SIGNALIZED INTERSECTION SAFETY: A CASE STUDY OF KOLLAM CORPORATION

A PROJECT REPORT

Submitted by

LEKSHMI SREEKUMAR

ROLL NO: TKM20CETE08

to

the APJ Abdul Kalam Technological University

in partial fulfillment of the requirements for the award of degree

of

Master of Technology

In

Transportation Engineering

DEPARTMENT OF CIVIL ENGINEERING

T.K.M College of Engineering, Kollam

July 2022

DECLARATION

I undersigned hereby declare that the project report “Signalized intersection safety: A case study of Kollam coporation ”, submitted for partial fulfillment of the requirements for the award of degree of Master of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by me under supervision of Prof. Sai Niveditha M.G. This submission represents my ideas in my own words and where ideas or words of others have been included, I have adequately and accurately cited and referenced the original sources. I also declare that I have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Kollam LEKSHMI SREEKUMAR.

7-07-2022

DEPARTMENT OF CIVIL ENGINEERING

T.K.M. COLLEGE OF ENGINEERING, KOLLAM

CERTIFICATE

Certified that this report entitled **'SIGNALIZED INTERSECTION SAFETY: A CASE STUDY OF KOLLAM CORPORATION** is the report of project presented by **LEKSHMI SREEKUMAR, Roll No. : TKM20CETE08** during **2021-2022** in partial fulfillment of the requirements for the award of the Degree of Master of Technology in Transportation Engineering of the A P J Abdul Kalam Technological University.

Guide:	Project coordinator:	Head of Department:
Prof. Sai Niveditha M.G.	**Dr. Kavitha Madhu**	**Dr. Sajeeb R.**
Assistant Professor	Associate Professor	Professor
Department of Civil Engineering	Department of Civil Engineering	Department of Civil Engineering

ACKNOWLEDGEMENT

I take this opportunity to express my deep sense of gratitude and sincere thanks to all who helped me to complete the project successfully.

I am deeply indebted to my guide**, Prof. Sai Niveditha M.G.**, Assistant Professor, Department of Civil Engineering for her excellent guidance, positive criticism and valuable comments.

I am greatly thankful to my project coordinators **Dr. Kavitha Madhu**, Associate professor, Department of Civil Engineering and **Prof. Jijin A**., Assistant Professor, Department of Civil Engineering for their constant supervision as well as for providing necessary information regarding the project.

I am greatly thankful to **Dr. Sajeeb R.**, Professor and Head of Department of Civil Engineering, for his kind support.

Finally, I thank my parents and friends who directly and indirectly contributed to the successful completion of my project.

LEKSHMI SREEKUMAR.

ABSTRACT

In developing nations like India, motorization is increasing along with economic growth. Road traffic deaths in urban India have consistently been a serious issue of concern. National Crime Records Bureau (NCRB) 2014 reports show that urban road traffic crashes within the state of Kerala, India, increased by 37% from 2009 to 2012. Nearly 20% of those crashes occurred at intersections. Crashes at signalized intersections formed 24% of the full reported crashes at intersections, and 40% of traffic related serious injuries and fatalities. Signalized intersections are major black spots in an urban road network

By collecting six years crash data of signalized intersections of Kollam corporation it is found that signalized intersection crashes are increasing each year crashes increased by 72% within six years. It implies the need to control the crashes occurring at signalized intersections of Kollam corporation. The current study investigates the formation of crash frequency prediction model and crash severity prediction model for signalized intersections of Kollam corporation by doing statistical analysis of the crash data. There are totally ten signalized intersections within the Kollam corporation. Six years of crash data of these intersections is collected and details regarding traffic control features, geometric features and traffic volume are also collected for formation prediction model for total crashes and grievous crashes. Six types of regression models are used to analyze crash frequency and the model which best fit the data is chosen as final prediction model for total crashes and grievous crashes. A detailed crash data of four years is collected to develop crash severity prediction model. Ordered probit model is employed to form crash severity prediction model and marginal effects are determined which helps to understand effects of each factors on severity levels. A better understanding of the crash causative factors aids to develop more targeted countermeasures for improving the safety and performance of signalized intersections. Assessment of safety at signalized intersection aid traffic and road safety engineers to adopt better solutions for reducing crashes. Modeling relationship between crash frequency, severity and it's determining factors help to achieve knowledge about crash occurrence mechanism and to come up with safety policies.

Keywords: *Signalized intersection, traffic control features, geometric features*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AIC	Akaike Information Criterion
AICC	Akaike Information Criterion Correction
BIC	Bayesian Information Criterion
CO	Countdown timer
CA	Presence of camera
CT	Cycle Time
F	Presence of Footpath
GCFP	Grievous Crash Frequency Prediction model
IN	Type of Intersection
L	Presence of exclusive Left lane
NB	Negative Binomial regression
PA	Parking on Carriageway
P	Number of Phase
RC	Road category
RM	Presence of Clearly Visible Road Mark
RT	Type of Road
S	Presence of Stop mark
TCFP	Total Crash Frequency Prediction model
TW	Volume of Two Wheelers
V	Traffic Volume
Z	Presence of Zebra crossing
ZINB	Zero Inflated Negative Binomial
ZIP	Zero Inflated Poisson
ZTNB	Zero Truncated Negative Binomial
ZTP	Zero Truncated Poisson

1. INTRODUCTION

Road traffic accidents are the world's leading reason behind death for individuals between the ages of one and twenty-nine. Throughout the globe, cars, buses, trucks, motorcycles, pedestrians, animals, taxis and other categories of travelers, share the roadways, contributing to economic and social development in many countries. Yet annually, many vehicles are involved in crashes that are chargeable for countless deaths and injuries. Globally, every year, about 1.25 million people are killed in vehicle crashes and approximately 50 million more are injured. Following current trends, about two million people might be expected to be killed in automobile crashes every year by 2030. Crashes often occur at intersections because these are the locations where two or more roads cross one another and activities like as turning left, crossing over, and turning right have the potential for conflicts leading to crashes. With economic prosperity, motorization is growing faster in developing countries like India. Road traffic deaths in urban India have consistently been a serious issue of concern. National Crime Records Bureau (NCRB) 2014 reports show that road traffic crashes are the leading reason of unnatural death among men and also the second highest reason for unnatural death of women in India (National Crime Records Bureau, 2014). Urban areas cover a small percentage of land area as compared to rural India, but the number of crashes is significantly higher in urban areas. While the first focus at signalized intersections is to separate the conflicting traffic, poor safety performance at signalized intersections has been a significant safety concern in urban road network, and for cities of developing countries demanding better understanding of the crash occurrence to develop targeted countermeasures. It is thus, essential to test the signal design parameters as well as the infrastructure planning, design, land use and operational condition, which are critical for the safety performance of the signalized intersections. Road crash prediction models are very useful tools in highway safety, given their potential for determining the crash frequency occurrence and also the degree of severity of crashes. Crash frequency refers to the prediction of the number of crashes that might occur on a selected road segment or intersection during a time period, while crash severity models generally explore the relation-ship between crash severity injury and the contributing factors like driver behavior, vehicle characteristics, roadway geometry, and road-environment conditions. Effective interventions to diminish crash include design of safer infrastructure and incorporation of road safety features into land-use and transportation planning.

1.1 PROBLEM STATEMENT

Indian states have been facing traffic related issues including traffic congestions and accidents for decades. Kerala has the fourth highest number of road accidents in the country. The state, which was ranked fifth from 2015 to 2018, stood fourth in the country in terms of the number of road accidents in 2019. According to the "Accident deaths and suicides in India in 2019" report by the National Crime Records Bureau, there were 4,370 fatalities and 44,467 injuries in 39,944 accidents in the State last year (NCRB). Road safety experts attribute the decline in fatalities in Kerala to a harsh enforcement effort by law enforcement, rising motorist awareness, and an increase in penalties. One of the primary places where accidents occur is at intersections. At crossings, there is a significant concentration of the opposing movements. Even though traffic is controlled by signals, intersections are where the majority of accidents occur. Intersections are sites where conflicting movements happen and are bottlenecks of traffic. So if a person drives the vehicle carelessly at an intersection it will definitely lead to an accident. So it is necessary to control the movements at an intersection. But the increasing intersection crashes indicates that even the implementation of signal at an intersection is not enough to control the accident rate. The increasing crashes at signalized intersections clearly indicate the need to develop for safety policies for efficient working of signalized intersections. But before developing safety policies one must have enough knowledge about factors which influence the intersection crashes. The development of crash frequency and crash severity prediction models will help to understand the influence of each factor on the crash occurrence and also help to know frequency and severity level of future crashes.

1.2 OBJECTIVES

The objective of proposed study is as follows:

- To identify factors that raise safety issues at signal controlled intersections of Kollam corporation
- To develop prediction models for total crash frequency and grievous crash frequency
- To develop crash severity prediction model for the signalized intersections of Kollam corporation
- To evolve safety conscious guidelines for efficient operation of signalized intersections of Kollam corporation

1.3 METHODOLOGY

The methodology planned in order to achieve the mentioned objectives is described below:

- **Study of literature:** Background study was done on the various way to conduct signalized intersection safety. Analyzed the existing statistical models to determine the appropriate statistical model to develop prediction models
- **Data collection**: Six years of accident data of the intersections are collected for the development of crash frequency prediction model. Four years (2018-2021) of detailed accident data of intersections is collected for formation of crash severity prediction model Geometric features, traffic control features are collected by visiting each intersections. 24 hour traffic volume of each intersection is also collected for the analysis.
- **Analysis of data**: Correlation analysis is conducted to determine the factors which are highly correlated. The independent variables which are highly correlated each other will be checked and among the highly correlated independent variables the variable which is highly correlated with dependent variable is considered for the analysis. Regression models like Poisson model, Negative binomial model, Zero inflated poisson model, Zero inflated negative binomial model, Zero truncated poisson model and Zero truncated negative binomial model are used to develop prediction model for total crash frequency and grievous crash frequency. The best model is selected based on goodness of fit details. Crash severity is divided into three severity levels like minor injury, grievous injury and fatal injury. Ordered probit model is used for the formation of crash severity prediction model. The influence of each factors in each severity level can be understood by analyzing marginal effects of variables on severity levels.
- **Interpretation of results and discussion**

1.4 SCOPE

- Crash data does not include damage only crashes. So such types of crashes are not considered for the study.
- There may be unreported crashes which are not available in the crash data.

2. LITERATURE REVIEW

There are several methods to analyze crash occurrence at signalized intersections. Some of the methods are:

- Analytical method based on crashes
- Analytical method based on conflicts
- Microsimulation models for safety evaluation
- Real safety models

2.1 ANALYTICAL METHOD BASED ON CRASHES

Given their ability to anticipate both the frequency of crashes and their severity levels, road crash prediction models are particularly helpful tools in highway safety. While crash severity models typically examine the relationship between crash severity injury and contributing factors like driver behavior, vehicle characteristics, roadway geometry, and road-environment conditions, crash frequency is prediction of the number of crashes that would occur on a selected road segment or intersection in a specified time period. Incorporating road safety features into land-use and transportation planning, designing safer infrastructure, enhancing vehicle safety features, improving post-crash care for victims of traffic accidents, and changing driver behavior, such as by establishing and enforcing stricter driving regulations, are all examples of effective interventions that can lower the crash toll. Mitra .S et al,2020 evaluated status of signalized intersection safety of the city Kolkata. They selected 52 signalized intersections for the study. This study specifically aimed to investigate the effects of geometric, infrastructural, traffic control and land use parameters on the number of crashes and their severity across 52 signalized intersections of Kolkata. Correlation analysis was done to know the relationship between the independent variables. Spearman’s rank order correlation, which is a non-parametric version of the Pearson’s correlation, was used. If two independent variables are highly correlated only one variable can be chosen for model formation. Otherwise it causes multicollinearity. Mitra .S et al,2020 selected the independent variables which were highly correlated with dependent variable, among the highly correlated independent variables. Traditional count data models such as the Poisson regression model and Negative Binomial regression model were employed for developing crash frequency prediction model. They employed negative binomial regression model for crash frequency prediction model and binary logit model for crash severity prediction model. Table 2.1 shows result of negative binomial regression model and Table 2.2 shows result of binary logit model.

Table 2. 1Result of Negative binomial regression model for total crashes(Mitra S et al, 2020)

Explanatory variable	Estimated coefficients	t statistic (p value)
Intercept	-0.336	-0.96(0.334)
Total approach volume	$1.90 * 10^{-5}$	6.41(0.000)
Major road blocked	0.031	2.62(0.009)
Right turn protected(1/0)	-0.289	-2.07(0.038)
Major road traffic configuration(One way=0;two way=1)	0.423	1.76(0.077)
Car parking on minor road(1/0)	0.260	1.75(0.08)
Log likelihood function	-107.309	
Restricted log likelihood function	-144.498	
Log likelihood ratio index	0.257	
Number of observations	52	

Table 2. 2 Result of binary logit model for fatal crash occurrence(Mitra S et al,2020)

Explanatory variable	Estimated coefficients	t statistic(p value)
Intercepts	-6.160	-2(0.045)
Right turn protected	-2.744	-2.05(0.039)
Non motorized traffic	0.097	2.07(0.038)
Minor road blocked	0.124	2.03(0.042)

Major road blocked	0.203	1.84(0.064)
Minor road straight	2.683	1.83(0.066)
Ratio of minor road to major road traffic volume	-3.660	-1.8(0.071)
Total approach volume	$4.787 * 10^{-5}$	1.78(0.075)
Log likelihood function	-22.095	
Restricted log likelihood function	-35.89	
Log likelihood ratio index	0.384	
Number of observations	52	

Another study conducted by Anjana.S et al,2014 examined the crash causative factors of signalized intersections under mixed traffic using advanced statistical models. In the Indian state of Kerala, three significant urban agglomerations—Trivandrum, Ernakulam, and Kozhikode—were chosen for signalized intersection selection. The database for this study is made up of data on traffic, geometry, and signal operation for 110 approach legs from 32 intersections. The approach level and intersection level are the two levels in the hierarchical structure of crash data. The hierarchical structure of crash data must be taken into account in a model for crash prediction. For modeling crash frequency, a hierarchical Poisson regression model is adopted as given below:

Level-1 model

$$Log(DA)= B0 + 0.053906 *(gd/gp)+ 0.000095 (ADTPL)+ 0.094084 *MW -0.074217 *(LM)$$

Level-2 model

$$B0= 0.47161 + 0.315259 *(CT) -0.495367 *(SC)+ 0.582814 * (PTW)$$

For severity prediction, a two level hierarchical logistic regression model is developed with crash level (level 1) and approach level (level 2). The developed model is as follows:

Level-1 model

$$Log(P/1-P) = B0 + 0.054110 *(D \text{ or } N) + 0.157667 *TC -0.677859 * (AV) -0.134989 *(VV)$$

Level-2 model

$$B0 = 3.856585-0.104175*(gd/gp) -0.000039 *(ADTPL)-0.277161*(PRT)-1.030969 *(ELTL)$$

ADTPL =Average daily traffic per lane width

LM= Presence of lane markings

MW= Median width

CT =Presence of countdown timer

SC= Surveillance camera

D= Day

N= night

PTW =Proportion of two wheeler in traffic stream

AV =Accused vehicle

VV = Victim vehicle

ELTL = Exclusive left turn lane

TC =Collision type

g_d = Design green time

g_p = Provided green time

Given their potential to determine both the frequency of accident occurrence and the underlying elements that might subsequently be addressed by transportation policy, traffic accident prediction models are particularly helpful tools in highway safety. Vehicle crash data can be used to model both the likelihood that a crash will occur and its severity level. Crash frequency is the estimation of the number of collisions that will take place on a particular road segment or intersection during a given period of time. The majority of collision severity approaches examine the connections between different injury categories and the contributing elements that include driver conduct, vehicle attributes, road geometry, and other aspects of the road environment. However, a number of problems, including lack of accessibility, poor spatial and/or temporal precision, underreporting, and misclassification, could impact crash data. Additionally, crashes are relatively rare events, so data must be collected for several years and/or in several different locations to obtain enough data to apply certain statistical models. The factors affecting crash frequency or crash severity can be classified as follows:

- Geometric features
- Traffic control features
- Vehicle features
- Traffic volume
- Driver features
- Built environment

Table 2.3 shows existing statistical models for crash severity prediction and table 2.4 shows statistical models for prediction of crash frequency.

Table 2. 3 Existing models for analyzing crash severity(Fred Mannering et al,2010)

Model type	Model specification
Binary response models	Binary logit model
	Random parameters binary logit model
Ordered probability models	Ordered logit model
	Ordered probit model
	Random effects ordered logit model
Unordered discrete models	Multinomial logit model
Continuous dependent variable	Nonlinear regression model

Table 2. 4 Existing models for analyzing crash frequency(Fred Mannering.,2010)

Model type	Advantages	Disadvantages
Poisson	Most basic model	Can't handle over dispersion and under dispersion: negatively influenced by low sample mean and small sample size bias
Negative binomial	Easy to estimate, can account for over dispersion	Can't handle under dispersion ; can be adversely influenced by low sample mean and small sample size bias
Poisson –lognormal	More flexible than Poisson gamma to handle over dispersion	Can't handle under dispersion; can be adversely influenced by low sample mean and small size bias; can't estimate a varying dispersion parameter
Zero inflated poisson and negative binomial	Handles datasets that have a large number of zero crash observations	Can create theoretical inconsistencies; zero inflated negative binomial can be adversely influenced by low sample mean and small sample size bias
Conway-Maxwell –Poisson	Can handle under and over dispersion or combination of both using a variable dispersion parameter	Could be negatively influenced by low sample mean and small sample size bias; no multivariate extensions available to data
Gamma	Can handle under dispersed	Dual state model with one state having a long term

	data	mean equal to zero
Generalized estimating equation models	Can handle temporal correlation	May need to determine or evaluate the type of temporal correlation a priori; results sensitive to missing values
Generalized additive models	More flexible than traditional generalized estimating equation models: allows non linear variable interactions	Relatively complex to implement; may not be easily transferable to other datasets
Random effects model	Handles temporal and spatial correlation	May not be easily transferable to other datasets
Negative multinomial	Can account for over dispersion and serial correlation; panel count data	Can’t handle under dispersion; can be adversely influenced by low sample mean and small sample size bias
Random parameters model	More flexible than traditional fixed parameter models in accounting for unobserved heterogeneity	Complex estimation process; may not be easily transferable to other datasets
Bivariate/Multivariate models	Can model different crash types simultaneously; more flexible functional form than generalized estimating equation models	Complex estimation process; may not be easily transferable to other datasets
Finite mixture /Markov switching	Can be used for analyzing sources of dispersion in the data	Complex estimation process; requires formulation of correlation matrix
Duration models	By considering the time between crashes allows for a	Complex estimation process; may not be easily

	very in depth analysis of data and duration effects	transferable to other datasets
Hierarchical /Multilevel models	Can handle temporal spatial and other correlations among groups of observations	Requires more detailed data than traditional crash frequency models; time varying explanatory variables are difficult to handle
Neural network, Bayesian neural network and support vector machine	Non parametric approach does not require an assumption about distribution of data; flexible functional form; usually provide better statistical fit than traditional parametric models	May not be easily transferable to other datasets; correlation results can be difficult to interpret

2.2 ANALYTICAL METHOD BASED ON CONFLICTS

Given that they occur far more frequently than crashes, traffic disputes are easier to measure. The most intriguing application for identifying hazardous conditions with a likelihood of collision is the traffic conflict technique. Using the traffic conflict technique, safety analysts may quickly spot and assess risky driving behaviors on the road while also examining how these behaviors relate to specific road features. A traffic conflict occurs when two or more road users are visibly approaching one another in space and time to the point that a collision is imminent if their motions are unaltered. Since traffic disputes only include dangerous contacts between vehicles, they might offer a more useful exposure measure for collisions. In order to employ traffic conflicts as a substitute for collisions in safety analysis, the relationship between collisions and conflicts must first be established. Sayed et al.,2016 developed conflict based safety performance functions to predict traffic collisions by type. They used Poisson gamma model for conflict and collision prediction model. The conflict prediction model developed was as follows:

$$E(AHC) = b_0 * PEV^{b1} * \exp(b_2 * AREA)$$

$$PEV = (AHV_1 * AHV_2)^{1/2}$$

- AHC = Average hourly conflicts
- AHV_1 = Major volume
- AHV $_2$= Minor volume
- AREA TYPE = (1=Urban area; 0=Suburban area)
- b_i = Conflict model coefficient

Collision prediction model formed was as follows:

$E(Y) = c_0 \, AHC^{c1}$

The result of collision prediction model obtained was shown in table 2.5

Table 2. 5 Parameter estimates for collision prediction model(Sayed et al,2016)

Parameter	Mean
$\ln(b_0)$	-3.033
b_1	0.857
b_2	0.618
$K_{conflict}$	7.428
$\ln(c_0)$	-0.992
C_1	0.874
$K_{collision}$	4.831

2.3 MICROSIMULATION MODELS FOR SAFETY EVALUATION

An intersection's safety performance is often assessed using a statistical analysis of the crash data gathered over a period of time both before and after the improvements have been made to the intersection. However, because damage-only crashes are more likely to go undetected, the data gathered could not be accurate. Thus, there may be a reliability problem with the conclusions drawn from such crash data. Similar to this, it takes a long time to gather and document crash reports, making it impossible to put the remedy needed to decrease potential crashes into action sooner. The fact that collisions happen right away before one can assess a location's risk is also problematic from an ethical standpoint. Traffic congestion has been demonstrated to be a reliable substitute for accidents. As a result, conflict data may be used in place of crash data. As a result, the method of calculating crashes based on conflicts has been used. The Federal Highway Administration's (FHWA's) Surrogate Safety Assessment Model (SSAM) and microsimulation are used to identify the conflict points (FHWA). The Surrogate Safety Assessment Model (SSAM), developed by the Federal Highway Administration (FHWA), is used to do a statistical analysis of the traffic flow data obtained from the microscopic traffic numerical simulation. SSAM computes protective surrogate steps by reading trajectory files produced by simulation software. The program's method and logic were developed based on Gettman and Head's 2003 research. By using this process, conflict analysis can be done objectively rather than subjectively. When compared to other traditional safety analysis approaches that need a lengthy wait for data collection, using simulation software and SSAM takes less time. SSAM is compatible with the following traffic simulation softwares:

- AIMSUN.
- Paramics.
- TEXAS.
- VISSIM.

SSAM classifies conflicts according to the conflict angle that SSAM extracts from path files constructed using the VISSIM for each pair of conflicting vehicles. Number and severity of conflicts are determined by SSAM. There are three categories for the conflict :(crossing, rear end, and lane change). The advantages of SSAM are outlined below:

- Compatible with as many traffic simulation models as possible;
- Provide tools for traffic engineers to do flexible safety analyses.
- Employ the best substitute safety measures you can (i.e., most representative of crash propensity).

SSAM provides statistical information (min, max, mean, and variance) of SSAM measures(TTC, PET, MaxS, DeltaS, DR, MaxD, Max DeltaV). The time before a collision, or time-to-collision (TTC), occurs if both vehicles continue travelling in the same direction and at the same speed. PET is the shortest post-encroachment time during the conflict. The interval between the first vehicle's last time at a position and the second vehicle's subsequent arrival at that same position is known as the post-encroachment period. A real collision is indicated by a value of 0 of PET. MaxS is either vehicle's top speed during the conflict (i.e.

while the TTC is less than the specified threshold). SSAM software, depending on classifying conflicts on the angle of conflict, follows.

- | conflict angle | > 85 ° the conflict is classified Crossing.
- | conflict angle | < 30 ° the conflict is classified Rear-end.
- 30 ° ≤ | conflict angle | ≤ 85 ° the conflict is classified Lane-change.
- Conflict angle unknown the conflict is considered Unclassified.

The methodology used to determine conflicts using SSAM as follows:

- Modeling intersection in microsimulation models like VISSIM.
- Running the simulation model
- Save simulated pathfiles in trj format
- Analyzing trj files in SSAM
- SSAM provides statistical values of SSAM measures

Abed et al,2021, did an assessment of safety at signalized intersections in Hilla city's urban areas using a micro-simulation model (VISSM) coupled with the Surrogate Safety Assessment Model (SSAM). Three signalized intersections were modeled using the micro-simulation VISSM (version 10) model by calculating the traffic flows and speeds extracted from field data. Also, geometric characteristics and timing of the signal were simulated to reach the real-world. Then the vehicle trajectory files were exported to SSAM (version 3). Several indicators for traffic conflicts were computed by SSAM, involving the max speeding (Max S), the rate of deceleration (DR), the time of post encroachment (PET), and time to collision (TTC). The number, type, and severity of conflicts were calculated. The conflicts were categorized into three types according to the conflict's angle, rear-end, lane change, and crossing conflicts.

Rear-end conflicts prevailed at all intersections until they reached 60% of the total conflicts at Bab- Al- Hussein intersection. The severity of conflicts for all intersection approaches ranged from 0.74, indicating a high risk, even up to 1.8, which refers to low risk. The intersections of 40 St. and Bab Al Hussein were classified as high-risk intersections, while the Bab Al-Mashed intersection was classified as a moderate-risk intersection.

Acharya Aet al,2020, predicted traffic conflicts of New Baneshar intersection, which is a signalized intersection. Plots between conflicts determined by SSAM observed conflicts are shown in figure 2.1, 2.2 and 2.3.The TTC value was analyzed for each interaction between vehicle and vehicle in the SSAM and is recorded as a traffic conflict if it is less than the threshold.

Using traffic simulation models for safety analysis can have several advantages. First, using simulated traffic conflicts overcomes quality and quantity limitations of crash and collision data. Second, simulation models and SSAM can be used to estimate simulated conflicts easily without observing them in the field. Third, traffic simulation is the easy way to explore new designs and innovative applications before implementing them in the real world.

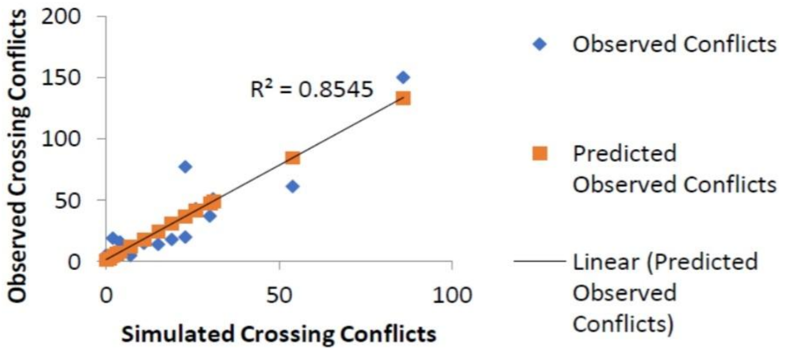

Figure 2. 1 Traffic conflicts for 5:00 -7:00 AM(Acharya et al, 2020)

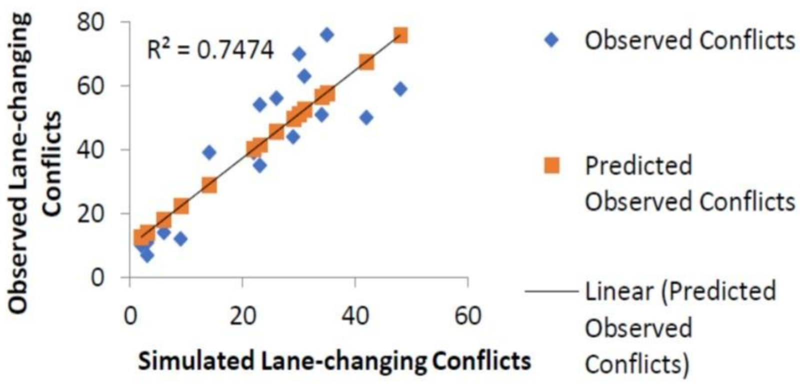

Figure 2. 2 Traffic conflicts for 9:00-11:00 AM(Acharya et al,2020)

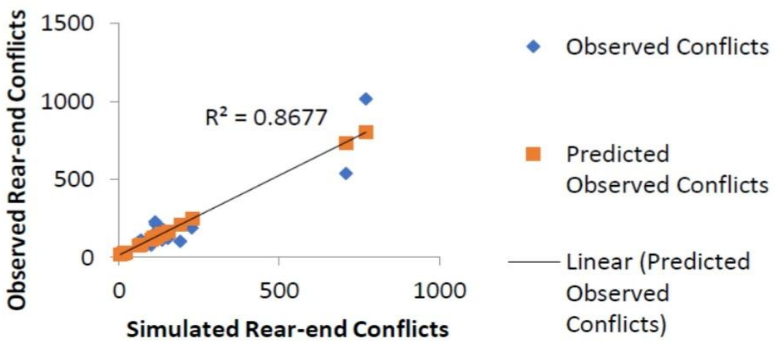

Figure 2. 3 Traffic conflicts for 1:00-3:00 PM(Acharya et al,2020)

The use of VISSIM, along with the SSAM is one of the fast way to investigate the safety impact of proposed solutions in both homogenous traffic and heterogeneous traffic.

2.4 REAL SAFETY MODELS

The use of traffic simulation models in safety studies has drawn several criticisms. First off, it was once assumed that drivers would act safely when creating the simulation models. In order to create a crash-free environment, vehicles in simulation models adhere to a set of rules (such as car-following models, gap acceptance criteria, and lane-changing behavior). It may produce unreliable results. Using these safe-moving vehicles to examine conflicts and near-misses may cause inaccurate results. Second, there are a lot of input variables in every simulation model. Any change in these parameters will cause significant changes on road user behavior in the simulation model and on the estimated number of traffic conflicts. Third, simulation models often use a variety of assumptions and methods to predict traffic (e.g., priority rules, conflicts areas, traffic distribution). Depending on the modeling strategy employed, the outcomes of simulated conflicts can differ dramatically. The failure to follow a priority rule, an abrupt lane shift by a vehicle in an intersection or while queuing, or uneven queuing up at left/right turn bay taper are some factors that often result in unrealistic crashes and abnormal movements in traffic simulations. Additionally, it was demonstrated that employing traffic simulation models without a sufficient calibration could lead to inaccurate results. However, despite ongoing research on calibrating simulation models for safety analysis, there are still a number of significant problems with using simulated traffic conflicts. The first problem is how difficult it is to calibrate the simulation model. The calibration process is time-consuming and typically necessitates the collection of genuine conflict/crash data, to obtain accurate conflict outcomes. Second, many simulation model parameters have a significant impact on the outcomes of simulated conflicts. The analysis primarily produces a wide range of estimated number of conflicts, which has poor conflict prediction and lower reliability. The third problem relates to the noted differences between simulated and real-world conflicts. Real-time safety models generally relate various traffic parameters to the risk of collision (i.e., crash potential) or to the number of traffic conflicts. Real safety models link traffic variables to the frequency of rear-end collisions at signal cycle level. The cycle-related traffic parameters (figure 2.1) include traffic volume (V), shock wave area (A), shock wave speed (S12), queue length (Q), and platoon ratio (P).

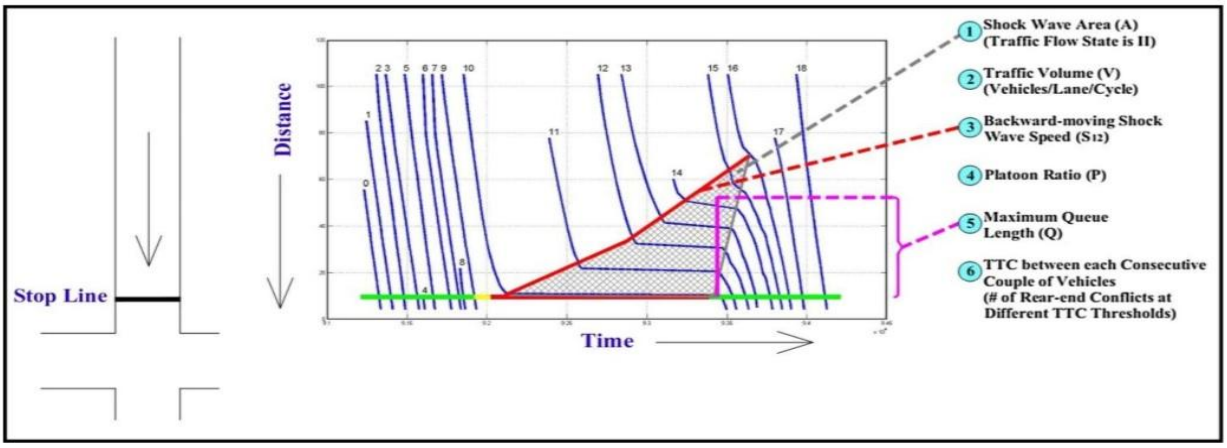

Figure 2. 4 Cycle related traffic parameters for real safety models(Essa et al,2020)

The real safety models were developed by Essa et al,2018. Generalized linear model (GLM) was used to model traffic conflict occurrence. They developed conflict-based safety performance functions (SPFs) for signalized intersections at the signal cycle level. Traffic video-data of six signalized intersections located in two cities in Canada was recorded. A video analysis procedure was done to collect rear-end conflicts and traffic variables at each signal cycle from the recorded videos. GLM used was as follows:

$E(Y) = V^{a1} \exp[\, a_0 + \sum b_j x_j \,]$

E(Y) = The preicted number of rear end conflicts per cycle

V = The traffic volume per lane per cycle

x_j = Any other explanatory variables (such as A,Q,S or P)

a_0 , a_1 , b_j : Th model parameters

Different models were created by combining the explanatory variables in various ways (V, A, Q, S12, and P). Numerous models have been developed to examine the effects of different explanatory variables and to make the suggested technique adaptable in diverse conditions where the ability to measure or estimate some explanatory variables is constrained. Developed models were shown in table 2.6.

Table 2. 6 Conflict based SPFs at cycle level (Essa et al,2018)

Model E(Y)	Variables
One variable (Exposure only) Model 1: $V^{1.563}\exp(-3.231)$	V
(Exposure+ One variable) Model 2: $V^{0.706}\exp(-1.797+0.501A)$	V,A
Model3: $V^{0.65}\exp(-2.046+0.0122Q)$	V,Q
Model4: $V^{1.637}\exp(-3.316+0.05S_{12})$	V,S_{12}
Model5: $V^{1.571}\exp(-1.768-1.266P)$	V,P
Combined model Model6: $V^{1.239}\exp(-1.624+0.294A-0.828P+0.119S_{12})$	V,A,P,S_{12}

The shock wave area, the maximum queue length, the shock wave speed, and the platoon ratio are important characteristics that influence the number of rear-end conflicts at the signal cycle. Using one of these factors or a combination of them, along with the traffic volume, in the conflict-based SPFs of signalized intersections is advised to improve the model fit and provide a better prediction of the number of conflicts beyond what can be expected from the traffic volume only. But these models have a number of drawbacks. These include:

- Validation of the developed SPFs using additional datasets
- Formation of advanced models that account for the unobserved heterogeneity across different sites using other statistical approaches such as the Full Bayes approach
- Incorporation of other conflict indicators like the PET, collision probability, safe stopping distance, among others and development of a safety index that comprises several conflict indicators to better reflect conflict severity. Also, more work is required to understand the relationship between collisions and conflicts.

2.5 SUMMARY OF LITERATURE REVIEW

Most of the accidents occurring at signalized intersections in Indian cities result in serious injuries and fatalities. Cars, buses, trucks, motorbikes, pedestrians, taxis, and other types of travelers all share the roads around the world, which helps nations' economic and social development. Nevertheless, millions of deaths and injuries occur every year as a result of collisions involving numerous cars. One of the simplest ways to assess the safety of junctions is to analyze crash data analytically. To create the regulations to prevent crashes and their negative impacts, traffic engineers must first identify the elements that affect crashes at crossings. Statistical approach to develop crash frequency and severity is easier when compared to other types of approaches to evaluate safety. The coefficients of parameters developed using prediction models help to understand the influence of each parameters on the crash easily. But limitation there is some limitations for the statistical approach based on crash data. There is a need to collect crash data for several years which is a time consuming process. The crash data may not contain damage only crashes. The development of traffic simulation models using software like SSAM helps to overcome this limitation. SSAM provides measures like PET, TTC etc which can be used to analyze collisions. But there is several limitations for traffic simulation approach. The first problem is how difficult it is to calibrate the simulation model. The calibration process is time-consuming and typically necessitates the collection of genuine conflict/crash data, to obtain accurate conflict outcomes. Second, many simulation model parameters have a significant impact on the outcomes of simulated conflicts. The analysis primarily produces a wide range of estimated number of conflicts, which has poor conflict prediction and lower reliability. The third problem relates to the noted differences between simulated and real-world conflicts. The real safety models are another approach for safety evaluation of intersections. These models link traffic variables to the frequency of rear-end collisions at signal cycle level. The rear end conflicts can be easily predicted using signal cycle parameters. There are mainly six types of models. One can use the suitable model based on the ability to measure variable. If all variables can be measured, then combined model can be used for the prediction. If one can measure only exposure variable and other one variable the models developed using exposure variable and one variable can be used. But this approach also has some demerits like validation of the developed SPFs using additional datasets, formation of advanced models that account for the unobserved heterogeneity across different sites using other statistical approaches ,incorporation of other conflict indicators like the PET, collision probability, safe stopping distance, among others and development of a safety index that comprises several conflict indicators to better reflect conflict severity. Also, more work is required to understand the relationship between collisions and conflicts.

3. METHODOLOGY

Detailed literature review has done and statistical analysis approach is selected to analyze safety of signalized intersections of Kollam corporation. Based on literature review some factors are selected to analyze severity and frequency of crashes. Traffic control features and geometrical features collected by visiting sites. Traffic volume data is also collected. Analysis is done using collected data. Correlation analysis is done to know presence of multcollinearity. After selecting factors prediction models were developed. Figure 3.1 shows the research methodology adopted for the study

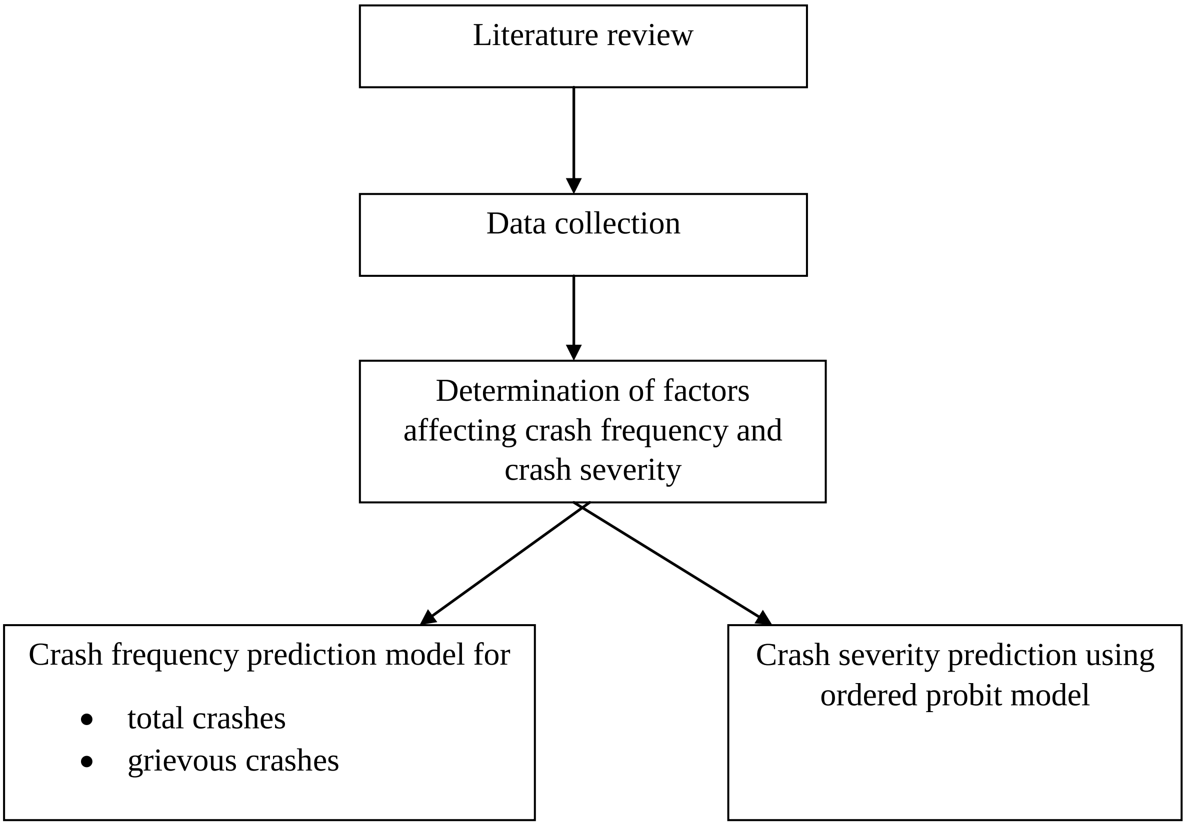

Figure 3. 1 Research methodology

3.1 STUDY OF LITERATURE

- The signalized intersection evaluation studies conducted at various cities of India is analyzed to understand the factors to be considered for the prediction. Statistical approach which uses crash data is one of the easiest methods to evaluate signalized intersection safety. Statistical approach helps to develop prediction models and also the influence of each variable on crash can analyzed easily.

- Based on literature review, it is found that there are different statistical approaches to model count data. To develop severity model in form of severity levels ordered probit model is suitable.

3.2 SELECTION OF STUDY AREA

The study area selected is Kollam corporation. Figure 3.2 shows the map of Kollam cooperation. There are totally ten signalized intersections in this study area. These are:

- Mevaram
- Ayathil
- Kallumthazham
- Kavanadu
- Kadavoor
- Kadappakada
- Taluk office junction
- High school junction
- Chinnakada
- Kappalandimukku

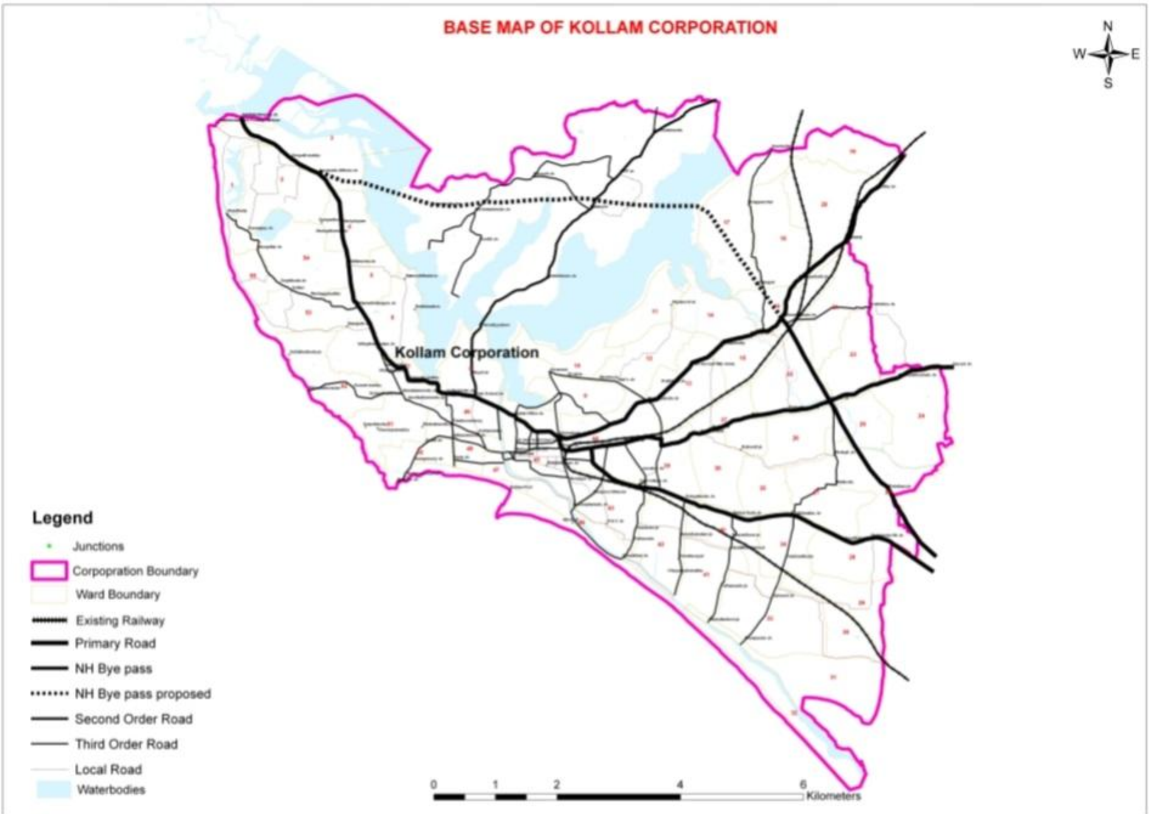

Figure 3. 2 Map of Kollam corporation (Kollam corporation.gov.in)

3.3 DATA COLLECTION

Data related to crashes, traffic control features and geometrical features were collected.

3.3.1 Geometrical Features

The geometric features considered for the study are as follows:

- Type of road (RT)
    a. Bypass
    b. Other
- Type of intersection(IN)
    a. Cross intersection
    b. Y intersection
    c. T intersection
    d. Roundabout
- Presence of footpath(F)
- Presence of zebra crossing(Z)

- Clearly visible road mark(RM)
- Presence of stop mark(S)
- Parking on carriageway(PA)
- Road category(RC)
    1. Major lane
    2. Minor lane
- Presence of exclusive left lane(L)

Geometric data of each arm of the intersections are collected. Table 3.1 shows a sample data of geometric data of each arm of one of the intersection. The value one correspond to a factor represents it's presence at the intersection and value zero represents the absence of the corresponding factor. The geometric features of each arm of all intersections are collected in same way and the data is shown in table A1 of appendix.

Table 3. 1 Geometrical features of Taluk junction

Location	Arm	Factors							
		RT	F	Z	RM	S	RC	L	PA
Taluk junction	Vellayittambalam	Other	0	1	0	0	Major	1	1
	KSRTC	Other	1	0	0	0	Minor	0	1
	Chinnakada	Other	1	0	0	0	Major	0	1
	Hospital road	Other	0	0	0	0	Minor	0	0

3.3.2 Traffic Control Features

Table 3.2 shows traffic control features of the signalized intersections of the study area.

Table 3. 2 Traffic control features

Location	No of phase	Surveillance camera	Countdown timer	Type of intersection	Cycle time(Seconds)
Mevaram	3	0	1	Y	118
Ayathil	4	1	1	Cross	119
Kallumthazham	4	0	1	Cross	126
Kadappakada	4	1	0	Cross	115
Chinnakada	4	1	1	Roundabout	137
Taluk office jn	4	0	0	Cross	120

High school n	5	1	0	Cross	120
Kappalandimukku	3	0	0	Cross	114
Kavanadu	3	1	1	T	118
Kadavoor	4	1	1	Cross	117

3.3.3 Crash Data

Crash data of six years (2016- 2021) of the selected intersections were collected for analyzing crash frequency of the signalized intersections. Detailed crash data of four years (2018-2021) is collected for analysis of crash severity. Four year crash data includes details regarding total number of accidents, date, time of accident, accident type(Fatal, grievous, minor injury), details of victims, gender of driver, age of driver, Involved persons(passenger, rider, driver, pillion rider), type of area(urban, rural), accident spot (near office complex, commercial area, institutional area, open area), type of road(other roads, bypass), type of collision , traffic violation(over speeding etc).

3.3.4 Traffic Volume Data

24 hour of traffic volumes of the ten signalized intersections are collected. From the collected volume data total volume, volume of car and volume of two wheelers are identified for each arms of the intersections. A sample data of traffic volume of each arm of the intersection is shown in the table 3.3. Collected traffic volume data is shown in table A2 of appendix. 24 hour volume, volume of two wheelers and volume of car are determined and considered as factors in the analysis. When compared to other type of vehicles the number of two wheelers and cars coming to road is higher. Two wheelers are the type of vehicle which contributes higher percentage of accident.

Table 3. 3 Traffic volume of Taluk junction

Location	Arm	Total traffic volume	Volume of car	Volume of two wheeler
Taluk junction	Vellayittambalam	11520	2961	3395
	KSRTC	10902	3065	4301
	Chinnakada	10765	2960	3781
	Hospital road	3811	972	1587

4. DATA ANALYSIS

4.1 CRASH FREQUENCY PREDICTION MODEL

Data is analyzed to develop two prediction models for crash frequency. One prediction model for total crashes and other for grievous crashes is developed. The factors analyzed to develop models are listed below.

1) Traffic volume (V)
2) Volume of two wheelers (TW)
3) Volume of car (Car)
4) Presence of footpath (F)
5) Presence of zebra crossing (Z)
6) Cycletime (CT)
7) Number of phase (P)
8) Presence of stop mark (S)
9) Presence of road mark (RM)
10) Presence of camera (CA)
11) Presence of countdown (CO)
12) Road category (RC)
13) Parking on carriageway (PA)
14) Presence of exclusive left lane (L)
15) Type of road (RT)
16) Type of intersection (IN)

4.1.1 Total Crash Frequency Prediction Model (TCFP)

Eighteen variables collected from the selected intersections were analyzed for the study to develop prediction model for total crashes. Correlation analysis is done to know presence of multicollinearity. If the independent variables are highly correlated each other it will results in multicollinearity. If we include two highly correlated independent variables in a model at a time it may leads to error in the estimated parameter coefficients. The table shows result of correlation analysis. The independent variables which are highly correlated each other and their correlation coefficient is shown in the table 4.1

Table 4. 1 Highly correlated independent variables TCFP

Independent variables	Correlation coefficient
Traffic volume & volume of car	0.843
Traffic volume & two wheeler	0.931
Volume of car & two wheeler	0.837
Cycle time & Type of intersection	-0.672
Number of phase & road mark	-0.882
Number of phase & type of intersection	-0.837
Zebra crossing & stop mark	0.728
Stop mark & road mark	-0.835
Stop mark & type of road	-0.672
Stop mark & left lane	0.674
Road mark & countdown	0.617
Road mark& left lane	0.672
Road mark & type of road	-0.822
Road mark & Type of intersection	0.739

Among the highly correlated variables the variables which are highly correlated with dependent variable is selected for model formation and other variables are neglected from the analysis.

Different models are used to develop the best prediction model. Prediction model using different models are developed. The different models which are used to form prediction model includes Poisson regression (Poisson), Negative binomial regression (NB), Zero inflated poisson regression (ZIP), Zero inflated negative binomial regression (ZINB), Zero truncated poisson model (ZTP) and Zero truncated negative binomial regression (ZTNB). Six types of regression models are used to form prediction model and the best model is selected based on goodness of fit statistics of the regression models. Table 4.2 shows the goodness of fit statistics for the six types of regression models.

Table 4. 2 Goodness of fit statistics of regression models in TCFP

Model	AIC	AICC	BIC
ZTP	118.4	124.6	133.3
ZTNB	118.7	126.6	135.4
NB	259.74	261.25	290.23
ZIP	266.84	271.84	321.74
ZINB	267.87	273.46	325.82
Poisson	273.9	275.13	301.34

From the goodness of fit statistics it is found that ZTP model fit well the data. AIC, AICC, BIC values of other regression models are higher than that of ZTP. So ZTP is selected as the best model. The results of ZTP model is shown in the tables 4.3, 4.4 and 4.5. Also the fig 4.1 shows that there is presence of excess zeros which leads to regression models like zero inflation model and zero truncated models which can deal with excess zeros.

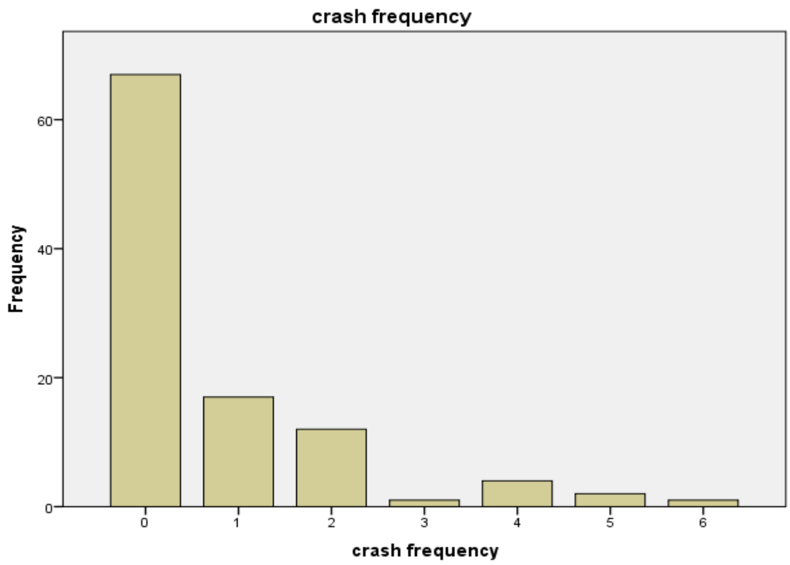

Figure 4. 1 Frequency of total number of crashes

Table 4. 3 Model information of ZTP model of total crash frequency prediction

Model information	
Type of model	Homogeneous regression mixture
Link function	Log
Distribution	Truncated Poisson
Estimation method	Maximum Likelihood

Table 4. 4 Fit statistics of ZTP model of total crash frequency prediction

Fit statistics	
-2 Log likelihood	100.4
AIC	118.4
AICC	124.6
BIC	133.3
Pearson statistic	43.81

Table 4. 5 Parameter estimates of total crash frequency prediction model

Effect	Estimate
Intercept	-5.4156
Volume of two wheelers	0.000136
Presence of zebra crossing	-0.7489
Presence of footpath	-1.6858
Presence of camera	-0.7062
Type of lane :Major lane (Reference : Minor lane)	0.7509
Presence of parking on carriageway	0.1003

Type of road :Bypass (Reference : Other)	0.2481
Number of phase	1.5915

The prediction model developed for total crashes can be written as follows:

Log (TCF/year/arm) = -5.4156 + 0.000136* TW -0.7489 Z-1.6858 F-0.7062 CA + 0.7509 major lane +0.1003 PA +0.2481 Bypass+1.5915 P

TCF =Total crash frequency

Result of ZTP shows that volume of car, parking on carriageway and number of phase increases the total crash frequency. If the road category is major road that may cause increase in total crash frequency. If the type of road is bypass there is a chance for increase in crash frequency. The factors like presence of zebra crossing, footpath, camera helps to reduce the total crashes. RMSE values determined for each types regression models are shown in the table 4.6 which indicates the ZTP model produce less error in prediction.

Table 4. 6 RMSE values of regression models in TCFP

Model	RMSE
ZTP	1.366
NB	1.731
ZIP	1.802
ZINB	2.185
Poisson	2.597
ZTNB	1.439

4.1.2 Grievous Crash Frequency Prediction Model (GCFP)

The crashes can be divided into minor, grievous and fatal crashes. Minor and fatal crashes occur rarely. But grievous occur frequently. So it is necessary to understand the factors affecting grievous crashes. The variables used to analyze grievous crash frequency are same as that of total crash frequency. Correlation analysis is done to check presence of multicollinearity. The result showed presence of multicollinearity. The independent variables which are highly correlated each other are shown in the table 4.7

Table 4. 7 Highly correlated independent variables in GCFP

Independent variables	Correlation coefficient
Traffic volume & volume of car	0.843
Traffic volume & two wheeler	0.931
Volume of car & two wheeler	0.837
Cycle time & Type of intersection	-0.672
Number of phase & road mark	-0.822
Number of phase & stop mark	-0.64
Number of phase & type of intersection	-0.837
Zebra crossing & stop mark	0.728
Stop mark & road mark	-0.778
Stop mark & type of road	-0.64
Stop mark & left lane	0.613
Road mark & countdown	0.617
Road mark& left lane	0.72
Road mark & type of road	-0.822
Road mark & Type of intersection	0.739

Among highly correlated variables the variables which are highly correlated with grievous crash frequency is retained and other variables neglected from the analysis. For grievous model too, different regression models were tested to find the best model which can predict grievous crash frequency. Six types of regression models like ZTP, ZTNB, ZIP, ZINB, Poisson and NB are used to find the best model which fit well the data. The goodness of fit statistics of the regression models are shown in the table 4.8.

Table 4. 8 Goodness of fit statistics of regression models of GCFP

Model	AIC	AICC	BIC
ZTP	89.82	98	103
NB	228.32	229.84	258.82
ZIP	228.91	233.9	283.8
ZINB	230.91	236.49	288.85
Poisson	240.81	242.05	289.26
ZTNB	91.79	102.3	106.5

AIC value is less for ZTP model when compared to all other models. So ZTP model can be considered as the best model for the prediction of grievous crash frequency prediction model. The result of ZTP model of grievous crash frequency prediction is shown in the table. 4.9 4.10 and 4.11.The fig 4.2 shows the frequency of each observation.

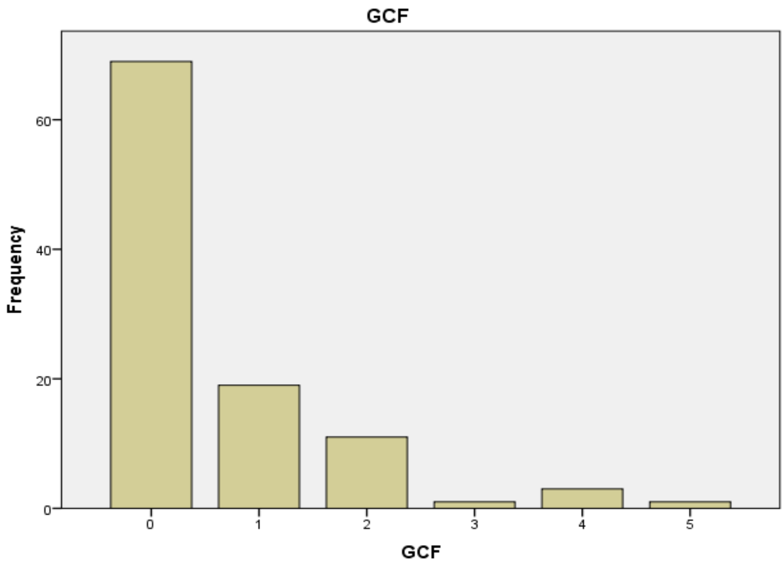

Figure 4. 2 Frequency of number of grievous crashes

Table 4. 9 Model information of ZTP model for grievous crash frequency prediction

Model information	
Type of model	Homogeneous regression mixture
Link function	Log
Distribution	Truncated Poisson
Estimation method	Maximum Likelihood

Table 4. 10 Goodness of fit statistics of ZTP model for grievous crash frequency prediction

Fit statistics	
-2 Log likelihood	71.82
AIC	89.82
AICC	98
BIC	103
Pearson statistic	29.26

Table 4. 11 Parameter estimates of grievous crash frequency prediction model

Effect	Estimate
Intercept	-5.099
Volume of two wheelers	0.000113
Presence of zebra crossing	-1.4475
Presence of footpath	-1.8431
Presence of camera	-0.7734
Type of lane :Major lane Reference : Minor lane	0.9307

Presence of parking on carriageway	-0.3541
Type of road : Bypass (Reference : Other)	0.3754
Number of phase	1.6431

Table 4. 12 RMSE values of regression models

Model	RMSE
ZTP	1.669
NB	1.933
ZIP	2.554
ZINB	3.861
Poisson	4.445
ZTNB	1.776

The result of ZTP indicates that volume of car, number of phase increases the crash frequency. Grievous crashes in major road are more compared that of minor road. Grievous crashes in bypass will be more when compared to other type of roads. Presence of zebra crossing, footpath and camera reduces grievous crash frequency. Presence of parking on carriageway increase total crashes but reduces grievous crashes. When there is parking on carriageway the moving vehicles on the road will be cautious. So the drivers may try to reduce the speed due to presence of parking which cause reduction in the occurrence of grievous crashes. But other crashes like minor or fatal crashes may occur because of the parking. RMSE values determined for each regression model is shown in the table 4.12. RMSE value is less for ZTP model which confirms goodness of fit of the model. The prediction model for grievous crashes is as follows:

Log(GCF/year/arm)=-5.099+ 0.000113 TW-1.4475 Z-1.8431F-0.7734CA+0.9307 major lane

-0.3541 PA +0.3754 Bypass +1.6431 P

4.2 CRASH SEVERITY PREDICTION MODEL

Crash severity is a important factor to be considered in accident studies. The factors which cause increase in crash frequency, but may impart less severity. So the factors which affect severity of crashes should be studied for the development of efficient policies to increase the safety. In this analysis crash severity is considered as dependent variable which is an ordinal variable. Three severity levels are considered: Minor injury level, grievous injury level, fatal. Ordered probit model is used for the formation of prediction model. The analysis is done in SAS. Ordered value and frequency of each severity levels is shown in table 4.13.

Table 4. 13 Frequency and ordered value of severity levels

Ordered value	Severity	Total frequency
1	Minor injury	8
2	Grievous injury	58
3	Fatal injury	2

The variables which are used for the analysis are:

- Gender of driver
    a. Female
    b. Male
- Hit and run
    a. Yes
    b. No
- Load category
    a. Passenger
    b. Goods
- Type of area
    a. Commercial area
    b. Near office complex
    c. Open area
    d. Institutional area
- Involved person

    a. Rider
    b. Driver
    c. Passenger
    d. Pillion rider

- Age of driver
    a. Age < 26
    b. 25 < Age < 46
    c. Age > 45
- Age of vehicle
    a. Age < 6
    b. 5 < Age < 11
    c. 10 < Age< 16
    d. Age > 15
- Type of collision
    a. Head on collision
    b. Hit from side
    c. Right turn collision
    d. Hit from back
    e. Rear end collision
    f. Skidding
    g. With parked vehicle
    h. Overturn
    i. Right angled collision
- Time of accident
    a. Night time
    b. Peak period
    c. Off peak
- Traffic violation
    a. Traffic light violation
    b. Speeding
    c. No violation
- Type of intersection

  a. Cross

b. Y
c. T
d. Roundabout

Correlation analysis is done to know the presence of multicollinearity. There are no highly correlated independent variables. So all the variables are considered for the development of prediction model. Table 4.14 shows parameter estimates of ordered probit model.

Table 4. 14 Parameter estimates of ordered probit model

Parameter	Category	Estimate
Intercept	Minor injury	5.0826
Intercept	Grievous injury	11.6751
Hit and run	Yes	-0.9119
	No	-
Gender of driver	Female	1.0743
	Male	-
Load category	Passenger	0.4170
	Goods	-
Type of area	Commercial area	-0.1086
	Near office complex	0.0453
	Open area	1.5076
	Institutional area	-
Involved person	Passenger	0.4297
	Pillion rider	-0.8197
	Rider	-0.5203
	Driver	-
Age of driver	Age < 26	-1.0579

	25 < Age < 46	-0.4224
	Age > 45	-
Age of vehicle	Age < 6	0.4699
	5 < Age < 11	0.3541
	10 < Age< 16	0.9609
	Age > 15	-
Type of collision	Head on collision	-0.3866
	Hit from side	0.1077
	Right turn collision	1.3717
	Hit from back	-0.0295
	Rear end collision	-0.1699
	Skidding	-1.157
	Overturn	-1.5741
	Right angled collision	-0.6301
	With parked vehicle	-
Time of accident	Night time	-0.4909
	Peak period	-0.6456
	Off peak	-
Type of traffic violation	Traffic light violation	-0.4461
	Speeding	-0.3921
	No violation	-
Type of intersection	Cross	-1.0959
	Y	-5.141

	T	-1.3667
	Roundabout	-

The estimates determined in such a way that probabilities modeled are cumulated over lowest ordered values. Here the lower ordered value is minor injury. So positive value represents the higher probability to get into lowest ordered value (minor injury level). Negative sign of coefficients represents lowest probability to fall under minor injury level. When consider a hit and run case the coefficient is negative which implies a hit and run case has low probability to fall under minor injury level. Hit and run accidents results in higher injury levels. When consider gender of driver the coefficient for female category is positive when reference category is male. Female drivers cause low severity like minor injury when compared to male category. It may be due to the fact that female drivers are much more cautious during driving and try to drive at low speeds. But male category creates severe accidents due to careless driving and the tendency to violate traffic rules is more among male category. If the load category is passenger it may cause lower severity levels compared to load category goods. Most of the vehicles which carry goods are heavy commercial vehicles and some of it carry dangerous goods like chemicals and also carry heavy materials. So the collision of a vehicle carrying heavy loads may create high impacts. In case of involved persons compared to driver the severity happen to passenger will be less. But severity happens to rider and pillion rider is more.

As age of driver decreases, possibility for minor injury accident reduces. It indicates younger people create more severe accidents. We know, nowadays accidents due to youngsters are more because of traffic violations like over speed, careless, lack of experience etc. Generally elder people may cause accidents because of health problems like low vision. By analyzing marginal effect it is found that elder people cause minor and grievous accidents. But they are 10.16 % less likely to be in fatal injury level. But, grievous crashes occur because of elder people, but possibility for fatal crashes is less.

In case of type of area when compared to institutional area the coefficient of commercial area is positive. Compared to institutional area the accidents occurring at open areas and office complex have a probability to fall under minor injury crashes. Crashes occur at institutional area are 32.64% more likely to be in fatal injury level and 13.22% and 19.41 % less likely to be in minor and grievous injury levels.

Compared to vehicles having age more than 15, the coefficient of other vehicles is positive which implies vehicles having less age have higher possibility to fall under minor injury level. Older vehicles are 9.84 % more likely to be in fatal injury level and 5.85% less likely to be in grievous injury level and 3.98% less likely to be in minor injury level.

Compare to collision of vehicle with fixed object the severity caused by collisions like hit from back, rear end collision, skidding ,overturn, right angled collision and head on collision results in severe accidents and are less likely to be in minor injury level.

Table 4. 15 Marginal effects of variables on each severity levels

Parameter	Marginal effect on minor injury level	Marginal effect on grievous injury level	Marginal effect on fatal injury level
Hit and run	-0.0813	-0.1193	0.2006
Female	0.0867	0.1272	-0.2139
Passenger	0.0326	0.0478	-0.0804
Commercial area	-0.1329	-0.1950	0.3279
Near office complex	-0.1264	-0.1856	0.3121
Institutional area	-0.1322	-0.1941	0.3264
Pillion rider	-0.0913	-0.1341	0.2255
Rider	-0.0582	-0.0854	0.1436
Driver	-0.0158	-0.0232	0.0391
Age of driver< 26	-0.0376	-0.0552	0.0928
Age of driver > 45	0.0412	0.0604	-0.1016
5 < Age of vehicle< 11	-0.0135	-0.0199	0.0334
10 < Age of vehicle< 16	-0.0126	-0.0185	0.0311
Age of vehicle> 15	-0.0398	-0.0585	0.0984

Head on collision	-0.0365	-0.0537	0.0902
Hit from side	-0.0041	-0.0060	0.0101
Right turn collision	0.1029	0.1511	-0.2540
Rear end collision	-0.0111	-0.0164	0.0276
Skidding	-0.0876	-0.1287	0.2164
Overturn	-0.1136	-0.1667	0.2804
Right angled collision	-0.0356	-0.0524	0.0882
With parked vehicle	-0.0058	-0.0085	0.0143
Night time	-0.0765	-0.1123	0.1889
Off peak	0.0366	0.0537	-0.0904
Traffic light violation	-0.0353	-0.0518	0.0872
Speeding	-0.0179	-0.0262	0.0441
Y intersection	-0.3123	-0.4583	0.7706
T intersection	-0.0291	-0.0427	0.0718
Roundabout	0.0867	0.1273	-0.2140

Accidents happening at night time and peak hour have less probability to be in minor injury level. Night time accidents are 18.89 % more likely to be in fatal injury level and 11.23% less possibility to result in grievous injury level and 7.65% less likely to be in minor injury level. Traffic violations like speed and traffic light violation cause severe accidents when compared to accidents happening without any violation. Accident due to traffic light violation has 8.72% more possibility to cause fatal injury.

5. CONCLUSION

Crashes occurring at intersections are more, because the conflicting movements are higher at intersections. The developed crash frequency prediction models and crash severity prediction model will help to determine future number and severity of accidents occurring at signalized intersections of Kollam corporation. The influence of factors on crashes can be also understood from these prediction models. Factors which reduce both total and grievous crashes include zebra crossing, footpath and camera. There is more chance for accidents in bypass road. Lack of sign boards near to intersections and lack of service road parallel to bypass are some of the reasons behind increased crashes at intersections of bypass road. Factors such as traffic volume and number of phase increase frequency of total and grievous accidents. Parking on carriageway increase total crashes, but reduces grievous crashes. When there is parking on carriageway the drivers will drive slowly because of the hindrance due to parked vehicles. So possibility for grievous crashes may be less, but minor injury crashes may occur, so total crashes will increase.

Crash severity is model reveals that male drivers, hit and run accidents, load category goods, traffic violations, younger drivers and old vehicles cause more severe accidents. Controlling these parameters will help to reduce severity happening during a crash. Female drivers are more cautious during driving when compared to male drivers. Male drivers have more tendency to violate traffic rules and also they try to overtake and drive fast compared to females. Younger drivers nowadays cause severe accidents due to careless driving. Use of old vehicles should be restricted by the authorities and proper maintenance of vehicles should be done by owners. Installation of crash reducing facilities like reflectors are required to reduce night time crashes. High priority should be given to intersection carrying high volume. There is a need to adopt new provisions to control the increasing two wheeler traffic. Installation of sign boards regarding speed limit near approach of intersection in bypass and service roads parallel to bypass may reduce crashes. Intersections near to commercial and institutional area should be controlled well. Installation of properly working surveillance camera is necessary to control violations at intersections. Vehicles carrying goods should reduce speed when approach to intersections.

The crash prediction will help for efficient working of signalized intersections and also help road authorities to improve the safety of intersections by understanding the depth of influence of factors which controls the crash occurrence.

REFERENCES

1) Abdullah, A.A., Jamil, S.A.M., Long, K.S., Mohamad, M. and Nor, M.E. 2017. Detecting overdispersion in count data: A zero inflated poisson regression analysis. *Journal of Physics.*
2) Anjana, S. and Ananeyulu, M. 2014. Safety analysis of urban signalized intersections under mixed traffic. *Journal of safety research*, Volume 9.
3) Aty, A.M., and Keller, J. 2004. Exploring the overall and specific crash severity levels at signalized intersections,. *Accident analysis and prevention*, 417-425.
4) Azad, A. 2017. Road crash prediction models: Different statistical modeling approaches. *Journal of transportation technologies,* Volume 7.
5) Chen, Y. and Huang, Z. 2020. Predicting wet road crashes using the finite mixture zero truncated negative binomial model,. *Journal of Advanced Transportation.*
6) Duncn, C.S., Khattak. A.J. and Council, F.M. 2018. Applying the ordered probit model to injury severity in truck passenger car rear end collisions. *Transportaion Research Record.*
7) Haghighatpour, J.P., Moayedfar, R. 2014. Pedestrian crash prediction models and validation of effective factors on their safety: A case study of Tehran signalized intersections. *Open journal of civil engineering*, (4),240-254.
8) Haque, M.E., Mallick, T.S. and Bari, W. 2021, Zero truncated poisson model: an alternative approach for analyzing count data with excess zeros, *Journal of Statistical Computation and Simulation.*
9) Hussain, T. 2020. A zero truncated discrete distribution: Theory and applications to count data. *Pakistan Journal of Statistics and Operation Research,* 16 (1).
10) Krishnan, M.J. and Anjana, S. 2014. Development hierarchical safety performance functions for urban mid-blocks, *Procedia-Social and Behavioral Sciences* ,Volume 11.
11) Milton, J.C. 1998. The relationship among highway geometrics, traffic related elements and motor vehicle accident frequencies..
12) Mitra, S. and Bhowmick, D. 2020. Status of signalized intersection safety- A case study of Kolkata. *Accident analysis and prevention.*
13) Mor, N., Sood, H. and Goyal, T. 2020. A statistical model for prediction of road accidents in the state of Punjab. *Journal of Interdisciplinary Mathematics*, 23(1).

14) Saffari, S.E., Adnan,R. and Greene, W. 2011. Handling of over dispersion of count data via truncation using poisson regression model. *Journal of Computer Science and Computational Mathematics*, 1(1).

15) Siddique, M.T. 2018, Accident severity analysis on national highways in Bangladesh using ordered probit model, *Academic journals,* 14(13).

16) Tay, R. and Shakil, M. 2007. Factors contributing to the severity of intersection crashes. *Journal of advanced Transportation* , Vol 41, 245-265.

17) Tayek, B.J.O. 2006. Development of severity models for vehicle accident injuries for signalized intersections in rural areas. KSCE Journal of Civil Engineering, vol. 3.

18) Unlu, H.K., Young, D.S., Yigiter, A. and Ozcebe, L.H. 2020. A mixture model with poisson and zero truncated poisson components to analyze road traffic accidents in Turkey, *Journal of applied statistics.*

19) Wong, C.K. 2019. Designs for safer signal controlled intersections by statistical analysis of accident data at accident black sites. *IEEE Access.*

APPENDIX

Table A1: Geometric features of intersections

Location	Arm	RC	Z	F	S	RM	PA	L	RT
Mevaram	Pallimukku	Minor	1	0	1	1	0	1	Other
	Ayathil	Major	1	0	1	1	0	1	Bypass
	Trivandrum	Major	1	0	1	1	0	1	Bypass
Kallumthazham	Kavanadu	Major	1	0	1	1	0	1	Bypass
	Chandanathoppe	Minor	1	0	0	0	0	0	Other
	Ayathil	Major	1	0	1	1	0	1	Bypass
	Chinnakada	Minor	1	0	0	0	0	0	Other
Kadappakda	Kottarakkara	Major	0	1	0	0	1	0	Other
	Kappalandimukku	Minor	1	1	0	0	0	0	Other
	Chinnakada	Major	0	1	0	0	0	1	Other
	Asramam	Minor	1	1	1	0	0	0	Other
Taluk junction	Vellayittambalam	Major	1	0	0	0	1	1	Other
	KSRTC	Minor	0	1	0	0	1	0	Other
	Chinnakada	Major	0	1	0	0	1	0	Other
	Hospital	Minor	0	0	0	0	0	0	Other
High school junction	Kavanadu	Major	0	1	0	0	1	0	Other
	Elamballor	Minor	1	0	1	0	0	0	Other
	Chinnakada	Major	1	0	0	0	0	0	Other
	Kottamukku	Minor	0	0	0	0	0	0	Other
Kavanadu	Shakthikulangara	Major	1	0	1	1	0	1	Bypass
	Kadavoor	Major	1	0	1	1	0	1	Bypass

	Vellayittambalam	Minor	1	0	1	1	0	1	Other
Chinnakada	Kottarakkara	Major	1	1	0	0	0	0	Other
	Alappuzha	Major	1	1	1	0	1	1	Other
	Asramam	Minor	0	1	0	0	0	1	Other
	Beach road	Minor	0	0	0	0	0	0	Other
Ayathil	Mevaram	Major	1	0	1	0	0	0	Bypass
	Kannanalloor	Minor	0	0	0	0	0	0	Other
	Kallumthazham	Major	1	0	1	0	1	0	Bypass
	Chemmamukku	Minor	0	0	0	0	0	0	Other
Kappalandimukku	Kadappakada	Minor	1	1	1	1	0	1	Other
	Mevaram	Major	1	1	1	1	1	0	Other
	KSRTC	Major	1	1	1	1	1	1	Other
	Railway	Minor	0	0	0	0	0	0	Other
Kadavoor	Kallumthazham	Major	1	0	1	1	0	1	Bypass
	Elamballor	Minor	1	0	1	1	0	0	Other
	Kavanadu	Major	1	0	1	1	0	1	Bypass
	High school	Minor	1	0	1	1	0	0	Other

Table A2: Traffic volume data

Location	Arm	24 hour traffic volume	Volume of car (veh/day)	Volume of two wheelers(veh/day)
Mevaram	Pallimukku	9222	2968	4443
	Ayathil	8009	3163	3285
	Trivandrum	12697	4168	6131

Kallumthazham	Kavanadu	3258	1143	1481
	Chandanathoppe	6212	1423	3501
	Ayathil	6611	1857	3505
	Chinnakada	4662	1362	2132
Kadappakda	Kottarakkara	10262	2085	5805
	Kappalandimukku	8154	2087	4192
	Chinnakada	4183	896	1931
	Asramam	5458	1400	2423
Taluk junction	Vellayittambalam	11520	2961	3395
	KSRTC	10902	3065	4301
	Chinnakada	10765	2960	3781
	Hospital	3811	972	1587
High school junction	Kavanadu	9582	2835	3852
	Elamballor	4773	1259	2072
	Chinnakada	12429	2878	6127
	Kottamukku	2789	556	1349
Kavanadu	Shakthikulangara	14086	3832	6905
	Kadavoor	4917	1808	2053
	Vellayittambalam	6941	1791	3377
Chinnakada	Kottarakkara	5357	943	2818
	Alappuzha	13493	2404	6580
	Asramam	2803	611	1289
	Beach road	5424	985	2466

Ayathil	Mevaram	10765	2976	5964
	Kannanalloor	3043	507	854
	Kallumthazham	10381	3465	5716
	Chemmamukku	1003	143	260
Kappalandimukku	Kadappakada	6883	1721	3737
	Mevaram	16255	4806	8271
	KSRTC	1399	283	701
	Railway	15696	4650	7615
Kadavoor	Kallumthazham	3287	1244	1450
	Elamballor	4863	1365	2178
	Kavanadu	5892	1599	2608
	High school	4119	720	2675