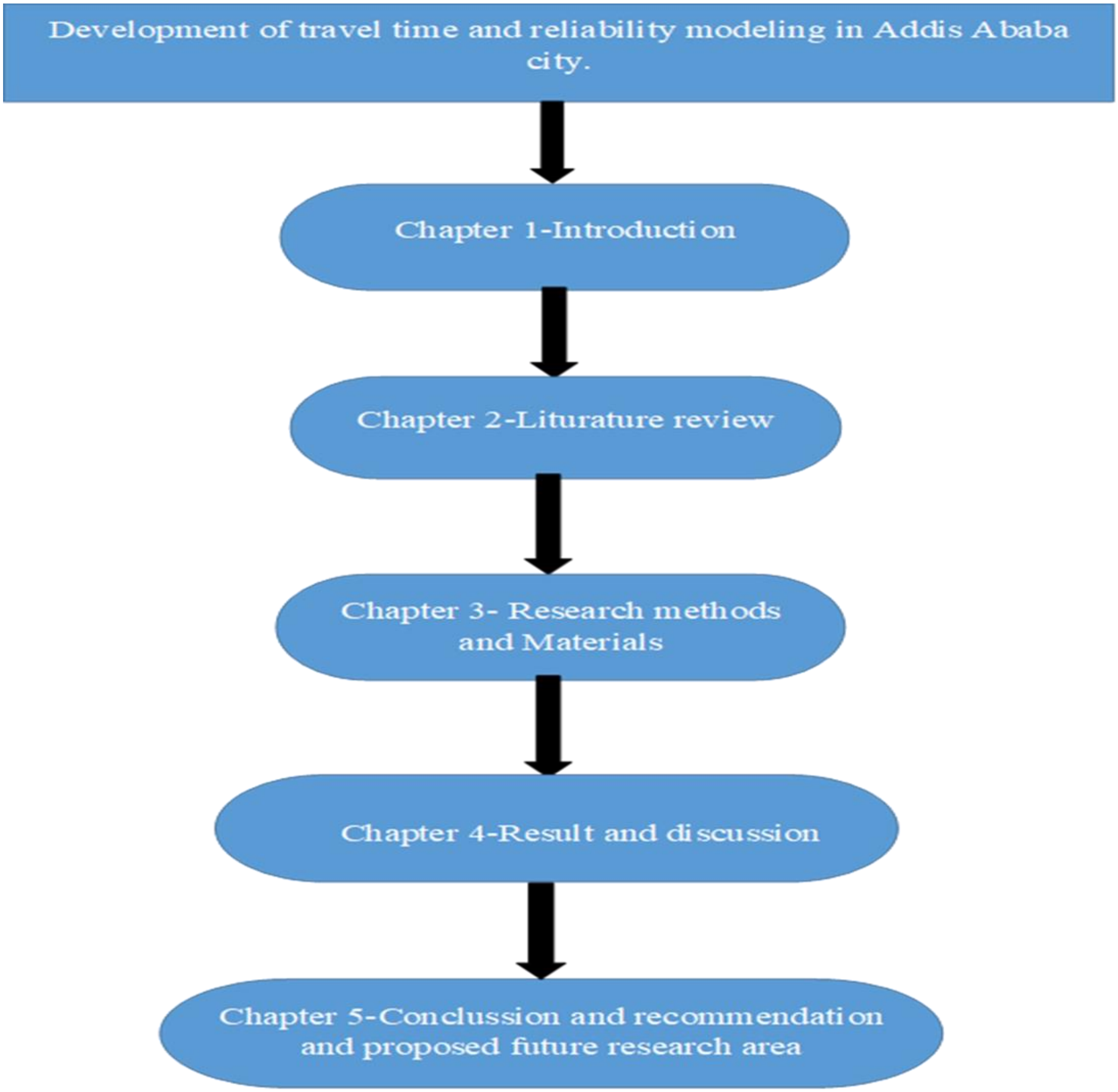


Figure 1.1: Thesis organization flow chat

CHAPTER TWO: LITERATURE REVIEW

2.1. Introduction

In the many cities where congestion is commonplace on the transportation system, drivers are accustomed to congestion and expect and plan for some increase in travel time, particularly during peak driving times. Many system users either adjust their schedules to avoid peak hours or budget extra time to allow for unexpected traffic congestion or incidents. However, problems arise when travel times are much higher than anticipated. Most travelers are less tolerant of unexpected travel time increases because those longer travel times cause travelers to be late for work or important meetings, to miss appointments, or to incur extra childcare fees. Moreover, shippers that face unexpected delays may lose money, disrupt just-in-time delivery, disrupt manufacturing processes, and lose their competitive edge on other shippers (Texas A&M Transportation Institute, 2006). Thus, transportation agencies should have a good grasp of those factors that impact travel time reliability and how travelers react to that variability and must understand how information can be used by travelers to accommodate that variability in their travel behavior. Increasing mobility and congestion results in an increase in travel time variability and in a decrease in reliability. Reliability becomes an important performance measure for transportation facilities. A variety of performance measures have been proposed to quantify it. Many of these indicators are based on percentiles of travel time **(*Guessous et al., 2014*).**

2.2. Travel Time Reliability Definitions

There are different definitions of TTR in the literature. Reliability as a concept was put forward in the literature in the early 1950s and was first implemented in the fields of communication and transport. In a study by***(Higatani et al., 2009)*** reliability was defined as "the probability that a system or component is performing its required function at a given point in time or over a stated period of time when operated and maintained in a prescribed manner. And it was realized to be a measure of how well a system meets its design objective during a given period without repair works. Florida Department of Transportation (2000) defined reliability as the percentage of travel that takes no longer than the expected travel time plus a certain acceptable additional time. California Department of Transport (DOT) (1998) defined reliability as the level of variability between the expected travel time (based on scheduled or average travel time) and the actual travel time experienced.***(Emam and Al-deek, 2006)*** proposed a definition for TTR, "a road segment is considered 100% reliable if its travel time is less than or equal to the travel time at the posted speed limit". Based on this definition, reliability is highly sensitive to the geographical location indicating the level of congestion and bottlenecks.

According to National Academy of Sciences, a typical definition for travel time reliability would be the consistency or dependability in travel times, as measured from day to day and/or across different times of the day***(Narang, 2012)***.

Travel time reliablitity is a measure of the stability of travel time and is subjected to fluctuation in flow and capacity***(Narang, 2012)*** .Generally, all definitions suggest that reliability is an indicator of operational constancy of a facility over an extended period, measured as some function of the amount of non-recurrent delay and/or recurrent delay that occurs during that period.

2.3. Importance of travel time reliability

According to U.S Department of Transportation of the Federal Highway Administration (FHWA), travel time reliability is significant to many transportation system users, whether they are vehicle drivers, transit riders, freight shippers, or even air travelers. Personal and business travelers value reliability because it allows them to make better use of their own time. Shippers and freight carriers require predictable travel times to remain competitive. Transportation planners and decision-makers should consider travel time reliability a key performance measure. Traffic professionals recognize the importance of travel time reliability because it better quantifies the benefits of traffic management and operation activities than simple averages**.**

Travel time reliability has been increasingly recognized as a key performance indicator for transportation roadways and transport systems. This has stimulated research into the development of measures to quantify the level of reliability or the extent of variability in travel times

(Doohee Nam, 2017) attempts a research to investigate the importance of travel time reliability. Travel time reliability is important to choose mode of transport to travelers in order to decrease travel time, travel cost and maximum delay.

Travel time reliability (TTR) is one of the important indexes for effectively evaluating the performance of road network (***Jiangfeng wang et al., 2017)*** It is also important to improve the travel efficiency of road network.

2.4. Travel time reliability measures

The measurement of travel time reliability is an emerging practice. However, a few measures appeared to have technical merit and were considered to be easily understood by non-technical audiences. Most of these measures compare days with high travel times to days with average travel times. Four recommended measures are as follows:

- A 90th or 95th percentile travel time;
- Buffer index;
- Planning time index; and
- Frequency the congestion exceeds some expected threshold (Texas A&M Transportation Institute, 2006).

The 90th or 95th percentile travel time is a time identified for a specific travel route that indicates how bad the delay will be on the heaviest travel days (Texas A&M Transportation Institute, 2006). These travel times are reported in minutes and seconds and were thought to be easily understood by commuters familiar with their trips. For this reason, this measure appeared to be ideally suited for traveler information. This measure has the disadvantage of not being easily compared across trips, as most trips will have different lengths. It is also difficult to combine route or trip travel times into a subarea or citywide average. Several reliability indices are presented below that enable comparisons or combinations of routes or trips with different lengths.

The buffer index represents the extra time cushion (or buffer) that most travelers add to their average travel time when planning trips to account for unforeseen delays and to ensure on-time arrival (Texas A&M Transportation Institute, 2006). The buffer index is expressed as a percentage, and its value increases as reliability gets worse. For example, a buffer index of 40 percent means that for a 20-minute average travel time, a traveler should budget an additional 8 minutes (20 minutes × 40 percent = 8 minutes) to ensure on-time arrival most of the time. In this example, the eight extra minutes is called the buffer time. The buffer index is computed as the difference between the 95th percentile travel time and average travel time, divided by the average travel time.

The planning time index represents the total travel time that a traveler should expect or plan on when an adequate buffer time is included (Texas A&M Transportation Institute, 2006). The planning time index differs from the buffer index in that it includes typical delay as well as unexpected delay. Thus, the planning time index compares near-worst-case travel time to a travel time in light or free-flow traffic. For example, a planning time index of 1.60 means that for a 15-minute trip in light traffic, the total time that should be planned for the trip is 24 minutes (15 minutes × 1.60 = 24 minutes). The planning time index is useful because it can be directly compared to the travel time index (a measure of average congestion) on similar numeric scales. The planning time index is computed as the 95th percentile travel time divided by the free-flow travel time.

2.5. Factors which affect travel time reliability

Surface transportation in urban areas, in general, experiences a problem of low travel time reliability. Non-recurring traffic congestion is one of the prime factors affecting the travel time reliability in case of surface transportation. There are function of variables affecting reliability such as volume-capacity ratio, route length, number of intersections, passenger loading, etc.(***Doohee Nam, 2017)***.

The travel time reliability depends strongly on traffic congestion, which has many sources such as bottlenecks (40%), traffic incidents (25%), bad weather (15%), work zones (10%), poor signal timing (5%), and special events (5%)***(Rich Taylor, 2010).***

The travel time reliability on arterial networks is usually not only a function of traffic flow, driver behavior, traffic composition, link capacity and speed limit, but also involves numerous other factors such as signal timing, roadway and intersection geometries, adjacent land use and development, median type, signalized intersection spacing, and conflicting traffic from cross streets ***(POLUS, 2012).***

There are factors influencing travel time reliability of motorway sections. These are primary and secondary factors. The primary factors are demand, weather conditions and accidents whereas the secondary factors are heavy goods vehicles, day of week, tome of day, local traffic rate and commuter traffic rate ***(Lohmiller, 2012).***

Several factors are likely to affect the reliability of arterial routes, one of the most important being traffic volume and average speed. Another potentially relevant factor is the amount of congestion and junction saturation levels, as related to traffic and geometric factors (e.g. Pavement width as reflected by average segment width),length of the facility(segment length) and segment grades.***(POLUS, 2012).***

2.6. Previous studies on travel time reliability regression modeling

(Chepuri, Borakanavar and Amrutsamanvar, 2018) conducted a research on examining travel time reliability under mixed traffic conditions: a case study of Urban Arterial Roads in Indian Cities. Travel time reliability indices along with descriptive statistical over selected periods are calculated and a new reliability measure called Reliable Buffer Index (RBI) is proposed for evaluating the performance of selected urban arterial road sections. This new measure has the credibility of explaining inherent causes and factors that affect travel time variations in short time intervals for a given traffic volume. The study also focuses on the determination of best fir or potential statistical distribution for analyzing travel time variations using car travel time data. The study result indicated that Buffer Times (BT) and Buffer Time Index (BTI) are the most effective measures that can capture the travel time variations. It was observed that Burr or lognormal distributions are the potential distribution for modeling car travel times under heterogeneous traffic conditions.

(***Jammula et al., 2018)*** has done a research on travel time prediction modeling in mixed traffic conditions in order to assess the impact of different travel modes on travel time. Data has been collected on a stretch of 14 km length in Warangal city, India using GPS probe vehicle along with video camera. Different private modes of transportation such as 2 wheeler, passenger car and 3 wheeler have been used as test vehicles for the collection of data in different traffic flow scenarios. Artificial Neural Network and a multi linear regression model have been developed to compare the estimated travel times with the field data. Two combinations of ANN model using single hidden layer, different numbers of neurons and epochs have been compared. The travel time of different modes has been compared and the effect of vehicle composition on travel time has been analyzed. The ANN model performs better than the regression model.

A dissertation was carried by ***(Fils, 2012)*** on Modeling Travel Time and Reliability on Urban Arterials for Recurrent Conditions This dissertation has identified 10 factors that influence travel time reliability on urban arterials by using a linear regression model. These factors could be attributed to driving behavior, driver group age, traffic congestion, inadequate lane capacity and month of the year, location of activities, special events or non-recurrent events (e.g., crashes, rain, and construction/maintenance activity). The reliability is measured in term of travel time threshold, which represents the addition of the extra time (buffer or cushion time) to average travel time when most travelers are planning trips to ensure on-time arrival. This dissertation has established baselines for researches to identify the other influencing factors which represent 32.6% of the observed data.

(Feng et al., 2011), conducted a study on Travel Time Reliability Modeling. In this paper, a novel multi-state model is proposed for travel time modeling and reporting. The model advances travel time modeling in two aspects. First, the multi-state model provides much improved model fitting as compared to single-mode models. Second and more importantly, the proposed model provides a connection between travel time distributions and the underlying travel time state. This connection allows for the quantitative evaluation of probability of each travel time state as well as the uncertainty associated with each state; e.g., the probability of encountering congestion on a given time of day and the expected travel time if congestion were experienced. A simulation study was conducted to demonstrate the performance of the model. The proposed model was applied to field data collected at San Antonio, TX. The variation of the model parameters as a function of time-of-day was also investigated. The simulation study and field data analysis confirmed the superiority of multi-state model over the state-of-practice single-mode travel time reliability models and ease of interpretation.

(*Rakha et al., 2015*) has studied on Bayesian Travel Time Reliability Models. This study built up and advanced the multi-state models by proposing regressions on the proportions and distribution parameters for underlying traffic states. The Bayesian analysis provides valid credible intervals for each parameter without asymptotic assumption. Two alternative approaches were proposed and evaluated. The first approach is a Bayesian multi-state travel time regression model which provides a regression for key model parameters to traffic volume; the second approach is a hidden Markov regression which not only provides a link between key model parameters and traffic volume, but also incorporates the dependency structure among traffic volume in adjacent time windows. Both approaches provide advanced methodology for modeling traffic time reliability under complex stochastic scenarios.

(*Bertini, 2008)* conducted a paper Using Travel Time Reliability Measures to Improve Regional Transportation Planning and Operations. This paper examines the uses of measured travel time reliability indices for improving real-time transportation management and traveler information using archived ITS data. Beginning with a literature review of travel time reliability and its value as a congestion measure, a content analysis of twenty regional transportation plans from across the nation is then described. Results from the content analysis indicate that travel time reliability is not currently used as a congestion measure, and that the most common measures of congestion were the volume-to-capacity ratio, vehicle hours of delay, and mean speed. Then, as a case study using data from Portland, Oregon, several reliability measures are tested including travel time, 95th percentile travel time, travel time index, buffer index, planning time index, and congestion frequency. The buffer index is used to prioritize freeway corridors according to travel time reliability. Metropolitan planning organizations should use travel time reliability in the following ways: 1) incorporate it as a system-wide goal; 2) evaluate roadway segments according to travel time reliability measures; and 3) prioritize roadway segments using those measures.

2.6.1. Regression Parameter Estimation Methods

Two parameter estimators, Maximum Likelihood and Ordinary Least Squares which are equivalent for the normal theory regression model, are considered as the best linear unbiased estimators of the underlying population parameter***(Montgomery and Runger, 2012).***

- Maximum Likelihood Estimation

Maximum Likelihood Estimation is another popular and sometimes useful statistical estimation method is called maximum likelihood estimation, which results in the maximum likelihood estimates, or MLEs. ***(Washington, 2003).***

- Least Square Estimation

Least squares estimation is a commonly employed estimation method for regression applications. Often referred to as "ordinary least squares" or OLS, it represents a method for estimating regression model parameters using the sample data. **(*Washington, 2003*).**

2.6.2. Regression Model Goodness-of-Fit Measures

Goodness-of-fit (GOF) statistics are useful for comparing the results across multiple studies, for comparing competing models within a single study, and for providing feedback on the extent of knowledge about the uncertainty involved with the phenomenon of interest. Three measures of model GOF are discussed: R-squared, adjusted R-squared, and the generalized F test **(*Montgomery and Runger, 2012*).**

2.6.3. Interactions in Regression Models

Interactions in regression models represent a combined or synergistic effect of two or more variables. That is, the response variable depends on the joint values of two or more variables. A second-order effect is an interaction between two variables, typically denoted as X1X2. A third-order effect is an interaction between three variables, and so on. As a general rule, the lower the order of effect (first order being a non-interacted variable), the more influence the variable has on the response variable in the regression. In addition, higher-order effects are generally included in the regression only when their lower-order counterparts are also included **(*Washington, 2003*)**.

The presence of a significant interaction indicates that the effect of one predictor variable on the response variable is different at different values of the other predictor variable ***(Grace-Martin, 2000).*** Part of the multiple linear regression is the ability to estimate and test interaction effects when the predictor variables are either categorical or continuous **(*Stevens, 2010*)** .In a linear model representing the variation in a dependent variable Y as a linear function of several explanatory variables, interaction between two explanatory variables X1 and X2 can be represented by their product, which is the variable created by multiplying them together ***(Burrill, 2011)***. Algebraically such a model is represented by the following equation:

2.7. Travel time probability distribution models

Increasing mobility and congestion results in an increase in travel time variability and in a decrease in reliability. Reliability becomes an important performance measure for transportation facilities. A variety of performance measures have been proposed to quantify it. Many of these indicators are based on percentiles of travel time. The knowledge of the distribution of travel time is needed to properly estimate these values. Congestion distorts the distribution and particular statistical distributions are needed.

Different distributions have been proposed in the literature. In a previous paper, we presented a comparison of six statistical distributions used to model travel time. These six distributions are the Lognormal, Gamma, Burr (extended by Singh -Maddala), Weibull, a mixture of two Normal distributions and a mixture of two Gamma distributions**(*Guessous et al., 2014*)**

Travel time distribution is an important basis for modeling travel time variability and reliability, which can be measured using several travel time distribution properties, such as standard deviation and coefficient of variation. Previous studies on travel time variability and reliability assumed that travel time distribution may follow either normal distribution or log-normal distribution particularly for freeways. However, several studies suggested that travel time data are skewed and that such data possess long upper tail***(Li, Chai and Tang, 2013)***. Numerous studies have been conducted to investigate the probability distribution of travel time on freeways and signalized arterial roads.

(Li, Chai and Tang, 2013)studied Seven signalized arterials with lengths ranging from 417m to 2028m and 10 non signalized urban expressways with lengths ranging from 1600m to 4100m.A 15min interval is used to study how the travel time distribution on actual road varies over time. This interval provides sufficient data for most times of a day, and these data can be used for travel time distribution estimation. In addition, this interval is short enough to capture short-term. The data collected on June 16, 2011, were selected for used in this paper. For each road link, a total of 96 15 min travel time datasets may be used for analysis. Normal, log-normal, gamma, and Weibull distributions are fitted to these 15 min travel time datasets, and the Chi squares are selected for testing goodness-of-fit variations in travel time.

2.7.1. Probability distribution Goodness of fit tests

The goodness-of-fit test is widely used when checking whether the underlying population distribution differs from a specified distribution. (***Liu, 2016***).There are two commonly used goodness-of-fit tests such as, the Kolmogorov-Smirnov test and the Anderson-Darling test

- The Anderson-Darling Test

Andersen Darling (AD) test is the most popular and more powerful than K-S test under some situations. For instance, when testing the normality of the observed data, Anderson-Darling test provides one of the most powerful statistic for detecting a normal distribution adequacy.

Anderson and Darling (1952) further adapted the Cramer-von Mises test, and introduced a new test statistic A_2, calculated by

$$\boldsymbol{n}\int_{-\infty}^{\infty}\frac{[\boldsymbol{Fn}(\boldsymbol{x})-\boldsymbol{F}^{*}(\boldsymbol{x})]^{2}}{\boldsymbol{F}^{*}(\boldsymbol{x})-(\boldsymbol{F}^{*}(\boldsymbol{x}))^{2}}\,\boldsymbol{dF}^{*}(\boldsymbol{x})$$

- The Kolmogorov-Smirnoff Test

Kolmogorov-Smirnov (K-S) test was proposed by Kolmogorov and Smirnov (1933 and 1939). The K-S test statistic D measures the distance between the empirical distributions function (EDF) using the observed data and the hypothesized distribution function F(x).The test statistic of the K-S test can be written as

$Dn = \sup |Fn^*(x) - F(x)|$

2.8. Simulation (computer) based travel time reliability models

Simulation models are capable of replicating the behavior of real-world systems ***(Liu et al. , 2009)*** Traffic simulations are computer programs based on various types of traffic models such as discrete time/discrete event models, micro/macroscopic/macro models, deterministic/stochastic models, and Dynamic Traffic Assignment (DTA) models. These models simulate traffic movements over a user-defined transportation network and extract data such as travel time, delay, and speed. In addition, most of the traffic simulation models are micro models, as well as stochastic models***(Oregon Dot, 2006).***

Although these models are capable of analyzing various types of situations, they require significant effort to create the network, and calibrate and validate the models. These models can be used for both travel time estimation and prediction.

(Shen and Hadu, 2009) utilized a traffic simulator for travel time estimation on a freeway segment of about 8 miles in Florida, USA. They proposed the best model based on the traffic simulation results.

2.9. Research Gap

There are no local study using regression approach to model travel time and reliability modeling. But in this research try to investigate factors that affect travel time and reliability and develop multiple regression modeling and see interaction effect independent variables. Which used to evaluate strategies and tactics to satisfy the travel time reliability requirements of users of the roadway network those engaged in person transport in urban areas, monitor the performance of road network, evaluate future options and provide guidance on transportation planning, roadway design, traffic design, and traffic operations features.

CHAPTER 3: RESEARCH METHODS AND MATERIALS

3.1. Introduction

In this chapter, the research materials and methodology elements are presented. Therefore, selection of the study area, study design, and data collection methods are briefly described by sections as follows.

3.2. Description of study area

The study area selected for this research is Addis Ababa city which is the capital city of Ethiopia. Addis Ababa is not only the capital city of Ethiopia but it is also the seat of African Union head quarter and more than 100 embassies. Due to the fact that Addis Ababa is the political and economic center of the nation, it is the highly populated town in the country. Most of the economic and social developments in the country manifested at this capital city and hence all the benefits and aftermath of such economic and population growth affect Addis Ababa. In order to study traffic congestion in Ethiopia, there is no a best place like Addis Ababa due to many factors. Hence, this research focuses on the Addis Ababa city and this section of the research describes briefly the study area and the selected corridors.

3.2.1. Study Location

In order to conduct this study two congested corridors in Addis Ababa were selected. Those corridors separate in to seven segments. Among these 6 segments are used for the analysis purpose and one for validation. The attributes segments and corridors are illustrated shown below.

Table 3.1: Segment Attributes

S.NO	From	To	Length(m)
1	Estifanos	St. Joseph school	378
2	St. Joseph school	Leghar	705
3	Leghar	Telebar	374
4	Telebar	Buna ena Shay	422
5	Beharwi	ETV	242
6	ETV	Tikur Ambesa	245
7	Tikur Ambesa	Tewedros round about	596

Figure 3.1: Corridor 1 from Estifanos to Buna shay

Figure 3.2: Corridor 2 from Beharwi to Tewedros Round About

3.3. Study subjects

The primary study subjects of this research were vehicle travel time, traffic volume, segment grade, average segment width, % of heavy vehicles, number of intersection legs, signal cycle length, segment length and average speeds.

3.4. Research Materials

According to different studies, travel time of vehicles is affected by different factors. Those are segment length, average segment width, speed, % of heavy vehicles; number of intersection legs, signal cycle length, segment grade (up/downs) and traffic volume. In this study, eight variables are selected to capture the effect of travel time reliability. To capture the effect of segment geometry segment grade, segment length, number of intersection legs and average segment width are considered. % of heavy vehicle is used to account for the composition of vehicles. The effect of signal characteristics is taken through signal cycle length.

3.4.1. Data Requirement

Data required for the research are classified in to three categories and these are Traffic related data, Signal related data, and Geometric related data.

Traffic related data

- Travel time data
- Traffic volume data
- Average speed data
- % of heavy vehicles

Signal related data

- Intersection cycle length

Segment geometry related data

- Segment length (m)
- Segment grade (%)
- Average segment width
- Number of intersection legs

3.4.2. Sample size

Once the manually recorded data was acquired, extracting travel time taken by an individual vehicle to travel a specified length of a road section was determined by tracing every individual vehicle. Since many vehicles negotiate the entry point at a time, vehicles were selected randomly but statistically significant sample size was determined for each 15min of count. The sample size was determined according to the procedure and equation on the handbook. According to travel time data collection Handbook the sample size for manually transcript travel time data is given by the equation;

$$\textit{Sample size for travel time study } (n) = \left(\frac{Z\sigma/2\sigma}{E}\right)^{2}$$

Where: - E is margin of error, σ is standard deviation, ZC is the Z score. For a 95% confidence level, E is 0.05 and σ is 0.5.

However, the handbook using the above statistical equation provides a sample sizes for different traffic conditions and level of confidence. Accordingly, for congested traffic condition at 95%

confidence interval and 5% error the minimum sample size is calculated to be 385.

On the other hand, based on (***Vanvoorhis, C.W. and Morgan, B.L., 2007*** the minimum sample size recommended regression analysis is,

$$N=50+8*m$$

Where m is the number of factors considered in the model. Therefore, for maximum number of independent factors to be considered in this study, the minimum sample size is 114.

3.4.3. Sampling method

Convenience sampling method is a type of non-probability sampling that involves the being drawn from that part of the population that is close to the hand.in this research convenience method of sampling was applied to collect the number of vehicle travel time that were available at the time of data collection.

A total of 1897 travel time data was collected for regression model development and 702 vehicle travel time data was collected for travel time distribution modeling.

3.5. Methods of data collection

In order to achieve the objective of this research, only primary data was collected.

3.5.1. Primary data collections

For primary data collection two types of survey were conducted, namely manual methods of vehicle license plate survey and manual traffic volume count survey on the subject route with following details:

- ❖ **Primary Data type** – Vehicle travel time and traffic volume data.
- ❖ **Locations**– The recording were taken at entry and exit of each segment for travel time data and Traffic volume data was also taken parallel to travel time data at the mid-block.
- ❖ **Day of the week: -** This research is focused on the middle weekdays (i.e., **Tuesday, Wednesday, and Thursday**) for data collection. Monday, Friday, and holidays are excluded because of their high variation from typical day-to-day operating conditions during the middle of the week.
- ❖ **Time Periods: -** Morning Peak Period: Between 7:00 am to 9:00 am.

3.5.1.1. Manual methods Vehicle license plate survey

This survey where carried out at entry and exit location of the study area. The objective of this survey is to estimate travel time for the study area for the study period. Travel time can be defined as the period of time to transverse a route between any two points of interest. It is a fundamental measure in transportation and it is one of the most readily understood communicated measures indices used by a wide variety of transportation analysis including congestion management, transportation planning and traveler information's. Travel time data is analyzed in terms of three trip components recording during the trip including waiting time, walking time and in vehicle travel time (travel time in motion, waiting time in queue, and moving time in queue). But in this research by recording vehicle travel time in order to check the reliability of road sections and to develop travel time reliability models. In this survey, vehicle license plate numbers and their arrival times at entry and exit points of the section was collected. Subsequently, matching of the license plates at entry and exit points was done and travel time data was calculated from the difference in arrival times at these points. Manual travel time recording at entry and exit location of the study area to record the vehicle license plate for all categories of vehicles. (i.e., motor cycle, car, Minibus, Mid Bus, Bus, Large Bus, Medium Truck and Truck).

3.5.1.2. Traffic volume count survey

This survey was conducted in the morning peak hours for the same period, i.e. 7:00 am to 9:00 am to assess the traffic volume and composition of traffic in the study area. Consistently middle day of data collection (Tuesday, Wednesday, and Thursday) data have been collected to know the day-to-day variation of traffic for the particular period. In addition to the above travel time and traffic volume data collection techniques other field measurements were done to gather data on the geometrical features of road segments for regression model developments. These include, segment grade, average segment width, segment length signal cycle length, number of intersection legs and average speed were collected at the selected study sites using the following specific methodology.

- Average segment width: - measuring with roller meter
- Segment length: - measuring with roller meter
- Approach grade: - GPS72H leveling device was used to measure the grade at each segment
- Signal cycle length: - Recording at the road segments.
- Number of intersection legs: - collected by counting for each intersection.
- Average speed: - calculated from segment length and each travel times.

3.5.2. Data Extraction methods

Travel time data and traffic volume data was extracted manually from data collection sheet. Matching the license plate between entry and exit points and computing travel time from the difference in arrival time.

3.6. Data Analysis methodology

This section describes the overall data analysis methodology of the research. Minitab 18 statistical analysis software is used as a major data analysis and manipulation tool.

3.6.1. Probability distribution modeling

Travel time distribution modeling in travel time is the most useful indicator to measure the performance and reliability of a transportation system. And it is an important basis for modeling travel reliability. In this research Probability distribution is fitted to travel time data of two vehicle types (car and minibus). Before fitting any probability distribution, the data is tested for outliers using Grubbs' test and are removed from the analysis. Then the data is tested for normality using Anderson-Darling normality test and for non-normal data other probability distributions are fitted. Probability-Probability (P-P) plots and Anderson-Darling test are considered to test the existing distribution models. The following are some common distributions that are usually used for fitting travel time data. Normal, Lognormal, 3-Parameter Lognormal, Weibull, 3-Parameter Weibull, Smallest Extreme Value, Largest Extreme Value, Gamma, 3-Parameter Gamma, Logistic, Log-logistic, 3-Parameter Log-logistic.

Grubbs' test with the following Null and Alternate Hypothesis is used to determine selected (car and minibus) vehicle travel time data outliers.

Ho: There are no outliers in the data set

Ha: There is exactly one outlier in the data set

Anderson-Darling Test with the following Null and Alternate Hypothesis is used to test for normality.

Ho: The data follow a specified distribution.

Ha: The data do not follow the specified distribution

3.6.1.1. Parameter estimation

There are various methods that used to estimate parameters which are maximum likelihood and least square. But maximum likelihood is the most popular and easy to apply so in this research maximum likelihood method of estimation was used.

3.6.1.2. Probability distribution Goodness of fit tests

The goodness-of-fit test is widely used when checking whether the underlying population distribution differs from a specified distribution. **(Liu, 2016)**.There are two commonly used goodness-of-fit tests such as, the Kolmogorov-Smirnov test and the Anderson-Darling test.

- The Anderson-Darling Test

Andersen Darling (AD) test is the most popular and more powerful than K-S test under some situations. For instance, when testing the normality of the observed data, Anderson-Darling test provides one of the most powerful statistics for detecting a normal distribution adequacy.

Anderson and Darling (1952) further adapted the Cramer-von Mises test, and introduced a new test statistic A2, calculated by

$$\boldsymbol{n}\int_{-\infty}^{\infty}\frac{[\boldsymbol{Fn}(\boldsymbol{x})-\boldsymbol{F}^{*}(\boldsymbol{x})]^{2}}{\boldsymbol{F}^{*}(\boldsymbol{x})-(\boldsymbol{F}^{*}(\boldsymbol{x}))^{2}}\,\boldsymbol{dF}^{*}(\boldsymbol{x})$$

- The Kolmogorov-Smirnoff Test

Kolmogorov-Smirnov (K-S) test was proposed by Kolmogorov and Smirnov (1933 and 1939). The K-S test statistic D measures the distance between the empirical distributions function (EDF) using the observed data and the hypothesized distribution function F(x).The test statistic of the K-S test can be written as

$$Dn=\sup|Fn^{*}(x)-F(x)|$$

3.6.2. Analysis of travel time reliability

Travel time reliability is defined as the consistency or dependability in travel time, as measured from day-to-day and/or across different times of the day. The planning time (PT), planning time index (PTI), buff time and buffer time index (BTI) will be used as performance indicators. **Planning time**: -represents the total travel time expected or ‘planned’ before a trip starts with a given probability. It is the 95th percentile of the measured travel time. It represents travel time on some of the heaviest traffic days.

Planning time index: - is the ratio of the planning time to the free-flow travel time. This index indicates the severity of traffic congestion as it represents the worst level of congestion at a given time of day in compare-son with the free-flow traffic condition.

Buffer time: - is defined as the difference between planning time and average travel time. It represents the extra time to ensure on-time arrival to the destination. This extra time accounts for any unexpected delay. In other words, the buffer time may be almost proportional to the variance of travel time.

Buffer time index: - is defined as the ratio of buffer time to the average travel time.

(Ch. Ravi Sekhar, E. Madhu, 2012) This article has used the use of reliability measures recommended by USDOT, NCHRP, SHRP2 and Lenox et al as buffer time and he stated that a segment is considered reliable when travel time threshold is equal to or lowers than the summation of the buffer time and the average travel time.

3.6.3. Travel Time Reliability Calculation Method

According to ***(Fils, 2012)*** Buffer time is a reliability measurement which is recommended by USDOT, NCHRP, SHRP2. The buffer time represents the extra time (buffer or cushion time) that most travelers add to their average travel time when planning trips to ensure on-time arrival. The buffer index, another reliability measure, is expressed as a percentage and its value increases as reliability gets worse. A segment is considered reliable when travel time threshold is equal to or lowers than the summation of the buffer time and the average travel time.

Buffer time and buffer index will be computed as follows:

Buffer time (sec) = $95\text{th percentile travel time (sec)} - \text{Average travel time (sec)}$

Buffer index (%) =$\frac{95\text{th percentile travel time (sec)} - \text{Average travel time (sec)}}{\text{Average travel time (sec)}}$

Where 95^{th} percentile travel time indicates how bad delay will be on the heaviest travel days and is a translation of a standard "I can be late to work 1 day a month (1 day out of 20 ± work days) without getting into too much trouble" (***Lomax et al., 2003***).

From the above two equations travel time reliability threshold (in seconds) for the research is deducted as follows:

Travel time threshold (reliability)$\leq$ $\text{Buffer time} + \text{Average travel time}$

3.6.4. Travel time reliability regression modeling

The linear relationship between travel time (dependent variable) and the influencing factors (independent variables) is computed by using the multiple linear regression equation. Multiple linear regressions are used to model travel time of a vehicles. A stepwise multiple linear regression analysis will be performed to model travel time reliability of vehicles at signalized urban segments with the following explanatory variables. The explanatory variables consist only continuous variables. Segment length, average segment width, traffic volume, % of heavy vehicles, segment grade, and number of intersection legs, signal cycle length and operating speeds were considered as continuous variables.

Dependent variable

- Vehicle travel time

Predictor variables

- Segment length (meter)
- Average segment width (meter)
- Number of intersection legs (Three leg, four leg and five legs)
- Segment grade in % (Grade)
- Traffic volume
- % of heavy vehicles (% of large bus, % of medium track and % of tracks)
- Signal cycle length(second)
- Average speed(m/s)

Multiple linear regression is one of the most widely studied and applicable statistical and economic techniques for the following reasons. The major part of a regression analysis is checking the assumptions. Before conducting any regression analysis, the assumptions should be checked. The following multiple linear regression assumptions are checked for violation before performing the regression analysis. In this research these are regression assumptions that should be checked before regression model development.

- Relationship between the independent and dependent variables to be linear.
- Multicollinearity
- Homoscedasticity (residuals have a constant variance along the fitted value).
- There should be no significant outliers
- residuals (errors) of the regression line are approximately normally distribute

3.6.4.2. Regression Model Goodness-of-Fit Measures

Goodness-of-fit (GOF) statistics are useful for comparing the results across multiple studies, for comparing competing models within a single study, and for providing feedback on the extent of knowledge about the uncertainty involved with the phenomenon of interest. In this research two measures of model GOF test was applied that are R-square and adjusted R-squared.

3.6.5. Interactions in Regression Models

Interactions in regression models represent a combined or synergistic effect of two or more variables. That is, the response variable depends on the joint values of two or more variables. A second-order effect is an interaction between two variables, typically denoted as X1X2. A third-order effect is an interaction between three variables, and so on. So, the researcher wants to check the synergetic effects of two independent variable. Algebraically such a model is represented by the following equation:

$$Y = \beta_O + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \epsilon$$

CHAPTER FOUR: RESULTS AND DISCUSSION

4.1. Introduction

In this chapter, the research findings are discussed analytically, statistically, graphically, in tabular form and qualitatively. In this regard, the chapters are divided into three parts: travel time distribution modeling, travel time reliability analysis and travel time reliability regression modeling.

4.2. Travel time distribution modeling

In this study, the distribution model is developed for the collected travel time data. Since the traffic flow condition at the selected road segment is Heterogeneous, it is not reasonable to treat all vehicle types under the same category. So, travel time data distributions will be fitted for each vehicle type. But in this study, I only consider two vehicle types because when I see from the traffic composition car and minibus was the dominant vehicle type. Before fitting any distribution to our data first we have to test it for normality. In order to try to fit any non-normal distributions to a data, the data should show a significant departure from normality. To determine this we should look at the probability plot of the data, histogram, and Anderson –Darling test results. The Null and Alternate hypothesis for Anderson-Darling test are as follows.

- Anderson-Darling Test

Ho: - Data values do come from the specified probability distribution
H1: - Data values do not come from the specified probability distribution.
Significance level: - 0.05

- Distribution fitting for travel time data of Cars for segment 1

Distribution fitting for travel time data of Cars for segment 1 shows that the shape of the curve on the histogram approximates to a bell shape with a peek value at the middle. This shows that the data is from a normal distribution. And the normal probability plot graph shows the pattern of the data approximately following the straight line strengthening the conclusion that the data is from normal distribution. The p-value from Anderson-Darling test result is *0.145* which is greater than the significance level ***(0.05)*** and this confirms that the data come from normally distributed population. And which leads to the conclusion to reject the alternate hypothesis which states that the data does not come from a normal distribution.

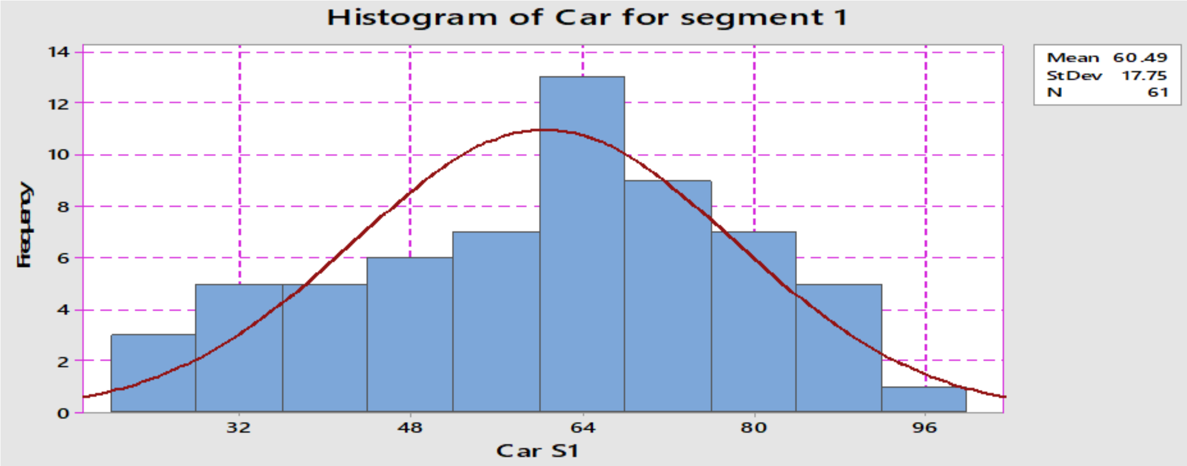


Figure 4.1: Histogram plot of car (segment 1)

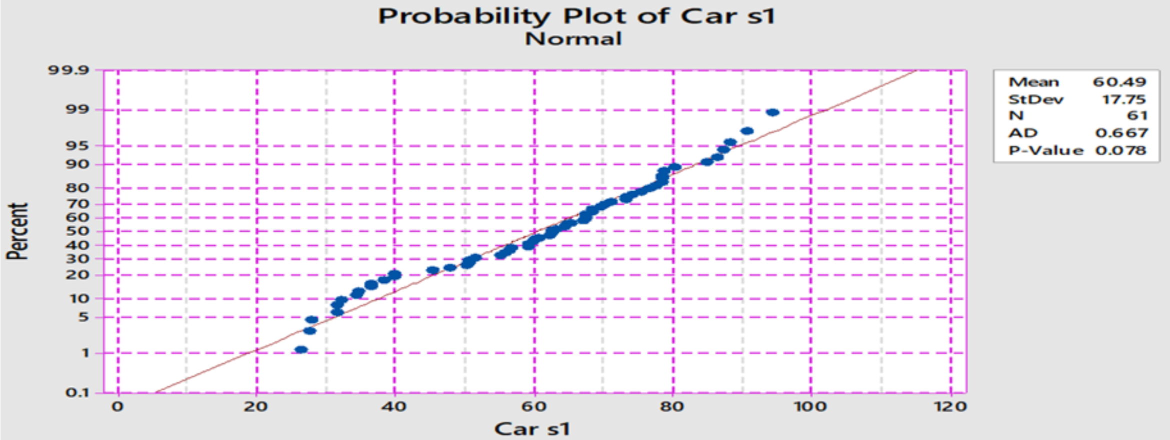


Figure 4.2: Normal probability plot of car (segment 1)

- Distribution fitting for travel time data minibus for segment 1

Probability plot of minibus shows that smaller and larger data values depart from the normality line The fit curve on the histogram is also a little bit skewed to the right implying a probable departure from normality***)*** but we have to look at Anderson-Darling test result, provided on the probability plot, to be sure that the data significantly departs from normality. The p-value from Anderson-Darling test is ***0.011***, less than significance level ***(0.05)*** which leads to the conclusion to accept the alternate hypothesis which states that the data does not come from a normal distribution.

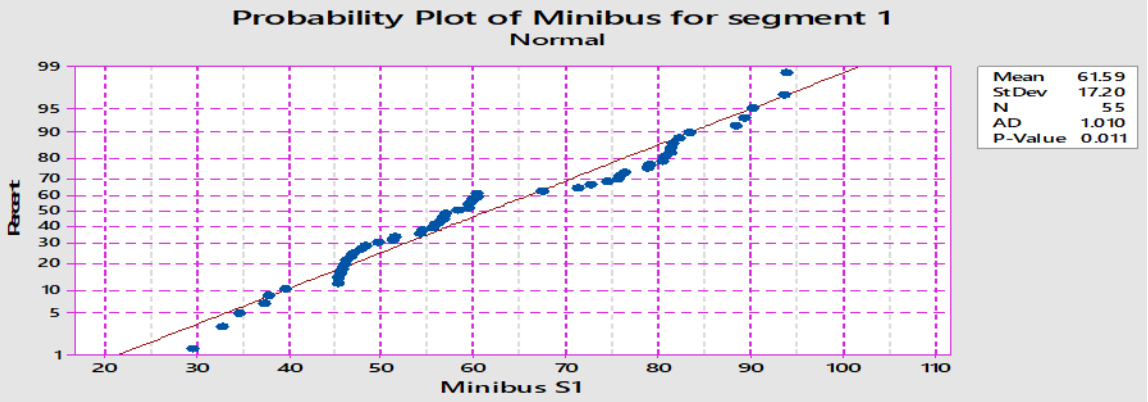


Figure 4.3: Normal probability plot of minibus (segment 1)

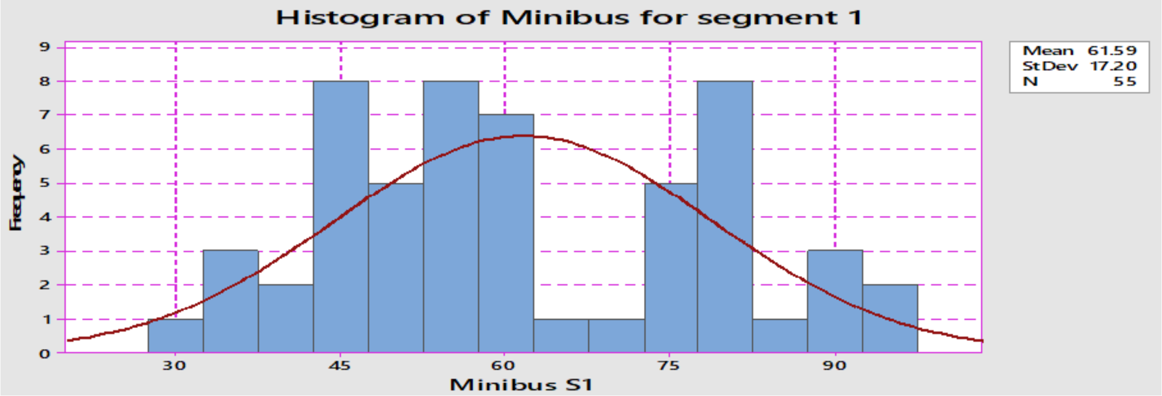


Figure 4.4: Histogram plot of minibus (segment 1)

As can be seen from figure # 4.3 and 4.4 the data is not normally distributed so it was tried to fit the data to other distributions. Probability plot and Anderson-Darling test are used to determine which probability distribution fit the data best. All possible distributions were tested to fit travel time data for minibus. Box-cox transformation, Lognormal,3-parametric lognormal, exponential,2-parametric exponential, Weibull, 3-parametric Weibull, smallest extreme value, gamma, 3-parametric gamma and logistic distributions were found to fit the data but we have to find out which probability distribution fit the data best. The smaller the Anderson-Darling test parameter. (AD), the better the fit and the higher the p-value the better the fit. The AD value for the lognormal distribution is greater than the AD value for the large extreme distribution but the p-value of the two distributions is very close to each other in this case we should carefully examine the probability plot of the distributions. If a certain theoretical probability distribution function is a good fit of a data, the data points cluster in to the straight line in the probability plot. Large extreme value distribution probability plot has a smaller number of data points which lie of the ***95%*** confidence interval limit and the points are also closer to the straight line than the lognormal probability plot. So the large extreme value probability distribution is a better fit for the travel time data of minibus for segment 1. The table below shows estimated distribution parameter and Anderson-Darling test result.

Table 4.1: Estimated Parameters and AD Test Result

Distribution	Estimated parameters		Anderson–Darling test result	
	Location	Scale	AD	P
Lognormal	4.081	0.29	0.737	0.052
Large extreme value	53.28	15.01	0.727	0.054

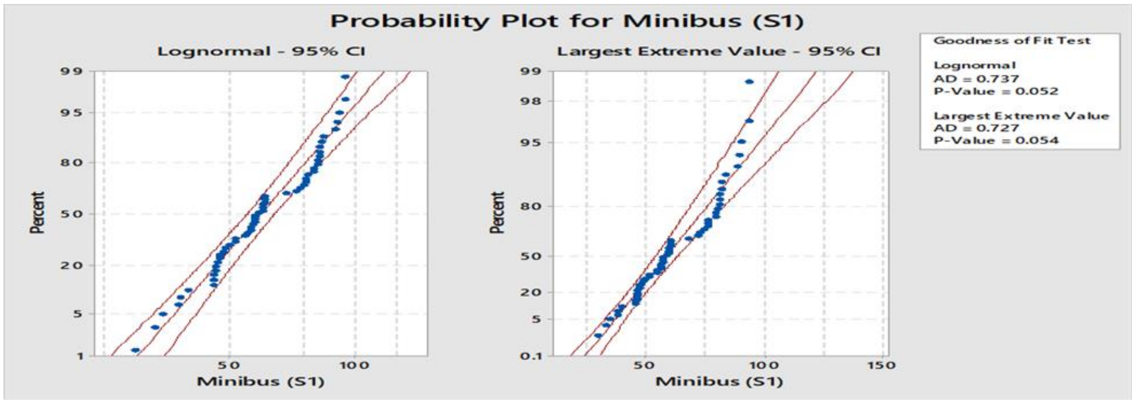


Figure 4.5: Probability Plot of minibus (segment1)

- Distribution fitting for travel time data of Cars for segment 2

Normality test result, probability plot and histogram of travel time data for segment 2 car indicate no significant departure from normality. The data points are gathered around the straight line. The Anderson-Darling test result implies that the data follow a normal distribution. The p-value from Anderson-Darling test result is greater than *0.075* which is greater than the significance level ***(0.05)*** and this confirms that the data come from normally distributed population. And which leads to the conclusion to reject the alternate hypothesis which states that the data does not come from a normal distribution.

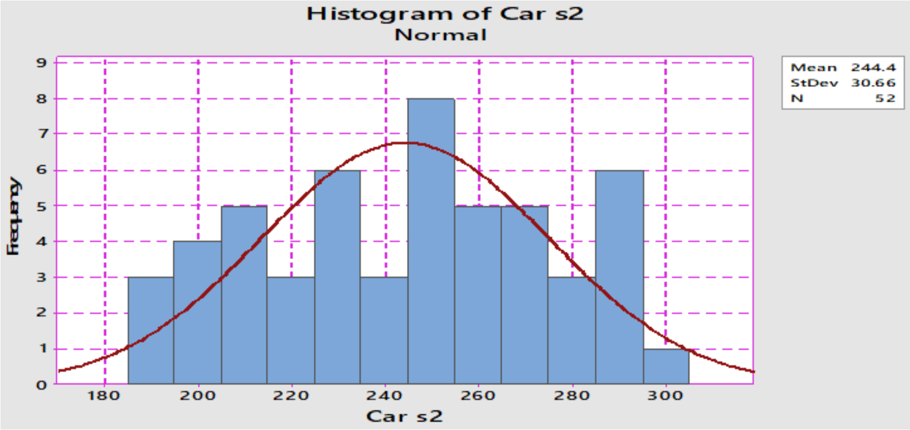


Figure 4.6: Histogram plot of minibus (segment 2)

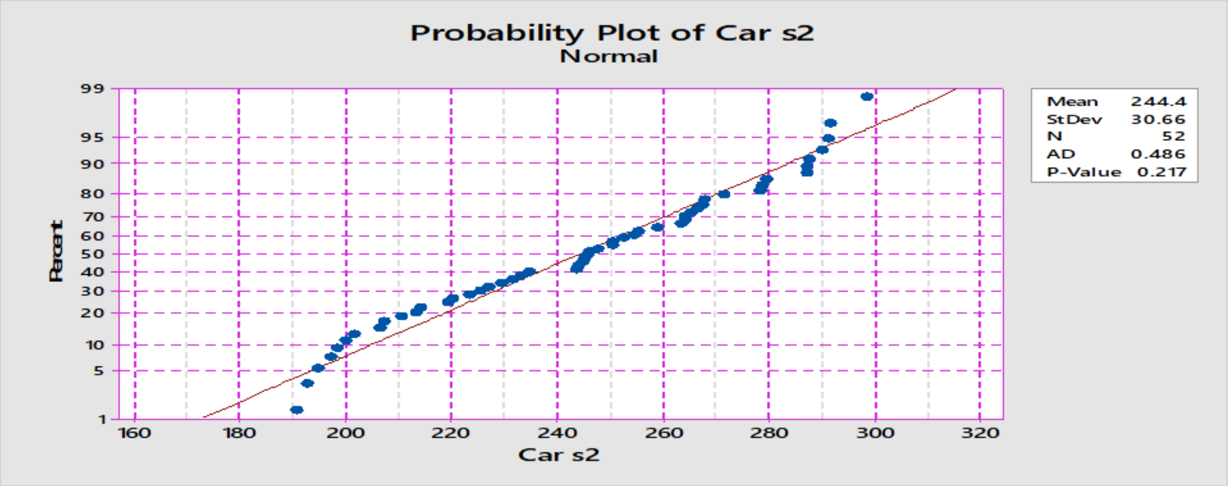


Figure 4.7: Normal probability plot of car (segment 2)

Histogram plot with a fit line signify that the data is skewed to the right and the p-value from normality test is ***<0.005*** which is significantly lesser than the significance level ***(0.05)***. The probability plot shows the data points are not close enough to the theoretical cumulative distribution function. The above points confirm that the data does not come from a normally distributed population. Since the data does not come from a normal distribution, we must find which probability distribution the data belongs to. All possible distributions were tested to fit travel time data for segment of minibus. Lognormal, Gamma, Log-logistic, and 3 Parameter Weibull probability distributions.

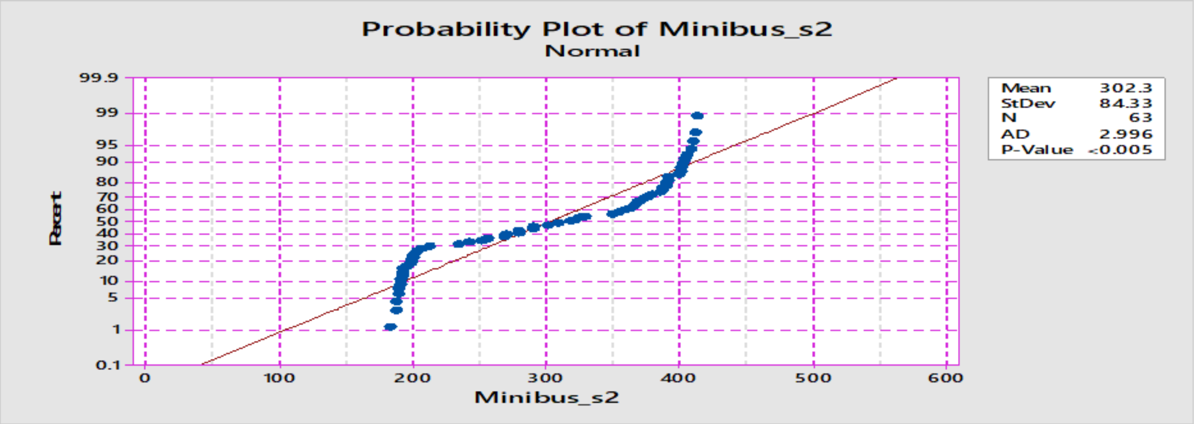


Figure 4.8: Normal probability plot of minibus (segment 2)

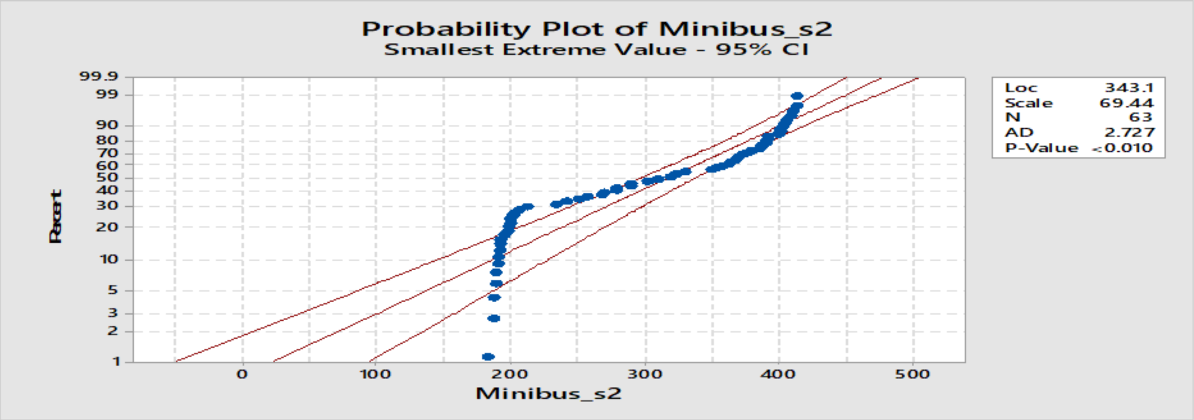


Figure 4.9: Probability Plot of minibus (segment2)

Among all possible probability distributions tried to fit the data only the smallest Extreme Value was found to fit the data. The table below shows estimated distribution parameter and Anderson-Darling test result.

Table 4.2: Estimated Parameters and AD Test Result

Distribution	Estimated parameters		Anderson –Darling test result	
	Location	Scale	AD	P
smallest extreme value	343.1	69.44	2.727	<0.01

- Distribution fitting for travel time data of car and minibus for segment 3

Normality test result, probability plot and histogram of travel time data for segment 3 car and minibus indicates no significant departure from normality. The data points are gathered around the straight line. The Anderson-Darling test result implies that the data follow a normal distribution. The p-value from Anderson-Darling test result is greater than ***0.005*** which is greater than the significance level ***(0.05)*** and this confirms that the data come from normally distributed population.

And which leads to the conclusion to reject the alternate hypothesis which states that the data does not come from a normal distribution.

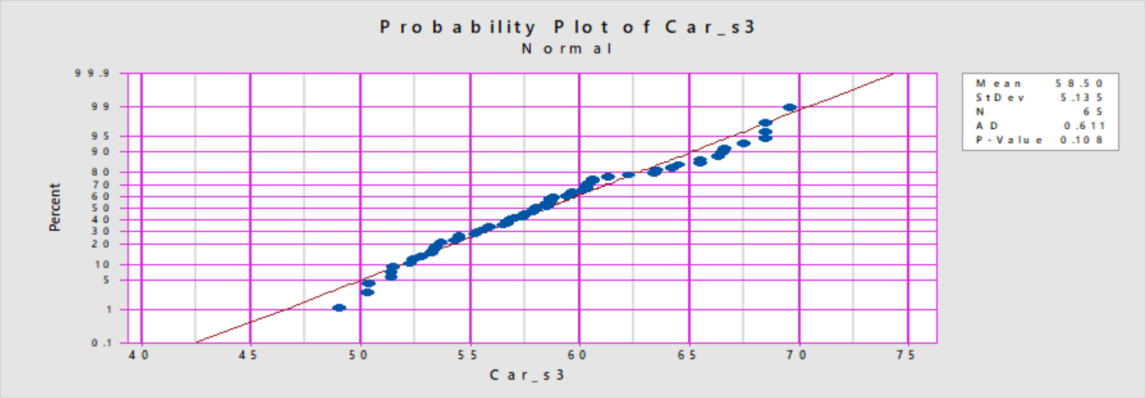


Figure 4.10: Normal probability plot of cars (segment 3)

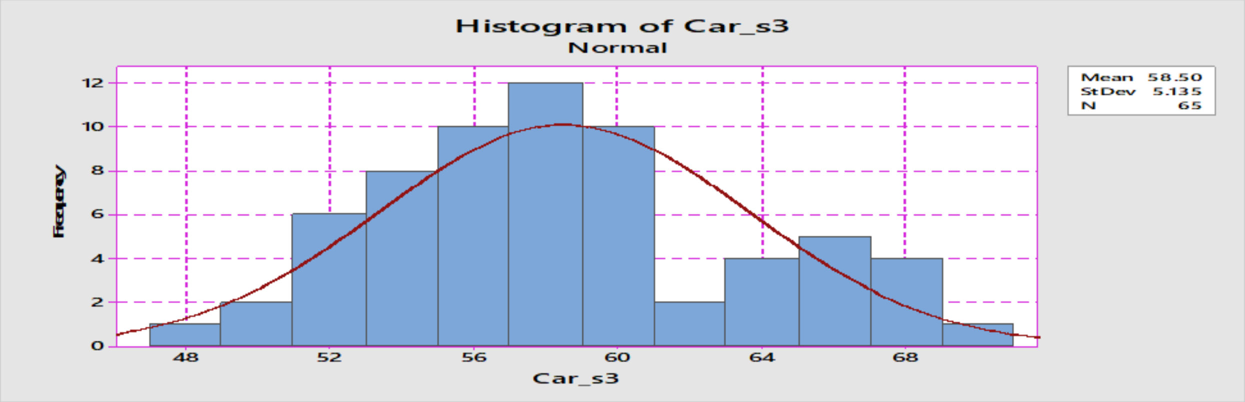


Figure 4.11: Histogram plot of car (segment 3)

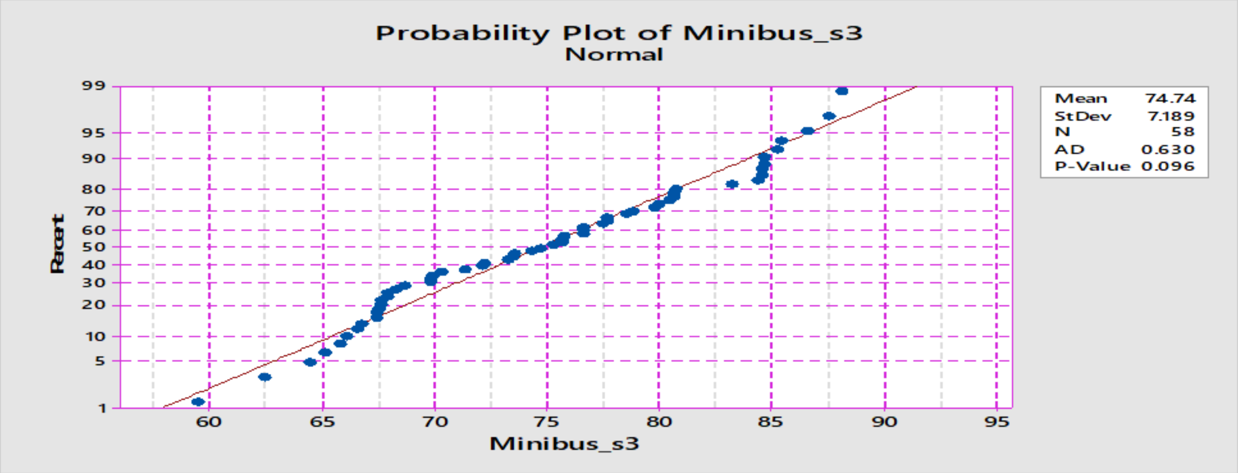


Figure 4.12: Normal probability plot of minibus (segment 3)

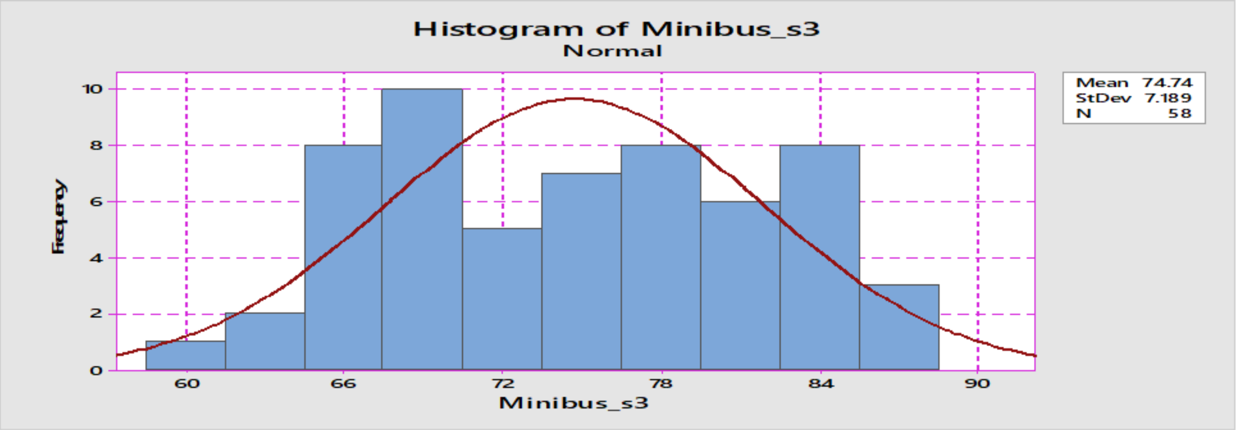


Figure 4.13: Histogram plot of minibus (segment 3)

- Distribution fitting for travel time data of car and minibus for segment 4

The probability plot shows the data points are not close enough to the theoretical cumulative distribution function. The above points confirm that the data does not come from a normally distributed population. Since the data does not come from a normal distribution, we must find which probability distribution the data belongs to. Among all possible probability distributions tried to fit the data only the smallest Extreme Value was found to fit the data.

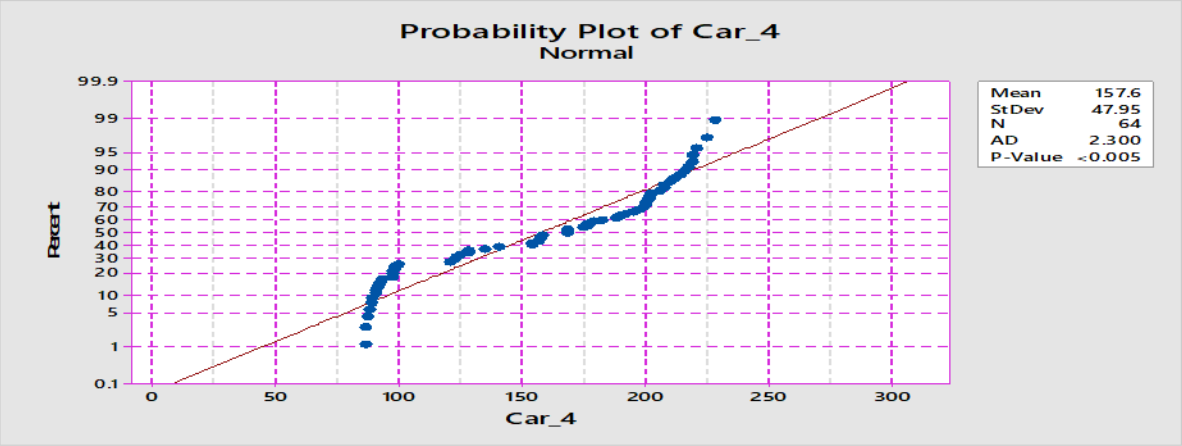


Figure 4.14: Normal probability plot of car (segment 4)

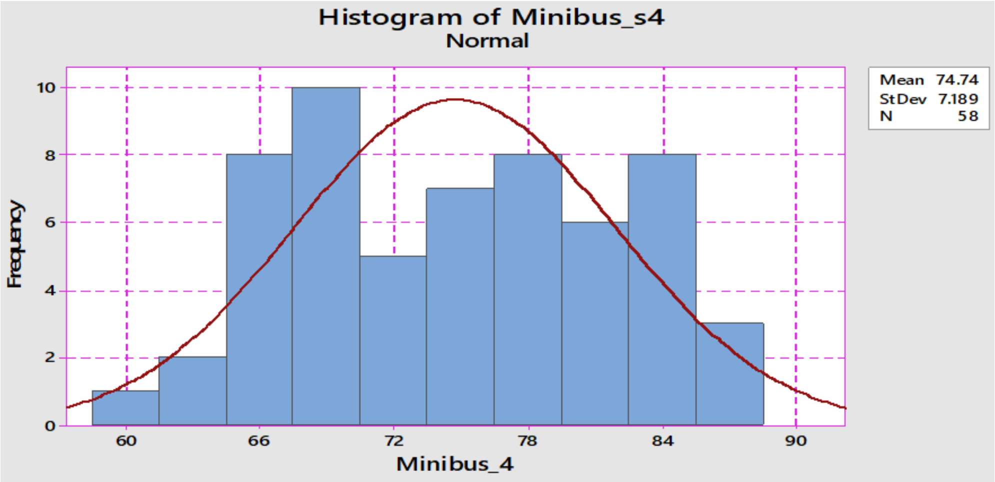


Figure 4.15: Histogram plot of minibus (segment 4)

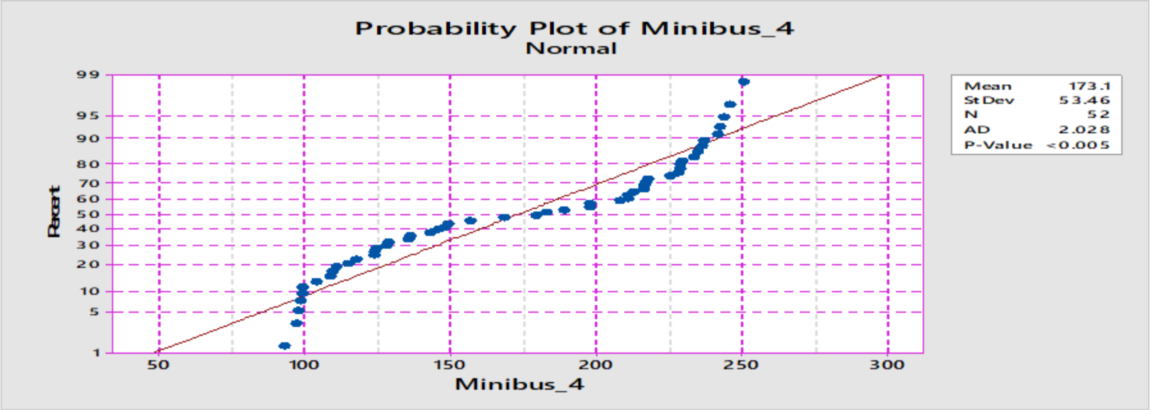


Figure 4.16: Normal probability plot of minibus (segment 4)

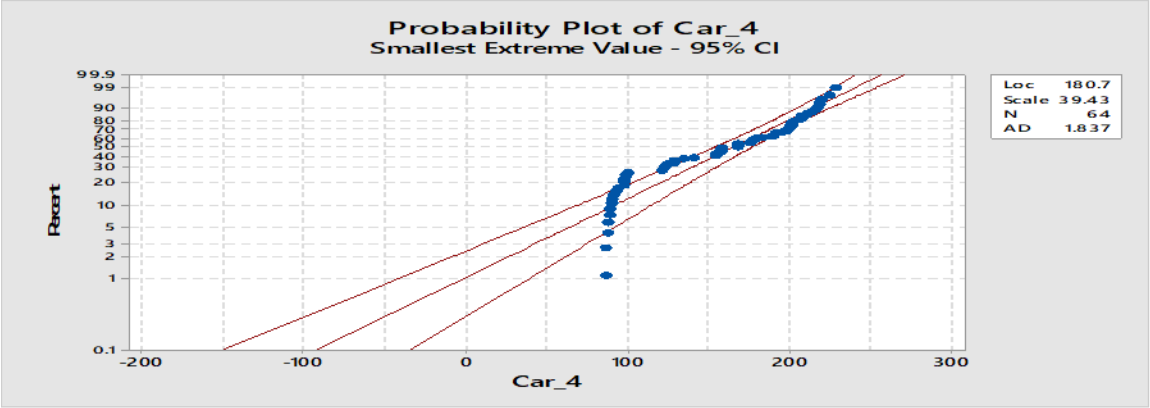


Figure 4.17: Probability Plot of car (segment4)

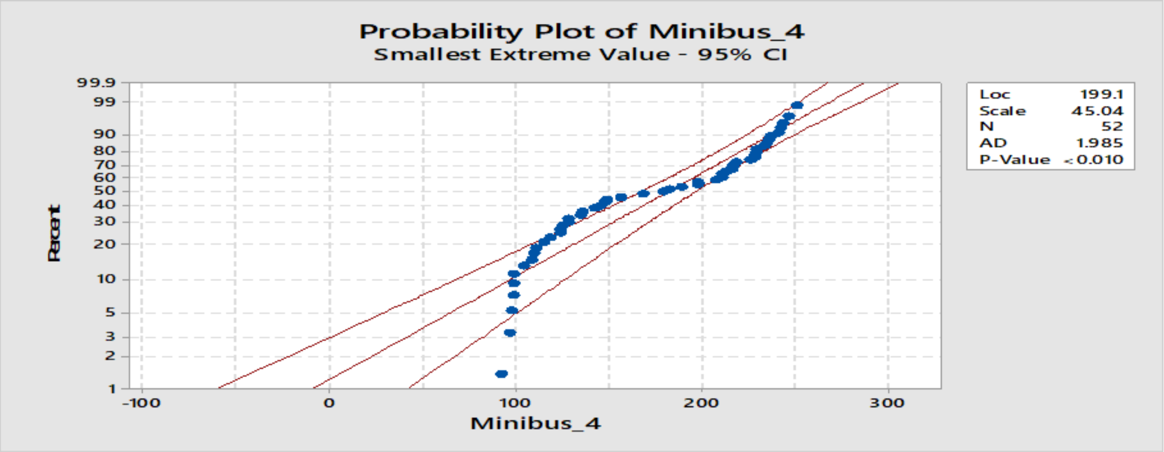


Figure 4.18: Probability Plot of minibus (segment4)

Table 4.3: Estimated Parameters and AD Test Result

Distribution	Estimated parameters		Anderson –Darling test result	
	Location	Scale	AD	P
smallest extreme value	180.7	39.43	1.837	<0.01

Table 4.4: Estimated Parameters and AD Test Result

Distribution	Estimated parameters		Anderson –Darling test result	
	Location	Scale	AD	P
smallest extreme value	199.1	45.04	1.985	<0.01

❖ Distribution fitting for travel time data of car for segment 5

The probability plot shows the data points are not close enough to the theoretical cumulative distribution function. The above points confirm that the data does not come from a normally distributed population. Since the data does not come from a normal distribution, we must find which probability distribution the data belongs to. All possible distributions were tested to fit travel time data for segment of minibus. Lognormal, Gamma, Log-logistic, and 3 Parameter Weibull probability distributions were found to fit the data but we have to find out which probability distribution fit the data best.

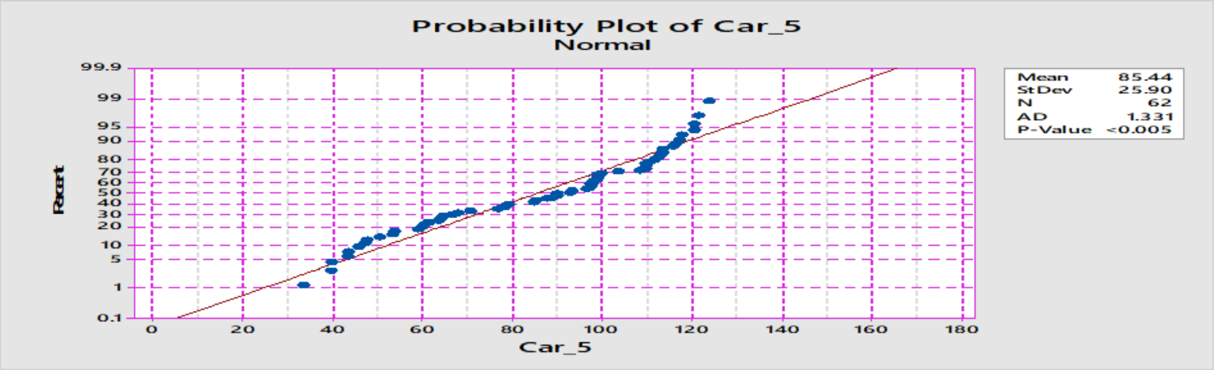


Figure 4.19: Normal probability plot of car (segment5)

Among all possible probability distributions tried to fit the data only the smallest Extreme Value was found to fit the data.

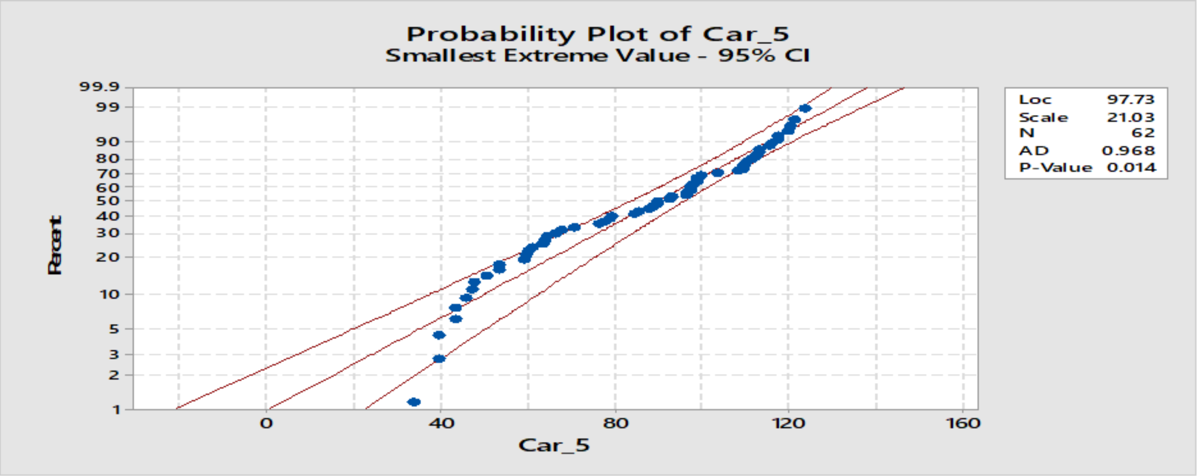


Figure 4.20: Probability Plot of cars (segment5)

Table 4.5: Estimated Parameters and AD Test Result

Distribution	Estimated parameters		Anderson –Darling test result	
	Location	Scale	AD	P
smallest extreme value	97.73	21.03	0.968	0.014

- ❖ Distribution fitting for travel time data of minibus for segment 5

Normality test result, probability plot and histogram of travel time data for segment5 car indicate no significant departure from normality. The data points are gathered around the straight line. The Anderson-Darling test result implies that the data follow a normal distribution. The p-value from Anderson-Darling test result is greater than ***0.005*** which is greater than the significance level ***(0.05)*** and this confirms that the data come from normally distributed population. And which leads to the conclusion to reject the alternate hypothesis which states that the data does not come from a normal distribution.

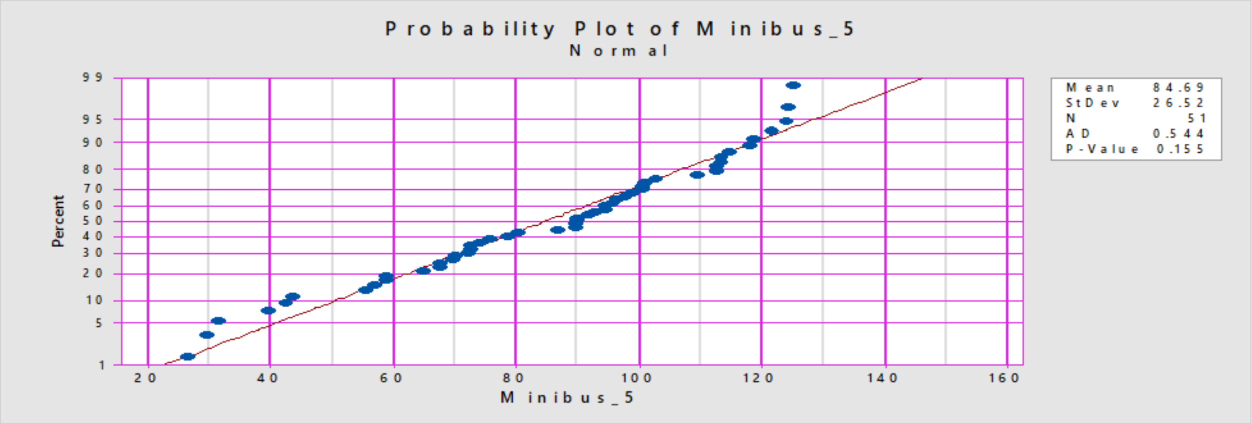


Figure 4.21: Normal probability plot of minibus (segment 5)

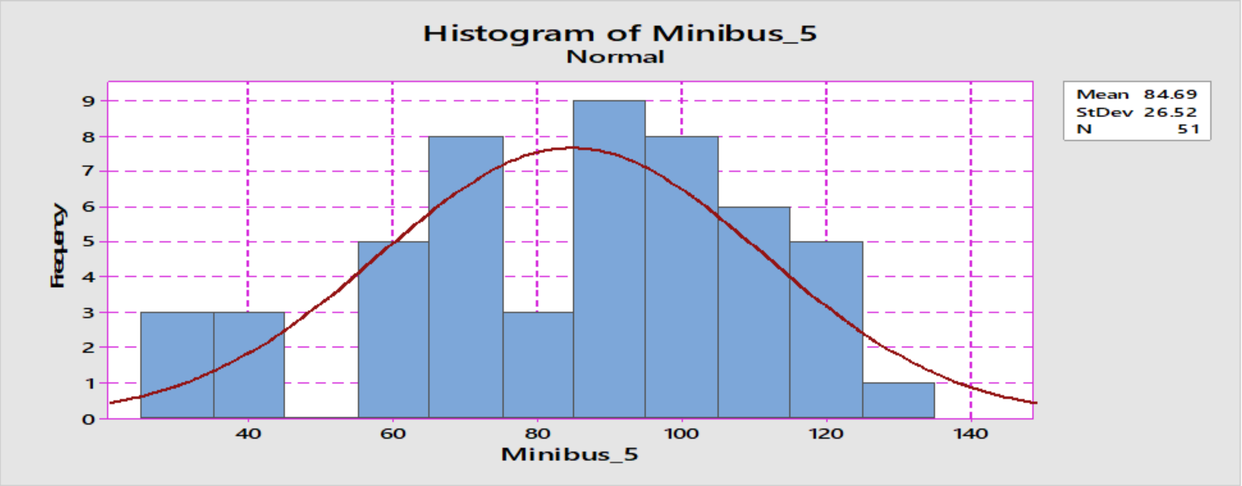


Figure 4.22: Histogram plot of minibus (segment 5)

- Distribution fitting for travel time data of car for segment 6

Normality test result, probability plot and histogram of travel time data for segment 6 car indicate no significant departure from normality. The data points are gathered around the straight line. The Anderson-Darling test result implies that the data follow a normal distribution. The p-value from Anderson-Darling test result is greater than ***0.005*** which is greater than the significance level ***(0.05)*** and this confirms that the data come from normally distributed population. And which leads to the conclusion to reject the alternate hypothesis which states that the data does not come from a normal distribution.

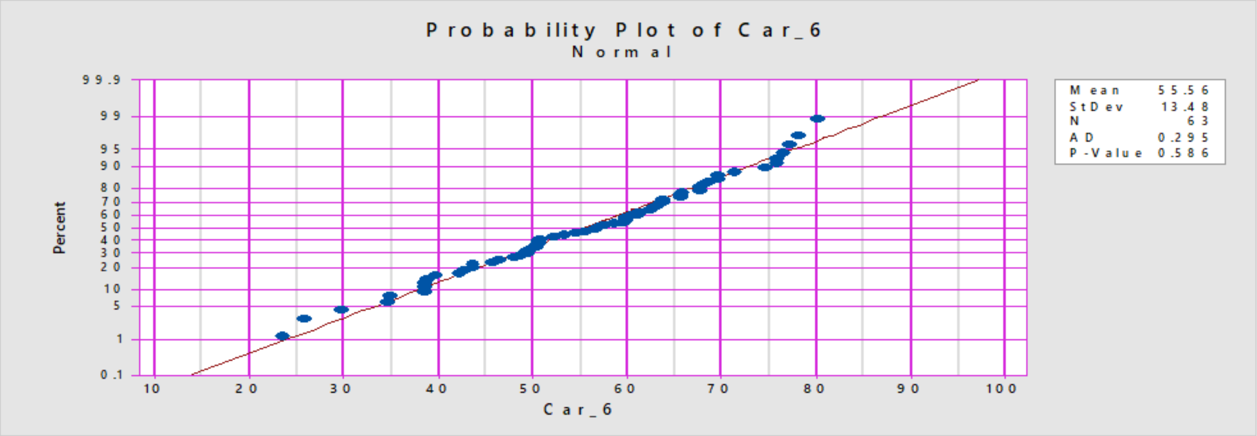


Figure 4.23: Normal probability plot of car (segment 6)

- Distribution fitting for travel time data of minibus for segment 6

Histogram plot with a fit line signify that the data is skewed to the left and the p-value from normality test is ***<0.005*** which is significantly lesser than the significance level ***(0.05)***. The probability plot shows the data points are not close enough to the theoretical cumulative distribution function. The above points confirm that the data does not come from a normally distributed population.

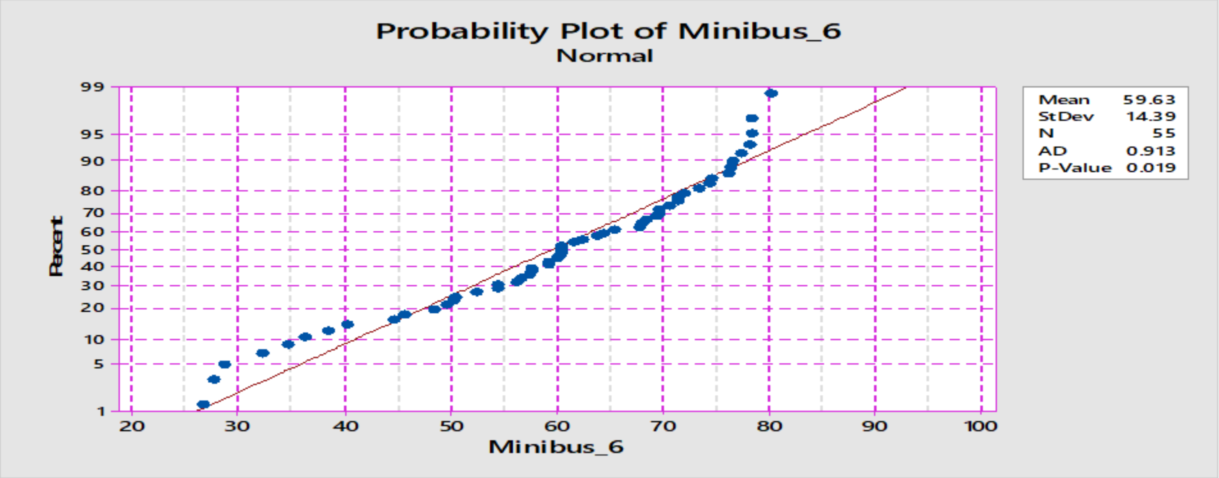


Figure 4.24: Normal probability plot of minibus (segment 6)

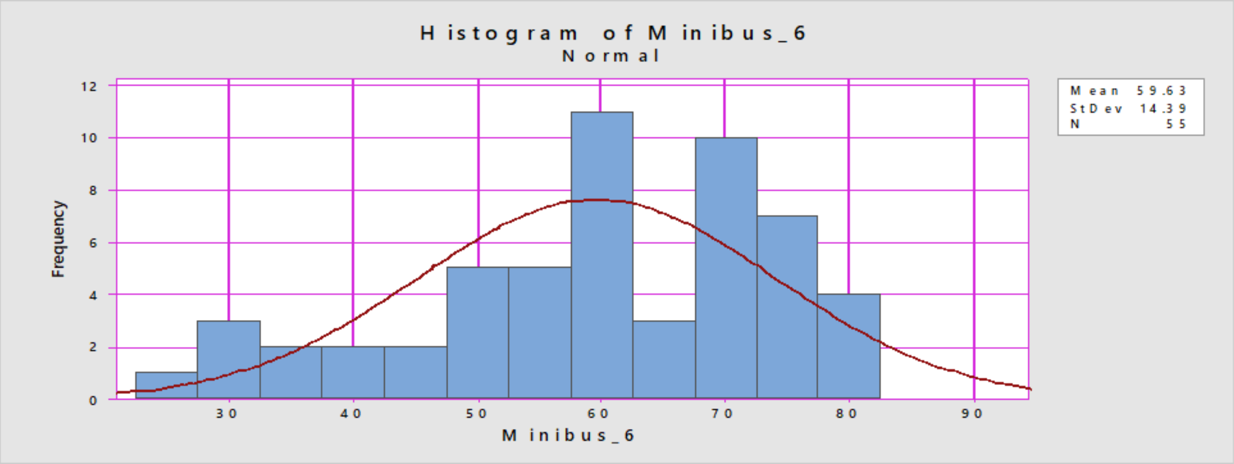


Figure 4.25: Histogram plot of minibus (segment 6)

Among the tested distributions Weibull, 3-Parameter Weibull and smallest extreme value probability distributions significantly fit the data. The Largest Extreme Value probability distributions have smaller AD value and larger p-value than the other distributions which proves that it's a better fit for the data. The following table shows the estimated parameters along with the Anderson-Darling test result. Smallest extreme value distribution probability plot has a less number of data points which lie of the ***95%*** confidence interval limit and the points are also closer to the straight line than the other probability plot. So the smallest extreme value probability distribution is a better fit for the travel time data of minibus for segment 6.

Table 4.6: Estimated Parameters and AD Test Result

Distribution	Estimated parameters		Anderson –Darling test result	
	Location	Scale	AD	P
Weibull		65.01	0.703	0.064
Smallest extreme value	66.25	11.25	0.379	>0. 250
3-parametric Weibull		13.69	0.379	0.288

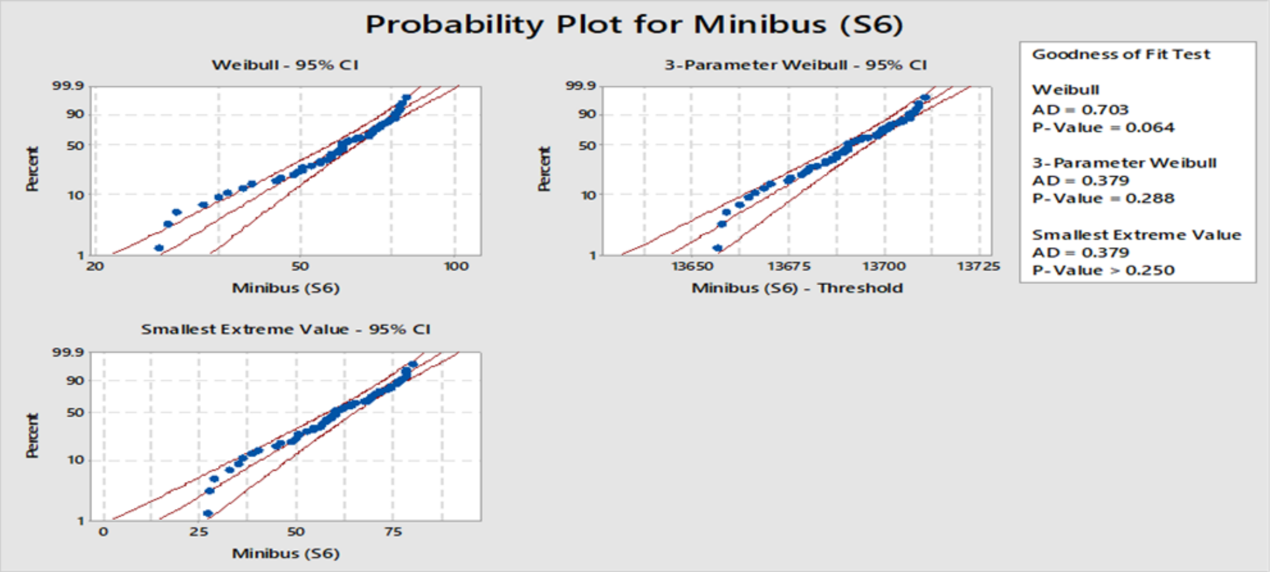


Figure 4.26: Probability Plot for minibus (segment 6)

- Distribution fitting for travel time data of car for segment 7

Histogram plot with a fit line signify that the data is skewed to the right and the p-value from normality test is ***<0.005*** which is significantly lesser than the significance level ***(0.05)***. The probability plot shows the data points are not close enough to the theoretical cumulative distribution function. The above points confirm that the data does not come from a normally distributed population.

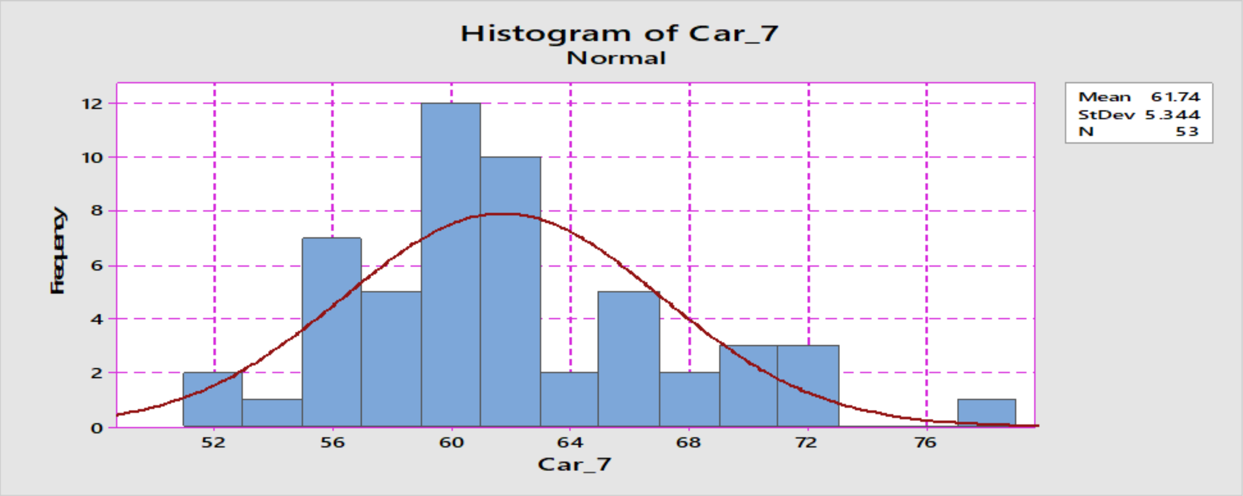


Figure 4.27: Histogram plot of car (segment 7)

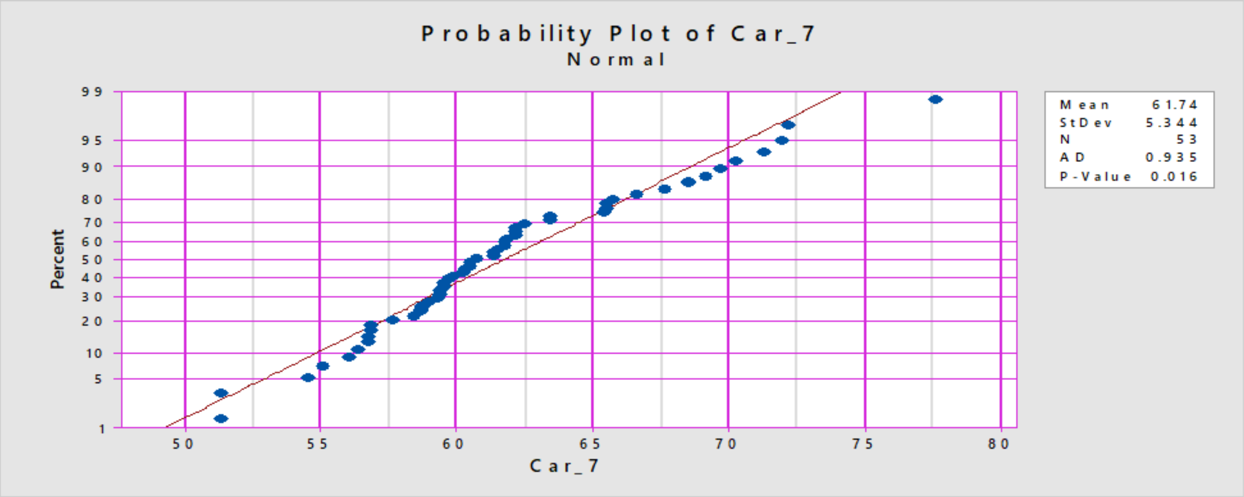


Figure 4.28: Normal probability plot of car (segment 7)

Among the tested distributions Lognormal, 3-Parameter Weibull, Largest Extreme Value and log logistic probability distributions significantly fit the data. The Largest Extreme Value probability distributions have smaller AD value and larger p-value than the other distributions which proves that it's a better fit for the data. The following table shows the estimated parameters along with the Anderson-Darling test result.

Table 4.7: Estimated Parameters and AD Test Result

Distribution	Estimated parameters		Anderson –Darling test result	
	Location	Scale	AD	P
Log normal		65.01	0.703	0.064
Largest extreme value	66.25	11.25	0.379	>0. 250
3-parametric Weibull		13.69	0.379	0.288
Log logistic	4.115	0.0469	0.557	0.104

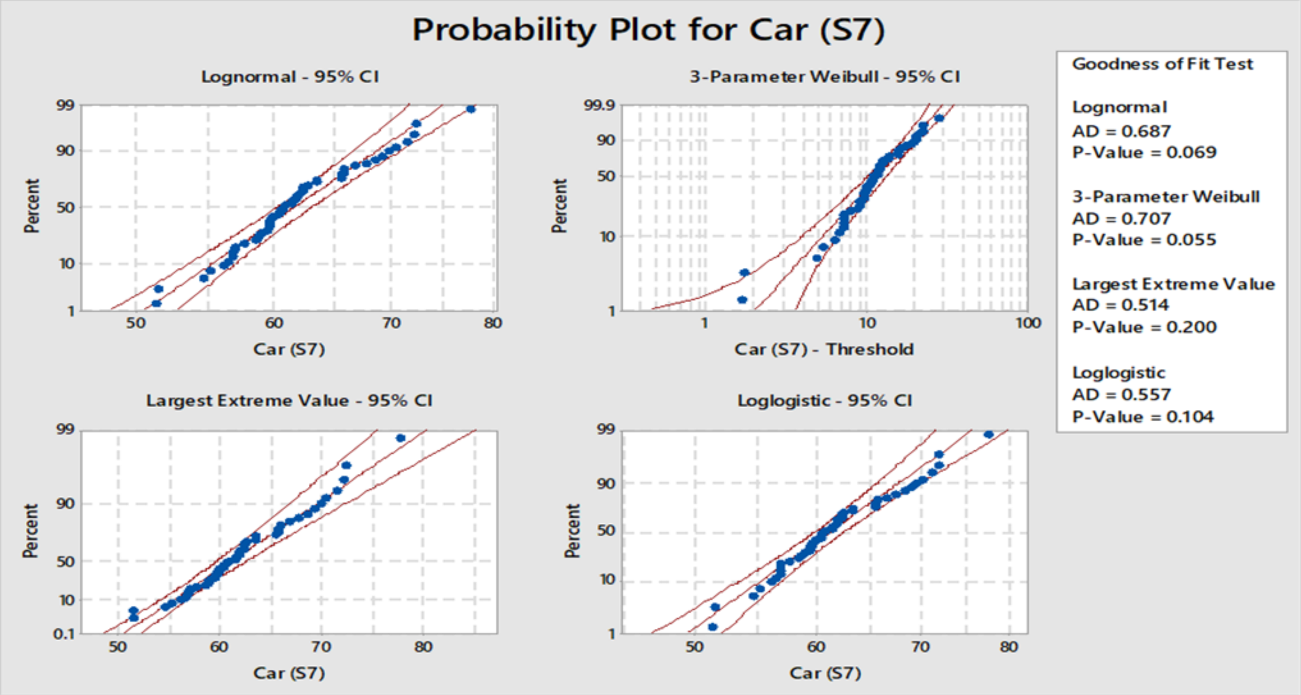


Figure 4.29: Probability Plot for car (segment 7)

- ❖ Distribution fitting for travel time data of minibus for segment 7

As it can be seen from the figure the shape of the curve on the histogram approximates to a bell shape with a peak value at the middle. This shows that the data is from a normal distribution. The normal probability plot graph shows the pattern of the data approximately following the straight line strengthening the conclusion that the data is from normal distribution. The p-value from Anderson-Darling test result is greater than ***0.005*** which is greater than the significance level ***(0.05)*** and this confirms that the data come from normally distributed population.

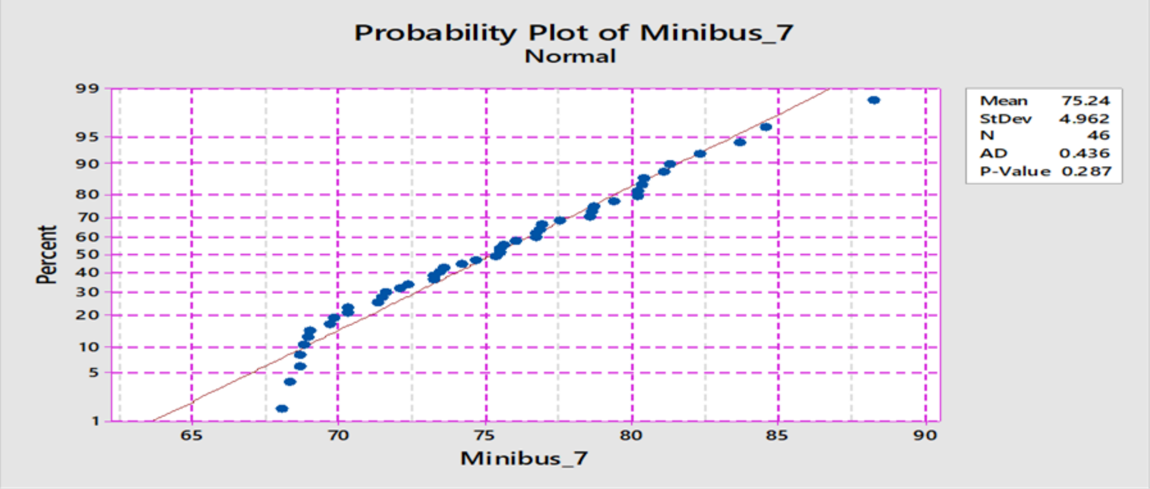


Figure 4.30: Normal probability plot of minibus (segment 7)

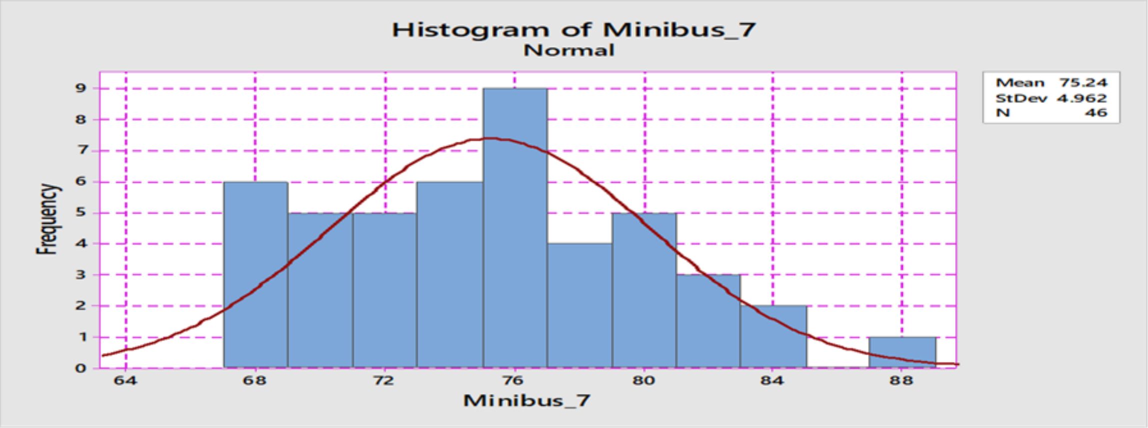


Figure 4.31: Histogram plot of minibus (segment 7)

4.3. Travel time reliability measurement data analysis

Reliability of travel time is an important topic for increasing safety, quality life for road users to produce less delay for their trips. Also, it's a good indicator for improving the overall system operations and management. Reliability measurement can be calculated as

Buffer Time =95th percentile travel time (sec)-Average travel time (sec)

Buffer time index= $\frac{\text{95th percentile travel time} - \text{Avrtagre travel time}}{\text{Average travel time}}$

Table 4.8: Travel Time reliability indices for each segment

Seg.No	Length(m)	ATT(sec)	BT(sec)	BTI(sec)	95th percentile travel time
1	378	60.08	31.43	0.52	91.52
2	705	264.58	138.39	0.523	402.98
3	372	71.08	19.23	0.27	90.3
4	422	171.34	74.63	0.43	245.38
5	247	88.82	42.49	0.49	131.32

6	245	58.33	24.59	0.42	82.92
7	596	78.98	15.39	0.195	94.37

Figure 4.31 presented the reliability measurement for all segment in terms of buffer time index for segment 1, 2, 3,4,5,6 & 7 respectively. Increasing the buffer time index results in worse reliability conditions. Segment (2) produced a buffer index of about 52.3% and 52 %, 49%, 43%, 42%, 27%f and 19.5% for segment 1, 5,4,6,3 and 7 respectively. These illustrated that the reliability get worst for segment2, 1, 5, 4, 6, 3, and 7 respectively. Also, buffer time of (31.43, 138.39, 19.23, 74.63, 42.49, 24.59 & 15.39) sec. for segment (1, 2,3,4,5,6 and 7) respectively is obtained based on average travel time for each segment, that mean additional 31.43,138.39,19.23,74.63,42.49,24.59 & 15.39 sec. Also, Figure (4.18) show the 95% percentile travel time for observed segments which presents the extra delay that perceived on each segment (91.52,402.98,90.3,245.38,131.32,82.92, and 94.37) sec for segment 1,2,3,4,5,6 and 7 respectively. Higher value for 95% travel time is obtained for segment 2.

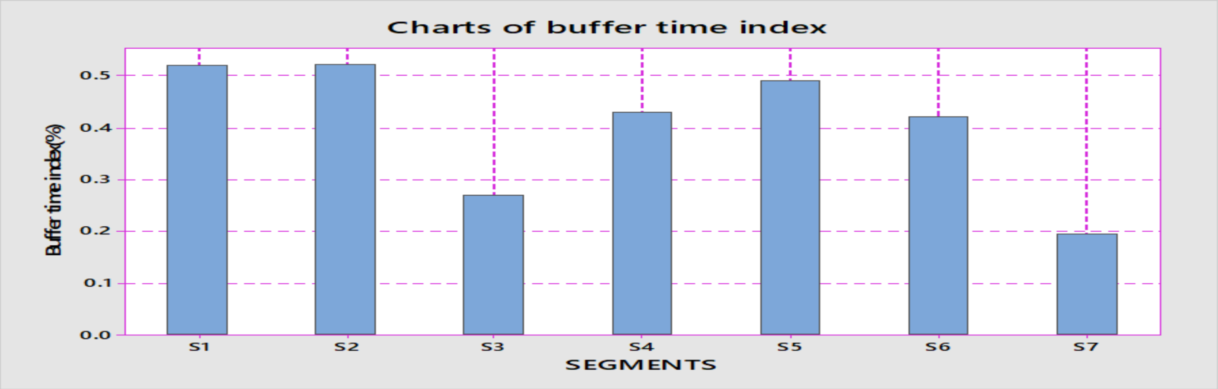


Figure 4.32: Buffer Time Index for segment 1, 2, 3,4,5,6 & 7.

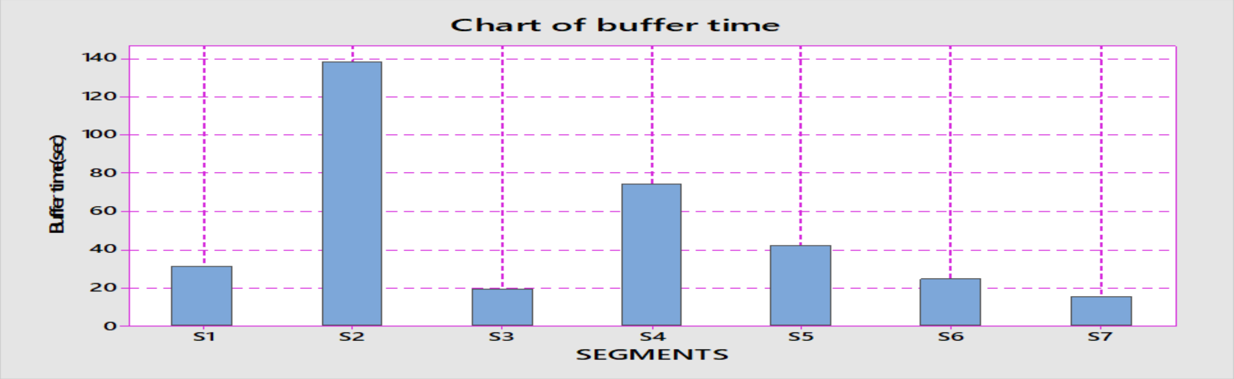


Figure 4.33: Buffer Time for segment 1, 2, 3,4,5,6 & 7.

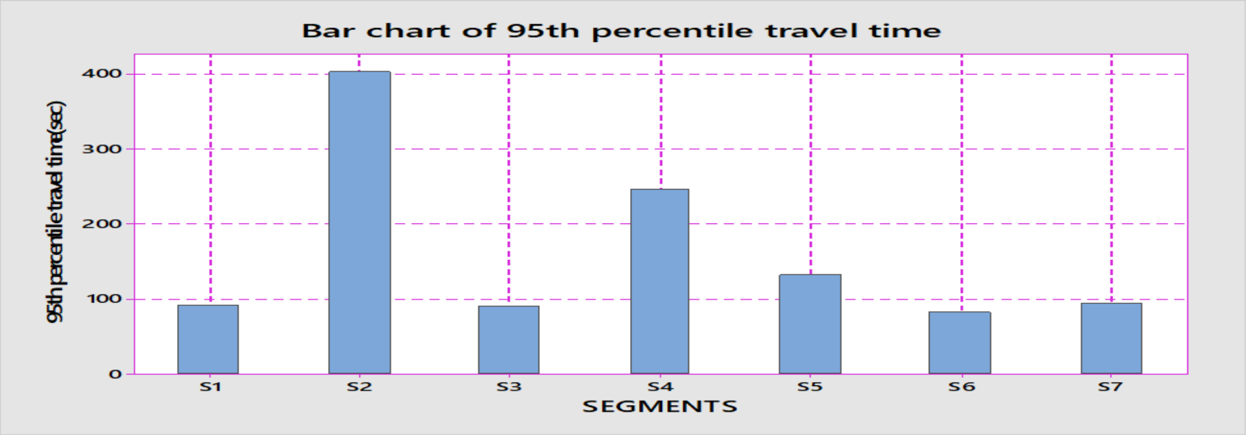


Figure 4.34: 95% Percentile Travel Time Results for segment 1, 2, 3,4,5,6 & 7

4.4. Travel time reliability Regression Modeling

Multiple linear regression technique is used to develop a mathematical model for travel time data. From previous study 15th percentile travel time considered as free flow travel time and when a segment said to be reliable all travel time data should be less than or equal to free flow travel times **(Fils, 2012).**.But in this research above 97.7% of travel time data was above free flow travel time .so the researcher wants to develop reliability models. The total travel time data extracted for the regression analysis was 1897.

4.4.1. Testing multiple linear regression assumptions

The major part of a regression analysis is checking the assumptions. Before conducting any regression analysis, the assumptions should be checked. The following multiple linear regression assumptions are checked for violation before performing the regression analysis.

Assumption 1: - relationship between the independent and dependent variables to be linear. There should be significant linear relationship between the dependent variable and each of the independent variables. Scatter plot of the each independent variable versus the dependent variable is used to check the assumption. As shown below in the figure scatter plot of dependent variables versus each of the independent variables is presented showing that there is no significant evidence for the existence of nonlinear relationship.

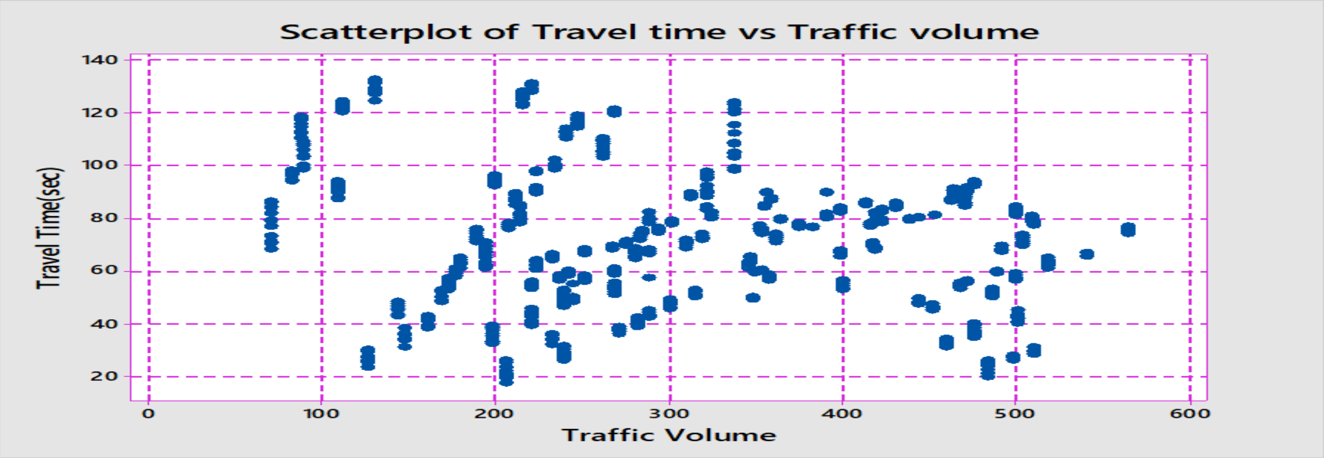


Figure 4.35: Scatter plot of travel time vs. Traffic volume

Other scatter plot of travel time vs. segment grade, average segment width, % of heavy vehicles, number of intersection legs, signal cycle length, segment length and average speed see appendix A.

Assumption 2:- Multicollinearity (linear regression assumes that there is little or no Multicollinearity in the data) Multicollinearity occurs when the independent variables are too highly correlated with each other. We can check Multicollinearity two ways: correlation coefficients and variance inflation factor (VIF) values. To check it using correlation coefficients, simply throw all your predictor variables into a correlation matrix and look for coefficients with magnitudes of .80 or higher. If your predictors are Multicollinearity, they will be strongly correlated. However, an easier way to check is using VIF values, which we will show how to generate below10.00. As we can see in the table below all the coefficient of correlation matrix ,0.8 and Variance inflation factors corresponding to each explanatory variables is much lesser than 10 (<<10) so it can be concluded that there is no a Multicollinearity problem.***(See table 4.9-4.10)***

Table 4.9: Predictor Variables Variance Inflation Factor

Variables	Traffic volume	Segment grade	Average segment width	% of heavy vehicles	Number of intersection legs	Segment length	Speed
VIF	1.71	2.45	3.16	1.24	2.6	2.46	2.83

Table 4.10: Correlation Matrix between Independent Variables

	Reliable travel time	Traffic volume	Segment grade	Average segment width	% of heavey vehicles	Number of intersectionlegs	Signal cycle length	Segment length	Average speed
Reliable travel time	1								
Traffic volume	-0.143	1							
Segment grade	-0.176	0.231	1						
Average segment width	0.033	0.151	-0.558	1					
% of heavey vehicles	0.08	0.139	0.129	-0.144	1				
Number of intersection legs	0.018	-0.177	-0.223	0.493	0.157	1			
Signal cycle length	-0.061	0.194	0.637	-0.815	0.24	0.544	1		
Segment length	0.172	0.228	-0.41	0.169	-0.102	-0.42	0.192	1	
Average speed	-0.278	0.447	-0.068	0.232	0.003	-0.286	0.225	0.621	1

Assumption 3: Homoscedasticity (variables along the line of best fit have to remain similar as moved along the line). To test the homoscedasticity (equality of variances) of our data, scatter plot of standardized residuals versus standardized fitted values is analyzed. As shown in the plot below, the spread of the residuals around the line seem to have approximately similar pattern.

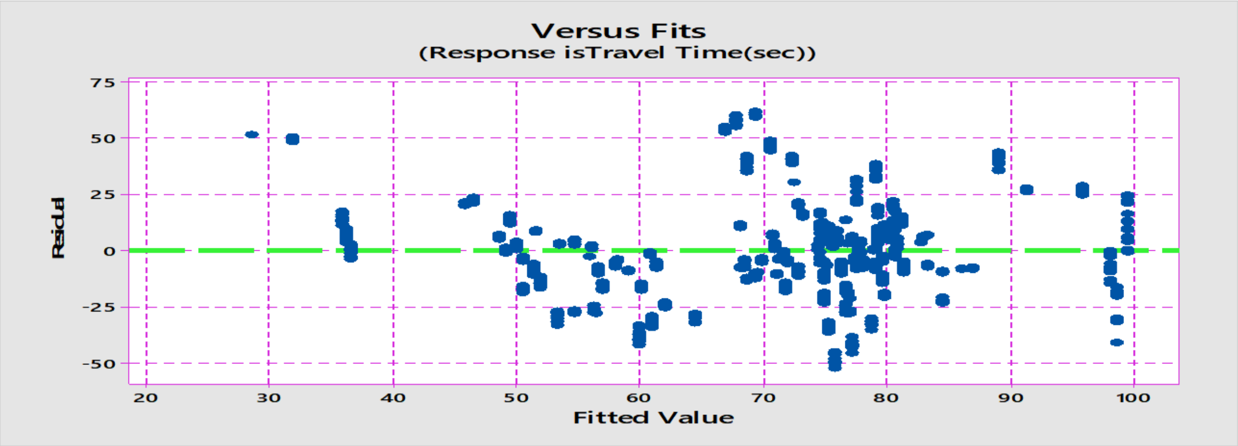


Figure 4.36: Scatter plot of standardized residuals vs. standardized fitted value

Assumption 4:- There should be **no significant outliers**. An outlier is an observed data point that has a dependent variable value that is very different to the value predicted by the regression equation. As such, an outlier will be a point on a scatterplot that is (vertically) far away from the regression line indicating that it has a large residual.

Assumption 5:- Finally, we need to check that the **residuals (errors)** of the regression line are

approximately normally distributed (we explain these terms in our enhanced linear regression guide). Two common methods to check this assumption include using either a histogram (with a superimposed normal curve) or a Normal P-P Plot Residuals are normally distributed. Figure-4.29 shows the histogram of the data. As it can be seen from the figure the shape of the curve on the histogram approximates to a bell shape with a peak value at the middle. This shows that the regression assumption is satisfied.

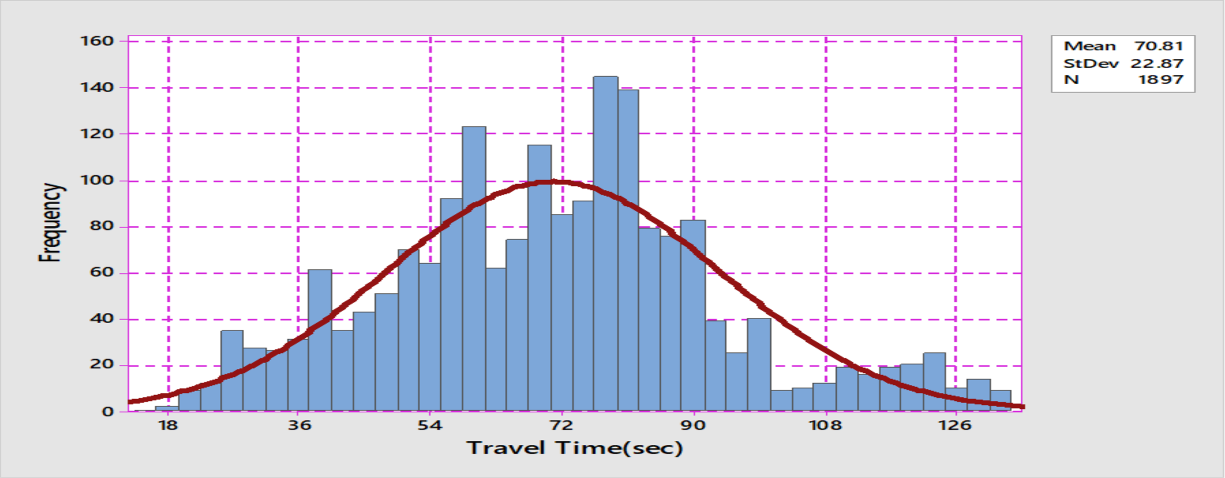


Figure 4.37: Histogram of Residuals with a Normal Curve

4.5. Regression Analysis

Stepwise multiple regressions were performed to model travel time of vehicle on urban segment with the following predictor variables. The predictor variables consist only continuous variables.

Dependent Variable

- Vehicle travel time

Predictor variables

- Traffic volume
- Segment grade
- Average segment width
- % of heavy vehicles
- Number of intersection legs
- Signal cycle length
- Segment length
- Average speed

Forward stepwise selection of terms is used for all candidate predictor variables. Only one variable is added at each step of the regression. As shown in the table below p-values for each added variable is less than the significance level which confirms that all predictor variables significantly predict travel time of vehicles and there is an increment in the adjusted R squared value at each step of the regression.

Table 4.11: Stepwise regression output

	Step 1		Step 2		Step 3		Step 4		Step 5		Step 6		Step 7		Step 8	
	Coef	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*
Constant	*90.94*		*79.6*		*71.63*		*55.73*		*47.06*		*45.18*		*93*		*103.40*	
Average speed	-3.415	*0.00*	7.697	0.00	-7.893	0.00	-8.2	0.00	-8.847	0.00	-8.614	0.00	-8.9	0.00	9.147	0.00
Segment length			0.096	0.00	0.101	0.00	0.1002	0.00	0.11269	0.00	0.1162	0.00	0.107	0.00	-0.346	0.00
%of heavy vehicles					1.622	0.00	1.801	0.00	1.818	0.00	1.949	0.00	1.754	0.00	-1.779	0.00
Average segment width							1.463	0.00	2.372	0.00	2.76	0.00	4.006	0.00	0.522	0.00
Segment Grade									1.6	0.00	2.151	0.00	1.427	0.00	-9.363	0.00
Traffic Volume											-0.0152	0.00	-0.01598	0.00	-0.01803	0.00
Signal cycle length													0.1519	0.00	-0.9725	0.00
Number of intersection legs															-7.29	0.00
R^2	0.487		0.567		0.674		0.685		0.724		0.767		0.803		0.846	
R^2_{Adj}	0.488		0.568		0.673		0.686		0.725		0.768		0.802		0.843	

Table 4.12: Final regression result

Model(Reliable travel time)	Coefficients				95 % Confidence interval	
	Beta	Stad.Error	t	Sig(p-value)	Lower bound	Upper bound
Constant	103.4	-167.26	21.03	0.00	97.58	111.9
Average speed	9.147	0.345	10.34	0.00	5.462	13.56
Segment length	-0.346	0.102	16.25	0.00	-0.501	-0.461
% of heavey vehicles	-1.779	0.023	2.45	0.00	-3.458	-2.569
Signal cycle length	-0.9725	0.48	13.45	0.00	-2.456	-1.698
Average segment width	0.522	1.67	-28.36	0.00	0.256	0.963
Traffic volume	-0.01803	0.04	36.45	0.00	-0.345	-0.136
Segment grade	-9.363	8.246	-47.36	0.00	-14.36	-11.63
Number of intersection legs	-7.29	14.26	25.46	0.00	-12.35	-9.26

Final regression equation written as follows

$TTR_{model}=103.4-0.01803TV-9.363SG+0.522ASW-1.779\%HV-7.29NIL-0.9725SCL-0.346SL+9.147AS$

Where: TTR=Travel time reliability, TV=Traffic volume, SG=Segment grade (%), ASW=average segment width (m), %HV=% of heavy vehicles, NIL=Number of intersection legs, SCL=signal cycle length (sec), SL=Segment length and AS=Average speed

5.7. Interactions in Regression analysis

Adding interaction terms to the regression model can greatly expand understanding of the relationship among variables in the model. Interaction regression analysis developed to predict the combined or synergistic effects of two independent variables.in this research the researcher wants to check all independent interaction effects in the regression model. Forward step wise multiple regression was performed with the following explanatory variables.

Dependent Variable

- TT (Travel time)

Predictor variables

- Traffic volume
- Segment grade
- Average segment width
- % of heavy vehicles
- Number of intersection legs
- Signal cycle length
- Segment length
- Average speed
- Traffic volume-Segment grade
- Traffic volume-% of heavy vehicles
- Traffic volume-Number of intersection legs
- Traffic volume-Signal cycle length
- Traffic volume –Segment length
- Segment grade-% of heavy vehicles
- Segment grade-Number of intersection legs
- Segment grade –Signal cycle length
- Segment grade-Segment length

- Segment grade-Speed
- % of heavy vehicles-Number of intersection legs
- % of heavy vehicles –Signal cycle length
- % of heavy vehicles-Segment length
- Number of intersection legs-Signal cycle length
- Number of intersection legs-Segment length
- Signal cycle length-Segment length

Table 4.13: Stepwise interaction regression analysis output

	Step 1		Step 2		Step 3		Step 4		Step 5		Step 6		Step 7		Step 8		Step 9	
	Coef	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*
Constant	*88.37*		*118.3*		*152*		*919*		*868.8*		*1330*		*93*		*917.80*		988.8	
Segment Grade	-13.39	*0.00*	30.71	0.00	89.9	0.00	264.9	0.00	385.3	0.00	438.6	0.00	-8.9	0.00	454.8	0.00	494.7	0.00
% of heavy vehicle	-4.504	0.00	-2.42	0.00	-3.4	0.00	-3.3	0.00	-1.356	0.00	-2.036	0.00	0.11	0.00	4.723	0.00	4.522	0.00
Segment Grade *% of Heavy Vehicle	-3.474	0.00	-3.15	0.00	-3.2	0.00	-2.05	0.00	-1.31	0.00	-1.526	0.00	1.75	0.00	-1.042	0.00	-1.06	0.00
Traffic volume			-0.09	0.00	-0.1	0.00	-1.92	0.00	-1.327	0.00	-1.131	0.00	897	0.00	-0.333	0.00	-2.39	0.00
Traffic volume *Segment Grade			-0.05	0.00	-0.1	0.00	0.005	0.00	0.0234	0.00	0.0173	0.00	453	0.00	-0.043	0.00	-0.14	0.00
Average Segment Width					-0.7	0.00	-75	0.00	-89.42	0.00	-102.2	0.00	-1.8	0.00	-52.53	0.00	-32.5	0.00
Segment Grade *Average Segment Width					-4	0.00	-21.3	0.00	-32.34	0.00	-36.08	0.00	-1.4	0.00	-35.38	0.00	-38.5	0.00
Number of intersection legs							-5.96	0.00	4.74	0.00	-51.31	0.00	-0.4	0.00	-98.65	0.00	-154	0.00
Traffic volume * Number of Intersection Legs							0.522	0.00	0.3966	0.00	0.3393	0.00	-0	0.00	0.436	0.00	0.754	0.00
Segment Length									0.4697	0.00	0.2725	0.00	-46	0.00	0.397	0.00	0.156	0.00
Average Speed									28.91	0.00	-56.85	0.00	-36	0.00	-37.05	0.00	-35.4	0.00
Segment Length *Average Speed									-0.096	0.00	-0.062	0.00	-0.1	0.00	-0.083	0.00	-0.08	0.00
Number of Intersection Legs * Average Speed											18.57	0.00	15.4	0.00	16.55	0.00	15.81	0.00
Traffic volume *Number of Intersection Legs													0.12	0.00	-0.11	0.00	-0.22	0.00
% of Heavy Vehicle *Average Speed															-1.008	0.00	-0.99	0.00
Traffic volume *Signal Cycle Length																	0.012	0.00
R^2	18.9		31.69		38.83		74.37		82.85		85.18		86.67		87.84		88.15	
R^2_{Adj}	18.74		31.47		38.55		74.22		82.71		85.05		86.55		87.73		80.03	

	Step 10		Step 11		Step12		Step 13		Step14		Step 15		Step 16		Step 17		Step 18	
	Coef	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*	*Coef*	*P*
Constant	*942.7*		*397*		*594.3*		*628*		*630.6*		*625*		*626.1*		*581.20*		608	
Segment Grade	547.8	*0.00*	477.1	0.00	497.3	0.00	503.8	0.00	511.4	0.00	507.9	0.00	508.8	0.00	487	0.00	560.2	0.00
% of heavy vehicle	5.363	0.00	5.645	0.00	5.48	0.00	6.07	0.00	6.38	0.00	11.6	0.00	12.53	0.00	107.8	0.00	49.32	0.00
Segment Grade *% of Heavy Vehicle	-1.0058	0.00	-1.0071	0.00	-1.0222	0.00	-0.864	0.00	-1.013	0.00	-1.373	0.00	-1.461	0.00	2.4	0.00		
Traffic volume	-2.297	0.00	-1.895	0.00	-2.016	0.00	-1.937	0.00	-2.098	0.00	-2.3	0.00	-2.364	0.00	-1.612	0.00	-2.081	0.00
Traffic volume *Segment Grade	-0.144	0.00	-0.1454	0.00	-0.1431	0.00	-0.1361	0.00	-0.1438	0.00	-0.147	0.00	-0.148	0.00	-0.1135	0.00	-0.1351	0.00
Average Segment Width	80.5	0.00	301.5	0.00	226.2	0.00	218.5	0.00	224.5	0.00	215.5	0.00	214	0.00	215.6	0.00	214.7	0.00
Segment Grade *Average Segment Width	-42.93	0.00	-37.12	0.00	-38.78	0.00	-39.43	0.00	-40.03	0.00	-39.71	0.00	-39.77	0.00	-38.03	0.00	-39.08	0.00
Number of intersection legs	-418.2	0.00	-856	0.00	-708.7	0	-695.6	0.00	-710	0.00	-686	0.00	-682.1	0.00	-679.2	0.00	-680.8	0.00
Traffic volume * Number of Intersection Legs	0.7167	0.00	0.8117	0.00	0.7789	0.00	0.7577	0.00	0.7828	0.00	0.8096	0.00	0.8189	0.00	0.702	0.00	0.775	0.00
Segment Length	-0.513	0.00	-1.1	0.00	-0.898	0.00	0.914	0.00	-0.935	0.00	-0.878	0.00	-0.868	0.00	-0.812	0.00	-0.846	0.00
Average Speed	-22.8	0.00	17.2	0.00														
Segment Length *Average Speed	0.0441	0.00	0.1756	0.00	0.1312	0.00	0.1353	0.00	28.91	0.00	0.1263	0.00	0.1241	0.00	0.126	0.00	0.1246	0.00
Number of Intersection Legs * Average Speed	67.25	0.00	133.6	0.00	111.3	0.00	110.5	0.00	0.1375	0.00	106	0.00	104.9	0.00	106.5	0.00	105.4	0.00
Traffic volume *Number of Intersection Legs	-0.2299	0.00	-0.2857	0.00	-0.2655	0.00	-0.2529	0.00	111.6	0.00	-0.265	0.00	-0.267	0.00	-0.2258	0.00	-0.2512	0.00
% of Heavy Vehicle *Average Speed	-1.0759	0.00	-1.1168	0.00	-1.101	0.00	-0.835	0.00	-0.2635	0.00	-0.153	0.00						
Traffic volume *Signal Cycle Length	0.01253	0.00	0.0105	0.00	0.01099	0.00	0.01042	0.00	-0.4723	0.00	0.0121	0.00	0.0124	0.00	0.00809	0.00	0.0108	0.00
Average Segment Width * Average Speed	-22.1	0.00	-53.21	0.00	-42.49	0.00	-42.35	0.00	0.01147	0.00	-40.81	0.00	-40.4	0.00	-14.07	0.00	-40.63	0.00
Traffic Volume *Average Speed			0.03751	0.00	0.02609	0.00	0.02416	0.00	-42.94	0.00	0.025	0.00	0.0251	0.00	0.02632	0.00	0.0256	0.00
Traffic Volume *% of Heavy Vehicles							-0.00625	0.00	0.02541	0.00	-0.009	0.00	-0.01	0.00	-0.00831	0.00	0.0051	0.00
% of Heavy Vehicle *Segment Length									-0.0058	0.00	-0.01	0.00	-0.012	0.00	-2.152	0.00	-1.21	0.00
Average Segment Width *% of Heavy Vehicles											-0.502	0.00	0.593	0.00	-0.515	0.00	-0.1749	0.00
% of Heavy Vehicle * Signal Cycle Length															0.023	0.00	0.1979	0.00
R^2	88.41		88.56		88.54		88.63		88.66		88.69		88.69		88.73		88.71	
R^2_{Adj}	88.28		88.43		88.42		88.49		88.52		88.55		88.55		88.58		88.57	

Table 4.14: Final Interaction regression results

	Coefficients				95 % Confidence interval	
Terms	Beta	Stad.Error	t	Sig(p-value)	Lower bound	Upper bound
Constant	608	55.8	10.89	0.00	498.5	717.4
Traffic Volume	-2.081	0.334	-6.24	0.00	-2.736	-1.427
Segment Grade	500.2	22.6	22.13	0.00	455.8	544.5
Average Segment Width	214.7	26.6	8.08	0.00	162.6	266.5
% of Heavy Vehicles	49.32	5.42	9.1	0.00	38.69	59.96
Number of Intersection Legs	-680.8	59.6	-11.42	0.00	-797.7	-563.9
Segment Length	-0.846	0.131	-6.46	0.00	-1.102	-0.589
Traffic Volume * Segment Grade	-0.1351	0.0155	-8.71	0.00	-0.1655	-0.1047
Traffic Volume * Average Segment Width	-0.2512	0.0198	-12.69	0.00	-0.2901	-0.2124
Traffic Volume *% of Heavy Vehicles	-0.00897	0.00167	-5.37	0.00	-0.01225	-0.00569
Traffic Volume * Number of Intersection Legs	0.775	0.0533	14.53	0.00	0.6703	0.8796
Traffic Volume *Signal Cycle Length	0.01078	0.00186	5.81	0.00	0.00714	0.01442
Traffic Volume *Average Speed	0.02561	0.00454	5.64	0.00	0.01671	0.03451
Segment Grade *Average Segment Width	-39.08	1.89	-20.68	0.00	-42.79	-35.38
Average Segment Width *% of Heavy Vehicles	-1.21	0.242	-5	0.00	-1.684	-0.736
Average Segment Width*Average Speed	-40.63	4.36	-9.32	0.00	-49.17	-32.08
% of Heavy Vehicles*Signal Cycle Length	-0.1979	0.021	-9.42	0.00	-0.2391	-0.1567
% of Heavy Vehicles*Segment Length	0.00506	0.00179	2.83	0.00	0.00155	0.00856
Number of Intersection Legs*Average Speed	105.4	10.5	10.04	0.00	84.8	126
Segment Length*Average Speed	0.1246	0.0243	5.13	0.00	0.0769	0.1723

Final interaction regression equation written as follows

$$TTR=608.0-2.081TV-500.2SG+214.7ASW-49.32\%HV-680.8NIL-0.846SL-0.1351TV*SG-0.2512TV*ASW-0.00897TV*\%HV-0.777TV*NIL-0.01078TV*SCL-0.2561TV*AS-39.08SG*ASW-1.21ASW*\%HV+40.63ASW*AS-0.1979\%HV*SCL-0.00506\%HV*SL+105.4NIL*AS+0.1246SL*AS$$

Where: TTR=Travel time reliability

TV=Traffic Volume

SG=Segment Grade

ASW=Average Segment Width

%HV=% of Heavy Vehicles

NIL=Number of Intersection Legs

SL=Segment Length

TV*SG=Traffic Volume with Segment Grade

TV*ASW=Traffic Volume with Average Segment Width

TV*%HV=Traffic Volume with % of Heavy Vehicles

TV*NIL=Traffic Volume with Number of Intersection Legs

TV*SCL=Traffic Volume with Signal Cycle Length
TV*AS=Traffic Volume *Average Speed
SG*ASW=Segment Grade *Average Segment Width
ASW*%HV=Average Segment Width with % of Heavy Vehicles
ASW*AS=Average Segment Width with Average Speed
%HV*SCL=% of Heavy Vehicles with Signal Cycle Length
%HV*SL=% of Heavy Vehicles with Segment Length
NIL*AS=Number of Intersection Legs with Average Speed
SL*AS=Segment Length with Average Speed

4.8. Model interpretation

As shown in the table above p-value of all explanatory variable is less than significance level. This implies that all the predictor variables significantly predict travel time reliability with R^2_{adj} =0.843 and it interpreted as all selected variables traffic volume, segment grade, average segment width, % of heavy vehicles, number of intersection legs, signal cycle length, segment length and average speed are significant and explain 84.3% of the observed data.

From the above regression coefficient, we have clearly understand that traffic volume, segment grade, % of heavy vehicles, number of intersection legs signal cycle length and segment length have an inverse relationship with travel time reliability of a road segments. When traffic volume increase congestion level of a road segment becomes worse this outcome increase vehicle travel time and driver drives below free flow speed and road segment reliability decrease

Segment grade has inverse relationship with travel time reliability. When approach grade increases vehicle travel time come to be increase. So, this result indicates that the reliability of a road segment decrease. Similarly, % of heavy vehicle, number of intersection legs, signal cycle length and segment length have inverse relationship with travel time reliability it implies that those factors increase travel time reliability of a road segment decrease.

Average segment width has positive coefficient implying positive relationship with travel time reliability. When the segment width is wider drivers feel comfortable and relaxed which makes vehicles to slow down or speed up depending on the driver's choice. Therefore, travel time becomes appropriate and road segment will be reliable.

Next, when we look at the coefficient of average speed, it is seen that a positive relationship with travel time reliability.

Next, when we look at the coefficient of number of regression in interaction effect, traffic volume*segment grade, traffic volume *% of heavy vehicles, Traffic volume *signal cycle length, traffic volume *segment length, % of heavy vehicles*Number of intersection legs, % of heavy vehicle signal cycle length and number of intersection legs*segment length it is seen that all have a negative relationship with travel time reliability.

The combined effect of traffic volume with segment grade,% of heavy vehicles, signal cycle length and segment length has negative relationship with travel time reliability as traffic volume and segment grade, traffic volume and % of heavy vehicle, Traffic volume and signal cycle length, traffic volume and segment length increase simultaneously, reliability of urban road segment will be decrease.

Finally, when we look at the combined effect of % heavy vehicle and number of intersection legs, signal cycle length, both have negative relationship travel time reliability of a road. As number of % of heavy vehicle, number of intersection legs and signal cycle length increase simultaneously, reliability of a road segment will decrease.

CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

In this study the researcher aimed to describe and recognized the nature of vehicles travel time at urban road segments and come up with a more general way of estimating vehicles travel time at urban road segment by developing probability distribution and multiple linear regression models.

- This study set out to determine the significant factors which affect travel time reliability in urban road segments and by calculating travel time reliability measurements, developing travel time probability distribution and multiple regression models.
- Travel time reliability measurements buffer time index, buffer time and 95th percentile travel time was calculated for each segment.
- From the analysis buffer time index result in worse reliability condition. Segment 2 produced a buffer time index of about 53.2%, 52% for segment 1, 27% for segment 3, 43% for segment 4, 49% for segment 5, 42% for segment 6 and finally for segment 7 which presented 19.5% buffer time index.
- The 95th percentile travel time about (91.52, 402.98, 90.3, 245.38, 131.32, 82.94 and 94.37) sec for segment 1, 2, 3,4,5,6, & 7 respectively. Higher value for 95th percentile travel time is obtained for segment 2.
- Travel time probability distribution model was developed for selected vehicles (cars and minibus). Distribution parameters were estimated from the data for the selected probability distributions.
- Normal probability distribution was found to fit travel time data of cars for segment 1,2,3,5 and 6.
- Large extreme value distribution fit to travel time data of cars for segment 7 and normally distribution was found of minibus.
- The smallest extreme value distribution was found to fit travel time data of minibus for segment 2, 4 and 6.
- From this we conclude that travel time data for different vehicle type doesn’t follow the same probability distribution.
- Multiple linear regression models were developed to predict travel time reliability of urban road segments. From the statistical analysis done and result obtained all variables were

found statistically significant to the model at 95% confidence level and the calculated R^2_{adj} for the regression is 0.843.

- Traffic volume, segment grade, % of heavy vehicles, number of intersection legs, signal cycle length and segment length are negatively correlated with travel time reliability and average segment width and average speed are correlated positively.
- Interaction effect of traffic volume*segment grade, traffic volume *% of heavy vehicles, Traffic volume *signal cycle length, traffic volume *segment length, % of heavy vehicles*Number of intersection legs, % of heavy vehicle signal cycle length and number of intersection legs*segment length it is seen that all have a negative relationship with travel time reliability.
- A developed regression model can also be used to identifying these factors is a key contribution to HCM and AASHTO where studies are being conducted to include reliability factors into roadway design features.
- The correlation among the influencing factors, which is necessary for traffic operations, traffic design, and long-range transportation planning. The correlation explains the contributing impacts of other factors on travel time reliability improvement on a corridor by modifying one factor.

5.2. Recommendations

We know that urban arterial road is the most important infrastructure its primary function of to deliver traffic from collector roads to freeways and between urban centers at the highest level of service.so, because of this and other factors everybody should be known factors which affect reliability of urban arterial segment in recurrent and non-recurrent conditions especially traffic engineers. In this study a linear regression model develops by selecting eight explanatory variables in recurrent conditions. But in the future, I strongly recommend that the researcher should be consider

- Travel time reliability factors in non-recurrent condition
- Pedestrian activity related to land use, parking characteristics along the road
- Pavement conditions
- Quality of signal timing (good coordination versus bad coordination)
- Vehicle per hour per lane (VPHPL)
- Left/Right turn lane major and minor roads.

In addition to this in this research travel time probability distribution conducted for selected vehicle types like car and minibus. But in the future travel time probability distribution should fit all types of vehicle because travel time for different vehicle type doesn’t follow the same probability distribution.

Measures that help to improve the travel time reliability

- All the segments on existing facilities and their attributes should be considered separately and traffic operations on the entire corridor should be coordinated. In the case of an already coordinated corridor, increasing the green split and retiming the signals can decrease travel time. However, increasing the green split could have an important effect on average delay.
- Provide reversible lane is a lane which traffic may travel in either direction depends on certain conditions. Typically, it’s meant to improve traffic flow during rush hours, this will improve travel time reliability of a road segment.
- Provide separate lanes for specific user groups like
    - Bus lane as part of a bus way system
    - Express toll lanes
    - Hov lane
- Widen roads

Reference

Akito Higatani , Toshihiko Kitazawa , Jun Tanabe , Yoshiki Suga, R. S. & Y. A. (2009) ‘Empirical analysis of travel time reliability measures in hanshin expressway network. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 13(1), pp. 28–38. doi: 10.1080/15472450802644454.

Al-Deek, E. B. E. and H. (2006) ‘Using Real-Life Dual-Loop Detector Data to Develop New Methodology for Estimating Freeway Travel Time Reliability. *Journal of the Transportation Research Board* (1959), pp. 140–150.

Feng Guo, Dengfeng Zhang, and H. R. (2015) ‘BAYESIAN TRAVEL TIME RELIABLITY MODELS. Journal of transportation Enginerring ,pp. 814–865.

Bertini, K. L. and R. L. (2008) ‘Using Travel Time Reliability Measures to Improve Regional Transportation Planning and Operations.Journal of transportaion research board, pp. 1–10. doi: 10.3141/2046-01.

Ch. Ravi Sekhar, E. Madhu, B. K. and S. G. (2012) ‘Analysis of travel time reliability of an urban corridor using micro simulation techniques. Journal of research articles (August), pp. 1–14.

Chen, C. *et al.* (2014) ‘Travel Time Reliability as a Measure of Service. Journal of Transportation Research Record 1855 (November). doi: 10.3141/1855-09.

Chepuri, A., Borakanavar, M. and Amrutsamanvar, R. (2018) ‘Examining Travel Time Reliability under Mixed Traffic Conditions : A Case Study of Urban Arterial Roads in Indian cities. Journal of Asian Transport Studies 5(1), pp. 30–46.

Doohee Nam, D. P. and A. K. (2017) ‘Estimation of Value of Travel Time Reliability. *Journal of Advanced Transportation*, 39(I), pp. 39–61.

Montgomery, D. C. (2012) *Applied Statistics and Probability for Engineers Sixth Edition*.

Emam, E. B. and Al-deek, H. (2006) ‘Using Real-Life Dual-Loop Detector Data to Develop New Methodology for Estimating Freeway Travel Time Reliability.Journal of the Transportation Research Board ,(1959), pp. 140–150.

Fils, P. B. (2012) ‘Modeling Travel Time and Reliability on Urban Arterials for Recurrent Conditions. *Graduate Theses and Dissertations.*

Guessous, Y. *et al.* (2014) ‘Estimating Travel Time Distribution Under Different Traffic Conditions.Journal of *Transportation Research Procedia*. Elsevier B.V., 3(July), pp. 339–348. doi: 10.1016/j.trpro.2014.10.014.

Higatani, A. *et al.* (2009) 'Empirical analysis of travel time reliability measures in hanshin expressway network. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 13(1), pp. 28–38. doi: 10.1080/15472450802644454.

Index, T. T. and Index, P. T. (2010) '2010 Urban Congestion Trends Enhancing System Operations.International Journal of transportation, pp. 1–8.

Jammula, J. K. *et al.* (2018) 'TRAVEL TIME PREDICTION MODELLING IN MIXED TRAFFIC. International Journal for Traffic and Transport Engineering, 8(1), pp. 135–147.

Li, R., Chai, H. and Tang, J. (2013) 'Empirical Study of Travel Time Estimation and Reliability. Journal of research Articles.

Liu, T. (2016) 'Power Comparison of Some Goodness-of-fit Tests. *In International Encyclopedia of Statistical Science* (pp. 52-54). Springer, Berlin, Heidelberg

Lohmiller, J. (2012) 'Factors influencing the travel time reliability of motorway sections. *Journal of the Transportation Research Board, 1678,* 42–47.

Mannering, S. P. W. M. G. K. F. L. and CHAPMAN (2003) *S tatistical AND E conometric M ethods FOR T ransportation D ata A nalysis Library of Congress Cataloging-in-Publication Data*.

Narang, R. V (2012) 'AC 2012-4278 : Introducing Reliability and Maintainability in Engineering and Technology.

POLUS, A. (2012) 'Modeling Travel Time and Reliability on Urban Arterials for Recurrent Conditions. Journal of transportaion Enginerring,pp. 12-35.

R, S. P. and Mathew, S. (2013) 'EVALUATION OF OPERATIONAL EFFICIENCY OF URBAN ROAD NETWORK USING TRAVEL TIME RELIABILITY MEASURES. International Journal of Research in Engineering and Technology, pp. 155–160.

Hesham Rakha, Sangjun Park, and F. G. (1909) 'Travel Time Reliability Modeling',Journal of Transportation Research Building , pp. 814–865.

APPENDIX

Appendix A

- Testing multiple linear regression assumption

Assumption 1:- Linear relationship between dependent variable and each of independent variables There should be linear relationship between the dependent variable and each of the independent variables. Scatter plot of each independent variables versus the dependent variable shows that there is no significant evidence for the existence of nonlinear relationship. (See ***Figure A-1-A7).***

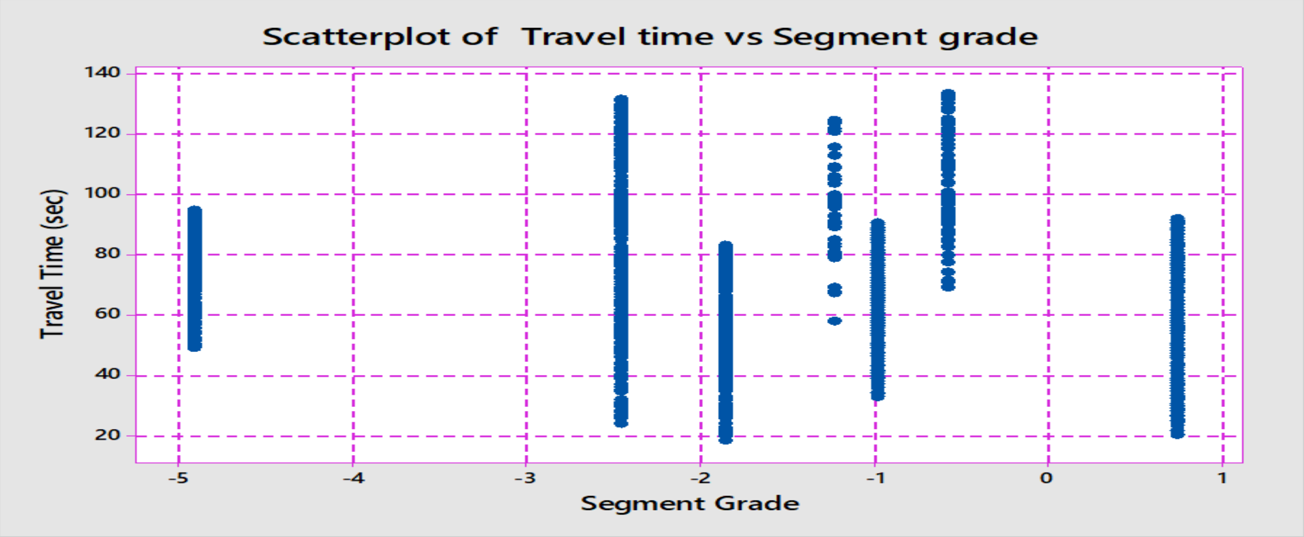


Figure A.1: Scatter plot of Reliable travel time vs. Segment grade

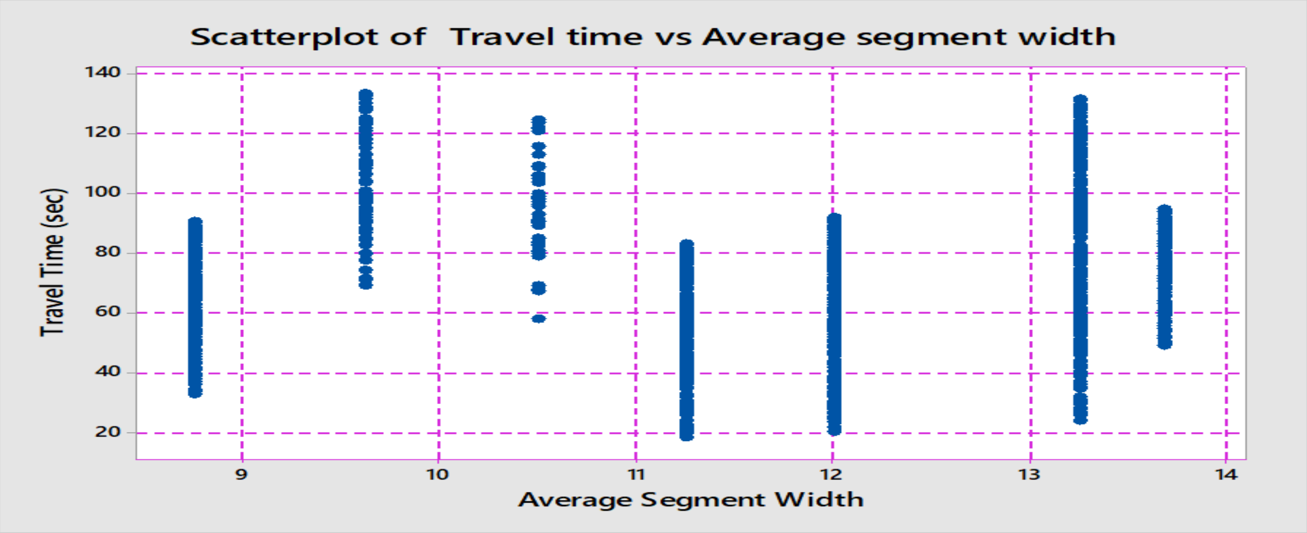


Figure A.2: Scatter plot of Reliable travel time vs. Average segment width

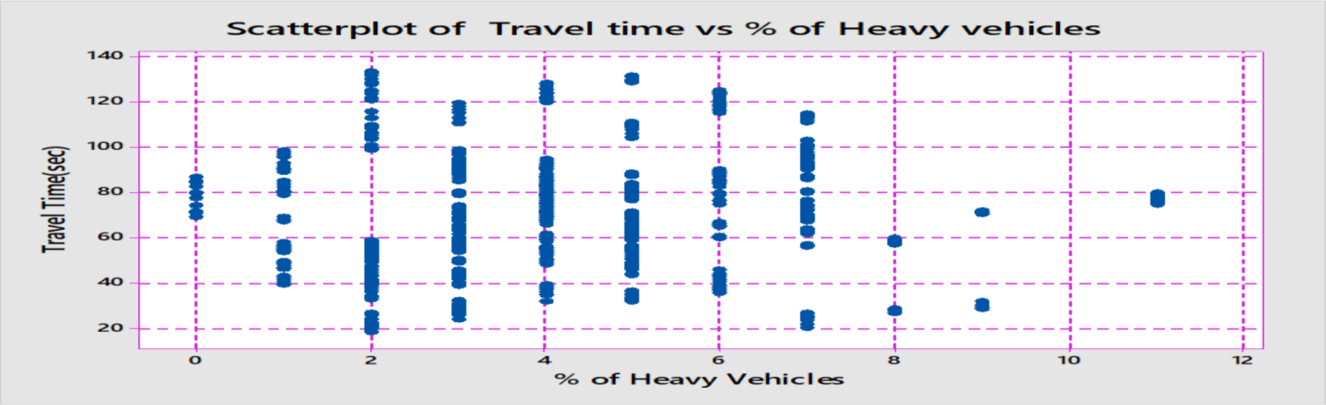


Figure A.3: Scatter plot of Reliable travel time vs. % of heavy vehicles

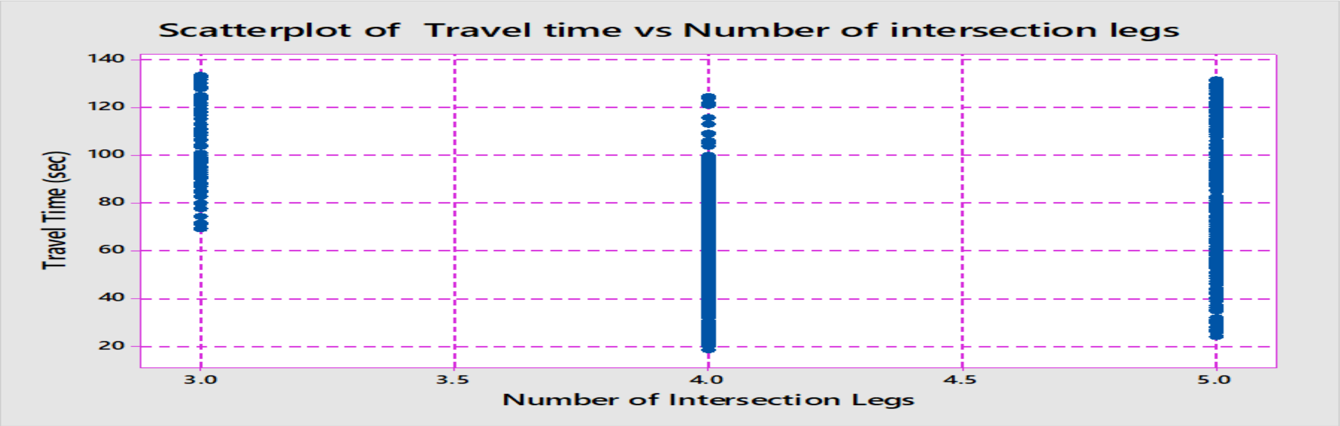


Figure A.4: Scatter plot of Reliable travel time vs. Number of intersection legs

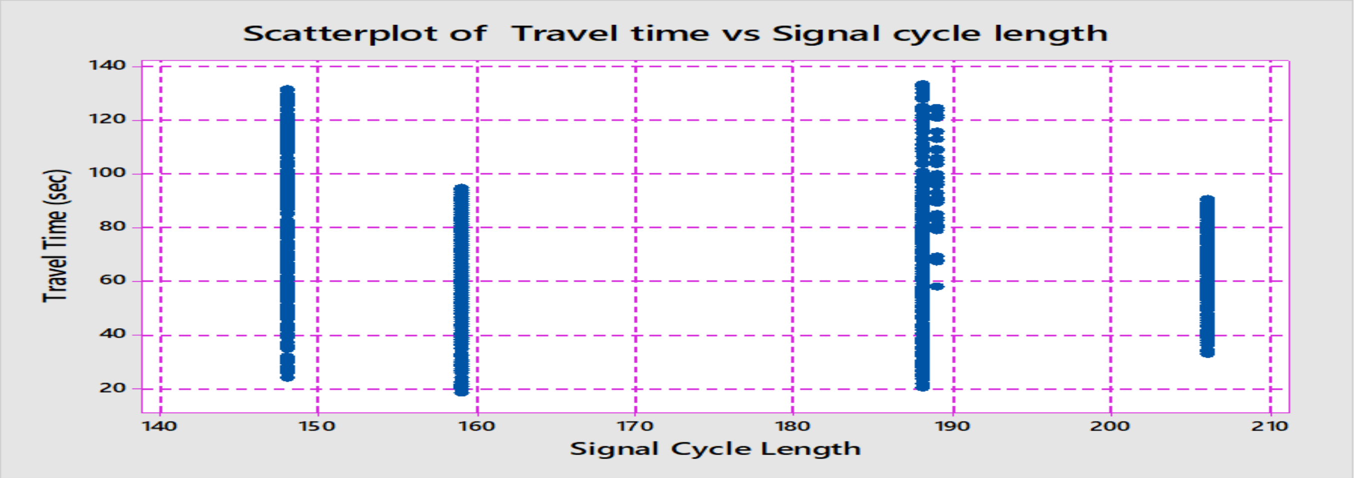


Figure A.5: Scatter plot of reliable travel time vs. Signal cycle length

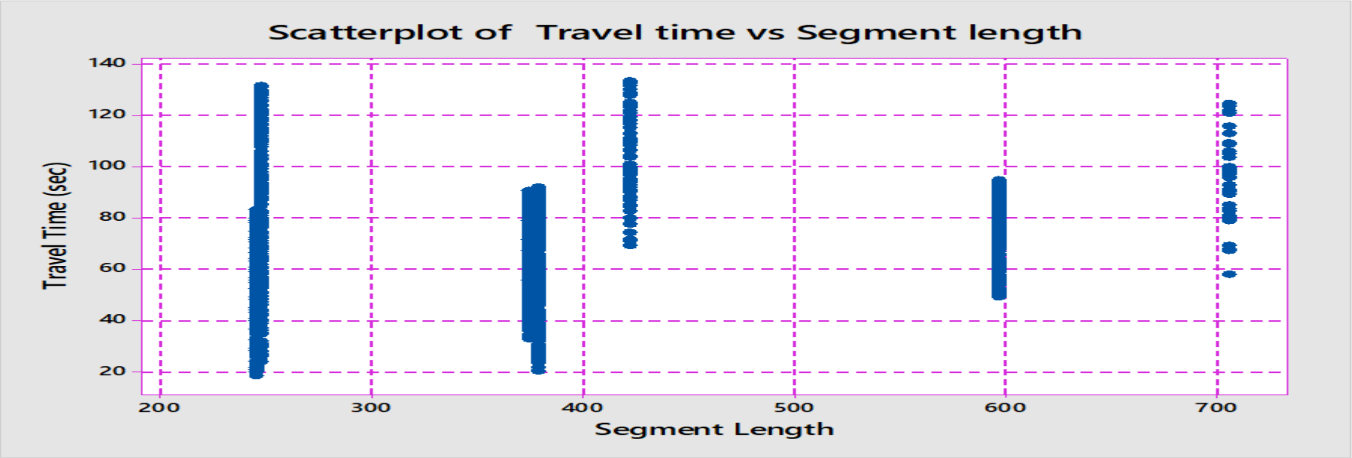


Figure A.6: Scatter plot of Reliable travel time vs. Segment length

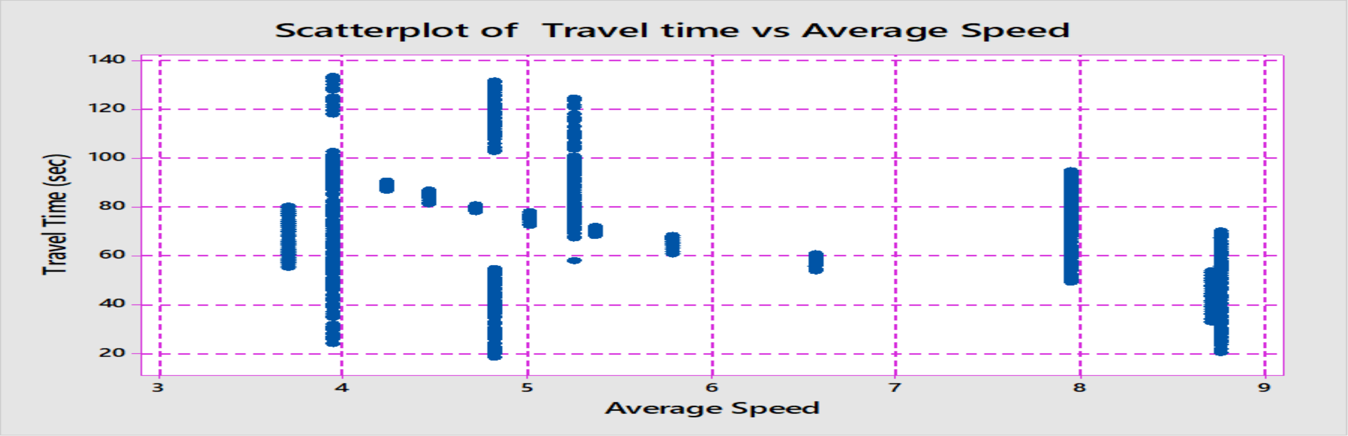


Figure A.7: Scatter plot of Reliable travel time vs. Average speed

Appendix B

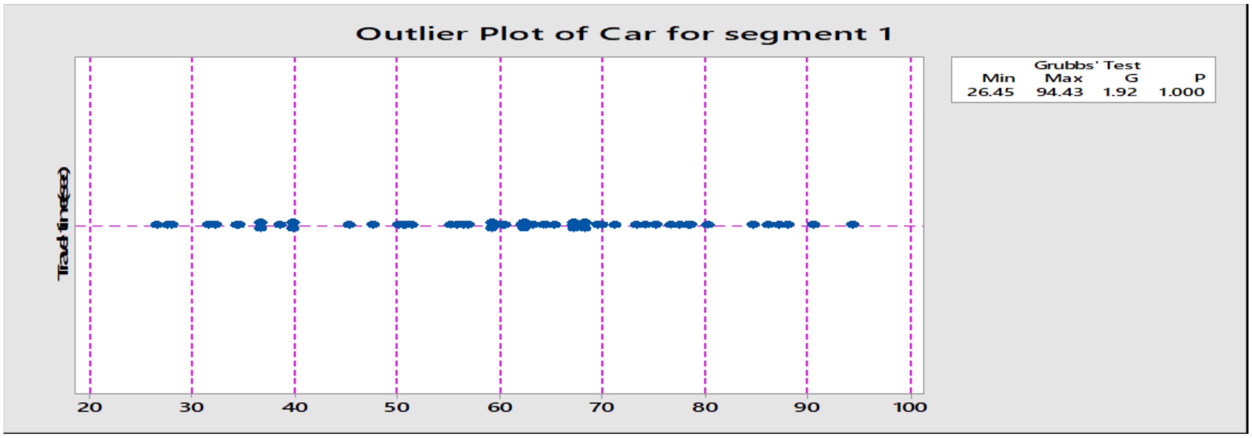
- Outlier Plot and Outlier Determination

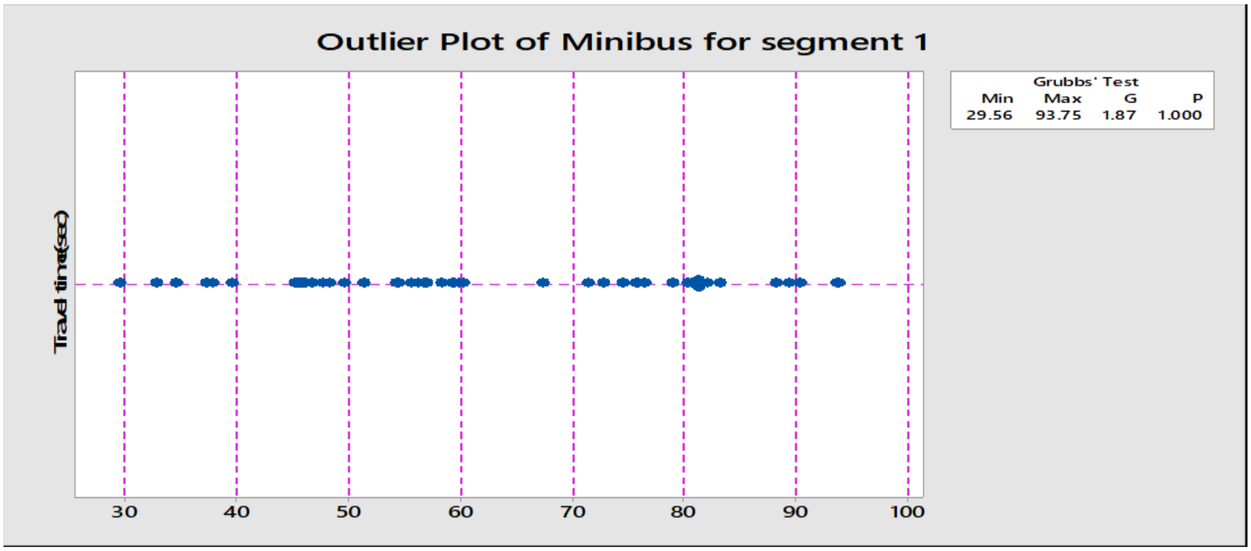
Before conducting the analysis, the data set should be tested for outliers and influential points. This could be done with Grubbs' test with the following Null and Alternate Hypothesis.

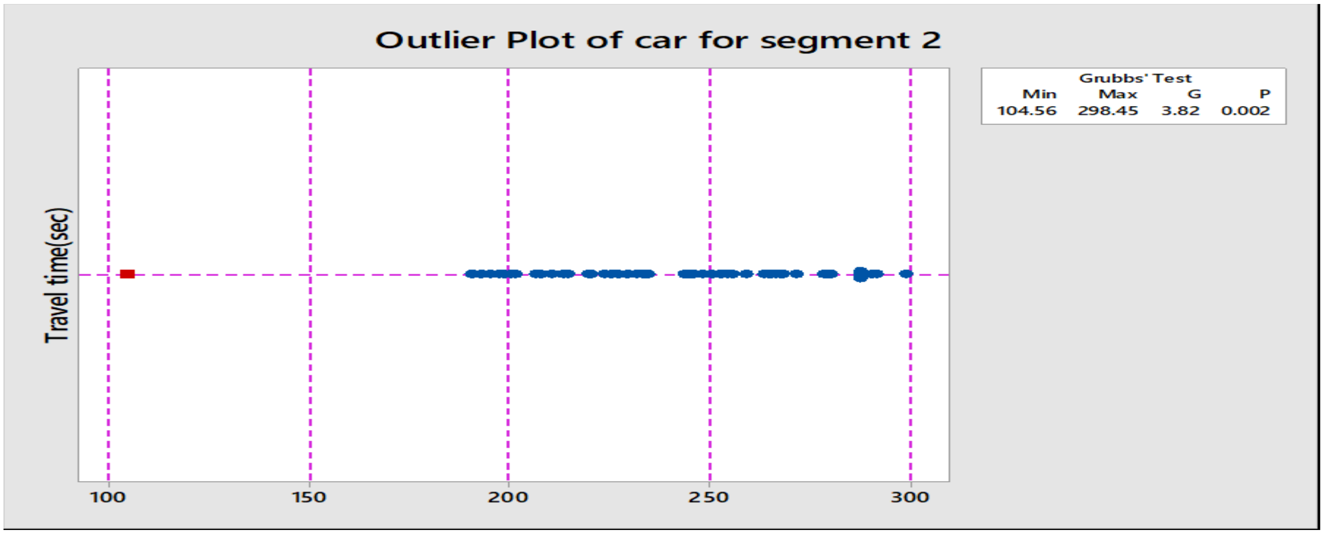
Ho: - All the data values come from the same normal population

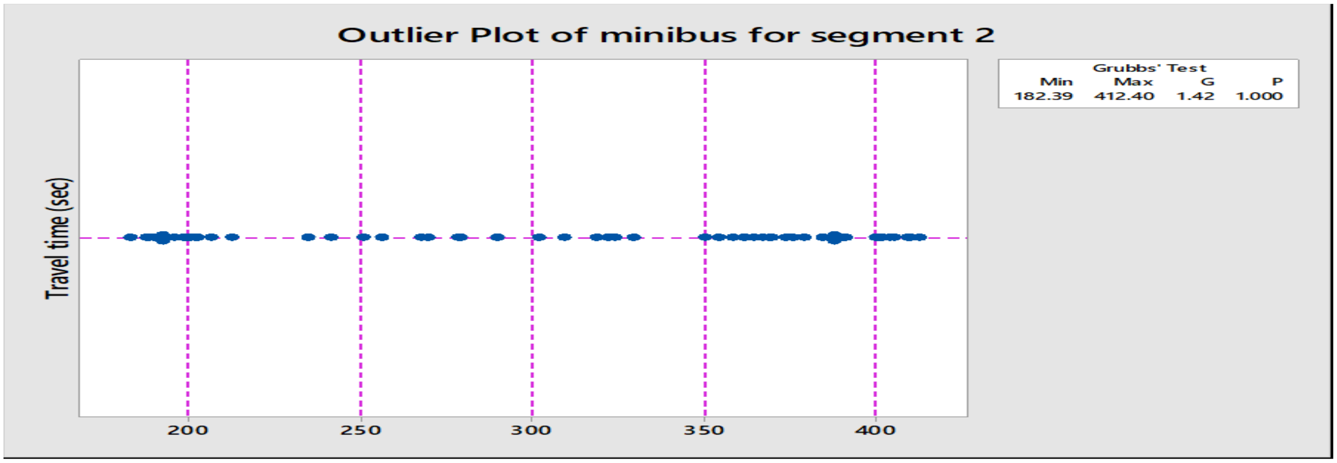
H1:- Smallest or largest data value is an outlier

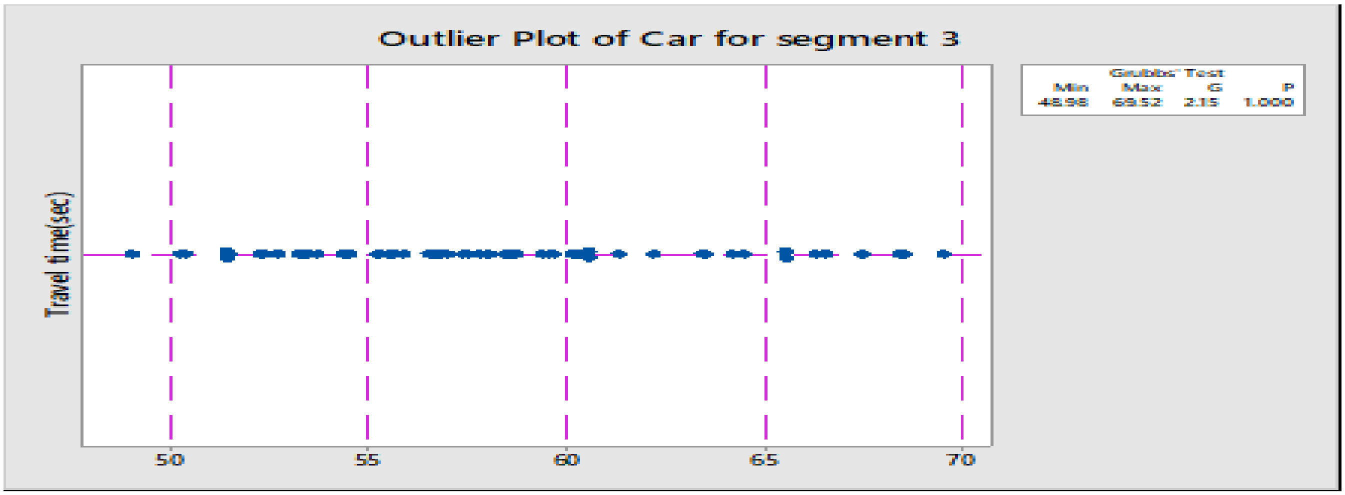
Significance Level: - 0.05

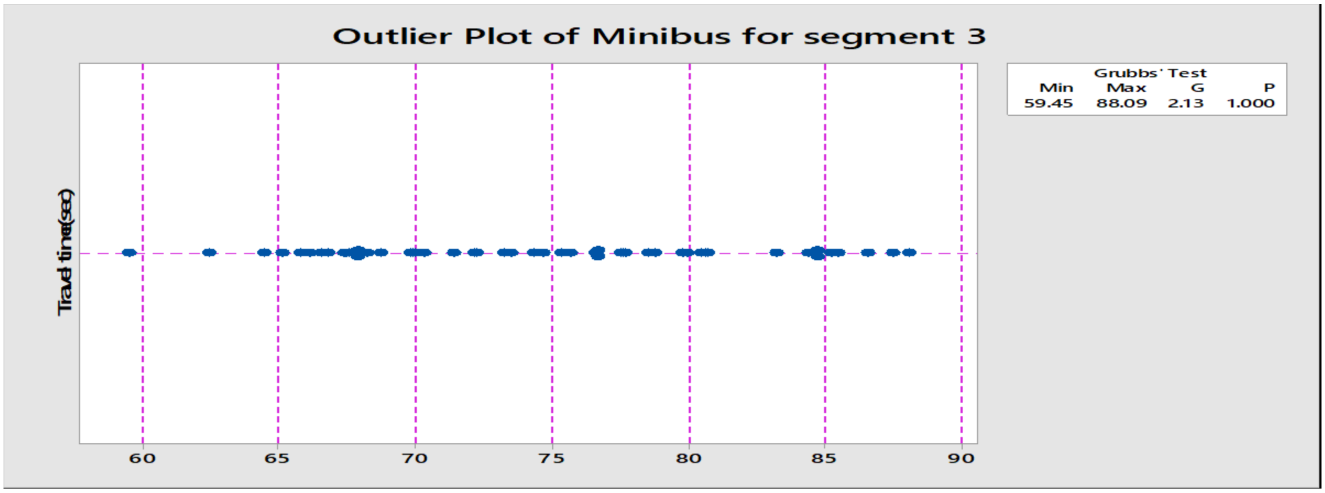


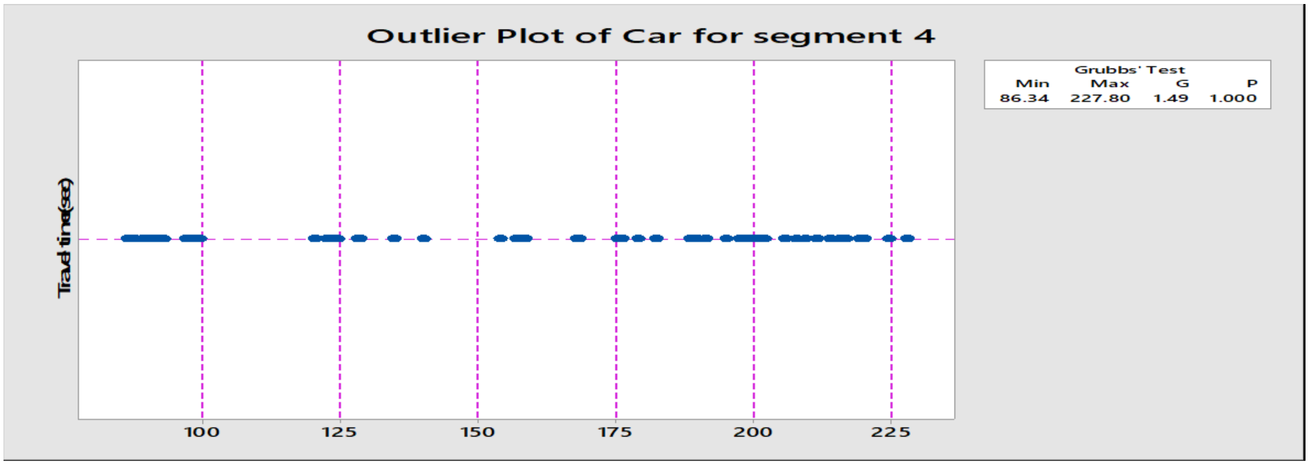


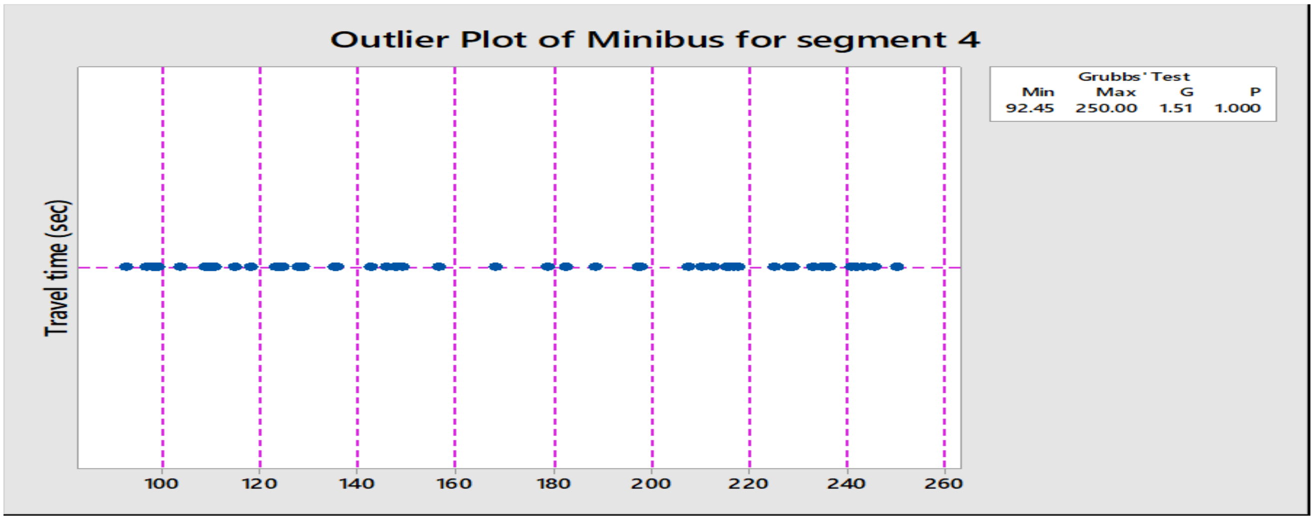

Outlier Plot of car for segment 2
Grubbs' Test
Min Max G P
104.56 298.45 3.82 0.002
Travel time(sec)
100
150
200
250
300

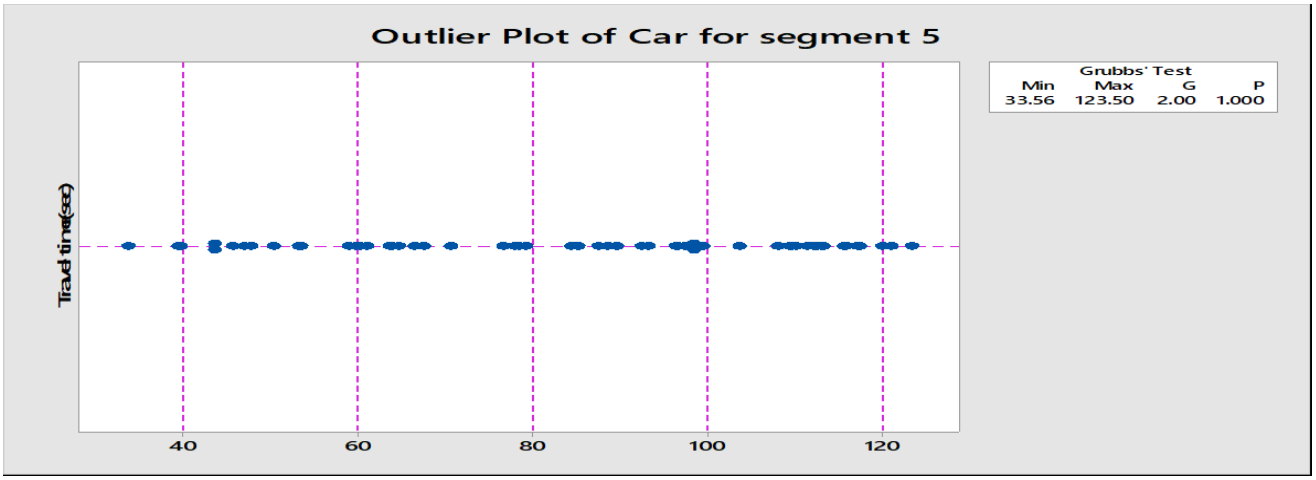

Outlier Plot of minibus for segment 2
Grubbs' Test
Min Max G P
182.39 412.40 1.42 1.000
Travel time (sec)
200
250
300
350
400

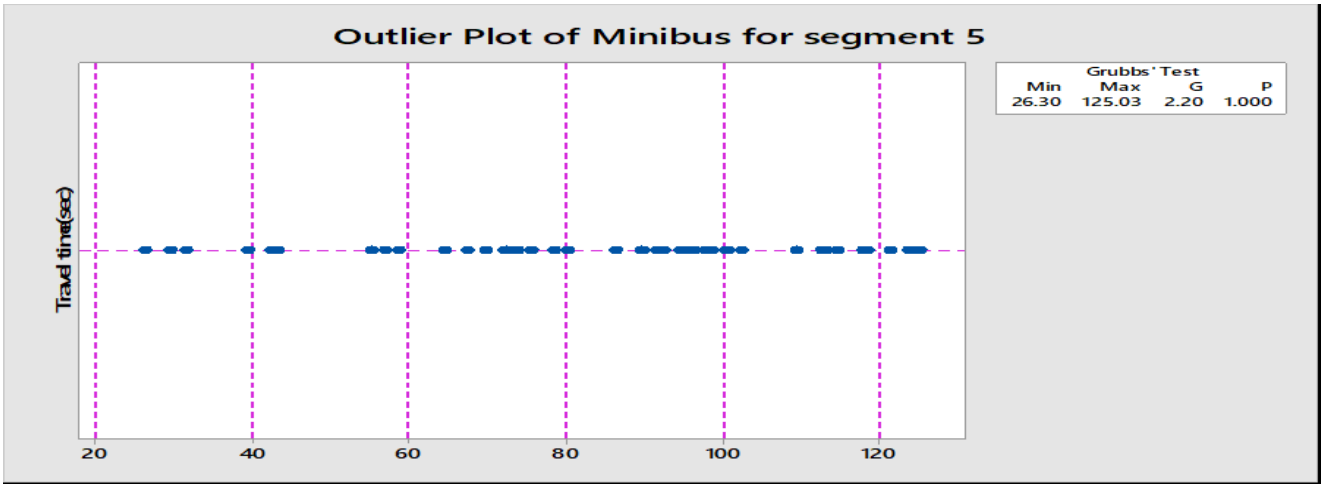

Outlier Plot of Car for segment 3
Grubbs' Test
Min Max G P
48.98 69.52 2.15 1.000
Travel time(sec)
50
55
60
65
70

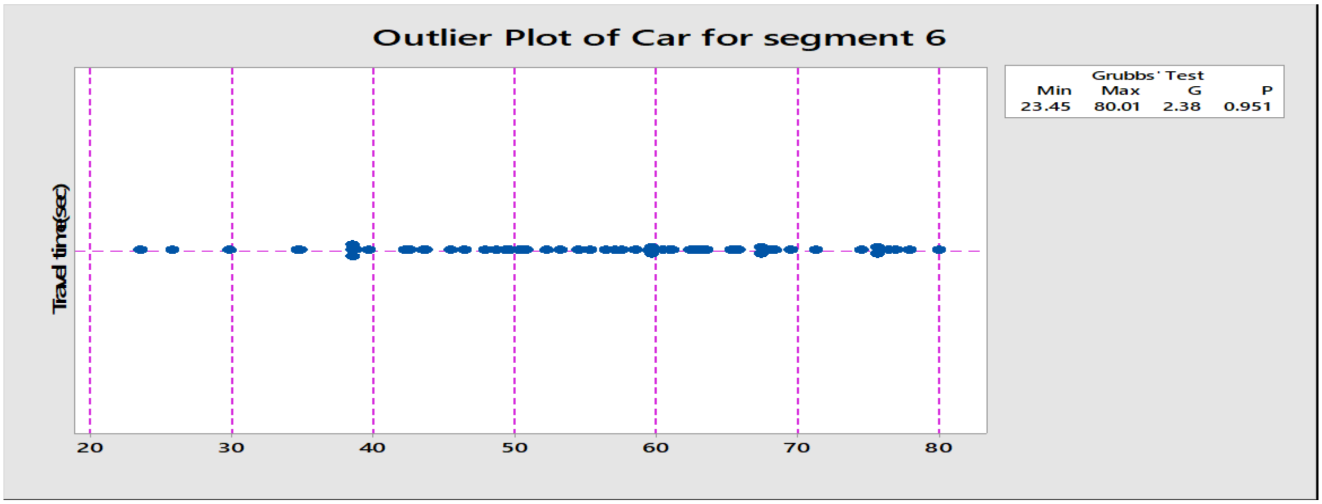

Outlier Plot of Minibus for segment 3
Travel time(sec)
60
65
70
75
80
85
90
Grubbs' Test
Min 59.45
Max 88.09
G 2.13
P 1.000

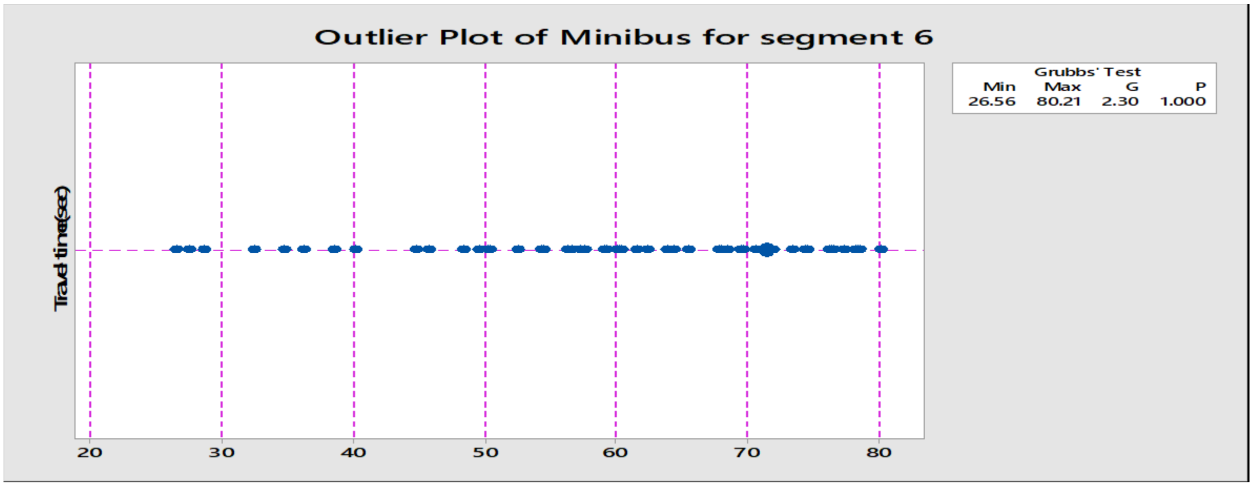

Outlier Plot of Car for segment 4
Travel time(sec)
100
125
150
175
200
225
Grubbs' Test
Min 86.34
Max 227.80
G 1.49
P 1.000

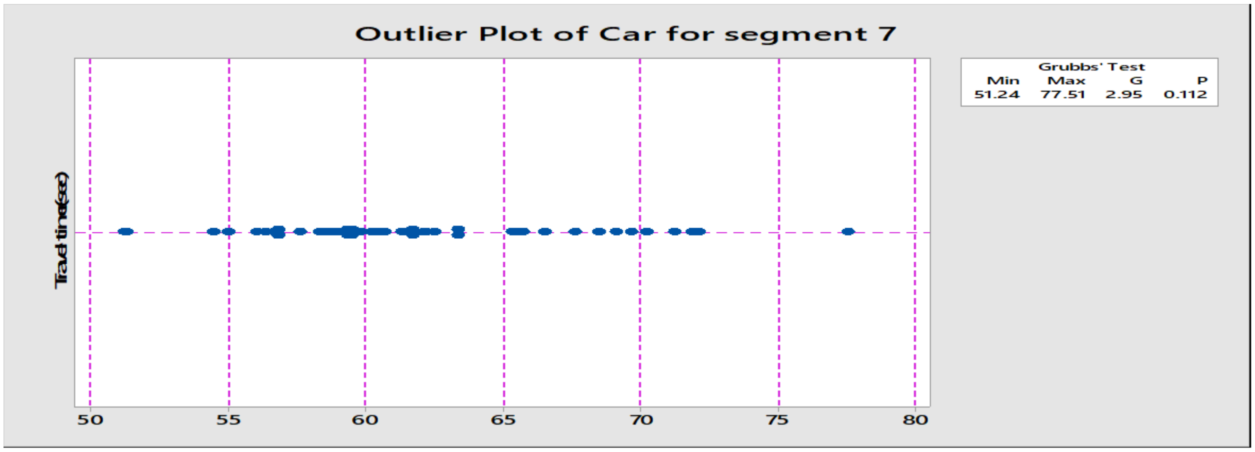

Outlier Plot of Minibus for segment 4
Travel time (sec)
100
120
140
160
180
200
220
240
260
Grubbs' Test
Min 92.45
Max 250.00
G 1.51
P 1.000


Outlier Plot of Car for segment 5
Grubbs' Test
Min 33.56
Max 123.50
G 2.00
P 1.000
Travel time(sec)
40
60
80
100
120


Outlier Plot of Minibus for segment 5
Grubbs' Test
Min 26.30
Max 125.03
G 2.20
P 1.000
Travel time(sec)
20
40
60
80
100
120


Outlier Plot of Car for segment 6
Grubbs' Test
Min 23.45
Max 80.01
G 2.38
P 0.951
Travel time(sec)
20
30
40
50
60
70
80





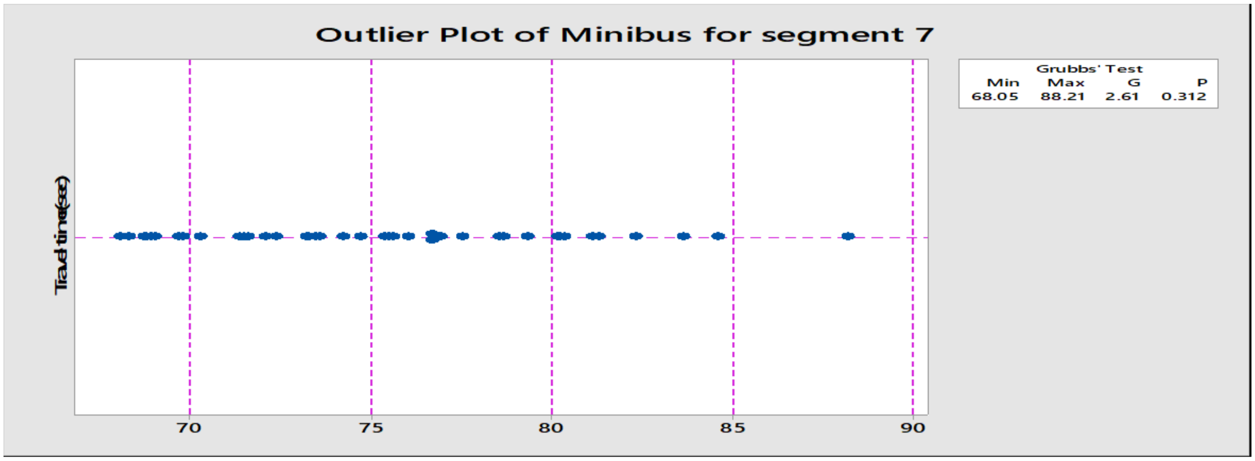


Table B.1: Grubb's Test Result

Variable	N	Mean	StDev	Min	Max	G	P
Car(S1)	61	60.49	17.75	26.45	94.43	1.92	1.00
Minibus (S1)	55	61.59	17.2	29.75	93.75	1.28	1.00
Car(S2)	63	302.3	84.3	182.4	412.4	1.87	1.00

Minibus (S2)	53	241.73	35.93	104.56	298.45	1.42	1.00
Car(S3)	65	58.5	5.135	49.98	69.52	2.15	1.00
Minibus (S3)	58	74.738	7.189	59.45	88.09	3.82	1.00
Car(S4)	64	157.64	47.95	86.34	227.8	2.13	1.00
Minibus (S4)	52	173.07	53.46	92.45	250	1.42	1.00
Car(S5)	62	85.44	25.9	33.56	123.5	1.49	1.00
Minibus (S5)	51	84.69	26.52	26.3	125.03	1.57	1.00
Car(S6)	63	55.56	13.48	23.45	80.01	2.38	0.951
Minibus (S6)	55	59.63	14.39	26.56	80.21	2.3	1.00
Car(S7)	53	61.741	5.344	51.24	77.5	2.61	0.112
Minibus (S7)	46	75.245	4.962	68.05	88.21	2.95	0.312

As we can see from the outlier plot of travel time data for two vehicle type car segment 2 of the plot is identified as an outlier. The p-value from Grubs test result in the above table for the above listed variables is less than 0.05 (level of significance) which leads to the conclusion to accept the alternate hypothesis which is Smallest or largest data value is an outlier. Based on this the following data point are identified as outliers and removed from the analysis.

Table B.2: Outliers

Variable	**Row**	**outlier**
Car(S2)	19	104.56