Study on the Behavior of Pedestrian Crossing at Signalized Intersection in Addis Ababa

Yemsrach Geremew, Shivakumar Mahto
Civil Engineering Dept.
Mizan Tepi University
Mizan Aman, SNNPR, Ethiopia

***Abstract*—Pedestrians are vulnerable road users therefore, it is necessary to analyze their behavior in order to define quality measures. Crossing behavior of pedestrians in Addis Ababa City is rarely in compliance with the pedestrian regulation. To study the crossing behavior of pedestrians in Addis Ababa 4 intersections namely Estifanos, Legehar, Banko de roma and post office intersections are selected. The first three intersections were used for the major study and the last one for validation study. Data collection was done using video recording method for the study of pedestrian crossing behavior and 385 respondents were interviewed to get pedestrian perception on the crossing environment. Student t-test and one way ANOVA were used to identify factors affecting pedestrian crossing speed while Mann-Whitney U test and Kruskal-Wallis one way ANOVA test were used to identify determining factors of pedestrian signal and crosswalk compliance. Gender, age group and pedestrian group size were found as a dominant factors affecting pedestrian crossing speed whereas age group and pedestrian group size are found as significant factors affecting pedestrian signal and crosswalk compliance. Multiple linear regression model is developed to study the effect of significant factors on pedestrian crossing speed. Except crossing signal phase, all variables were found statistically significant to predict pedestrian crossing speed. Gender and pedestrian group size are correlated negatively with crossing speed while pedestrian age group is correlated positively. Binary logistic regression models are also developed to study the effect of significant factors on pedestrian signal and crosswalk compliance. Pedestrian age group and group size were found statistically significant to predict pedestrian probability of signal compliance and crosswalk compliance.**

Keywords: Pedestrian, crossing behavior, crossing speed, compliance, signalized intersection.

I. INTRODUCTION

In developing countries a vital component of the transportation system is being marginalized the pedestrian, as planners, policy makers, and engineers focus on motor vehicles. Large percentages of the population in developing countries are still dependent on walking; yet do not have safe means to do so. With transportation fatalities predicted to become the 4th leading cause of early mortality globally by 2030, this is an important issue, and especially so in the developing world context. Once, pedestrians behavior in terms of safety when interacting with automobiles and safe utilization of types of signalization & infrastructure treatments is understood, we may then plan and design transportation systems that are safer for all users. (Nicholas N and Matin 2017)

Trying to particularly focus on the capital, commercial and political center of Ethiopia, the above statement holds true for Addis Ababa. The population of the city has nearly doubled every decade since the 1980s. The 2013 population size is estimated at 3.1 million and is estimated to reach 12 million in 2024. (Central Statistical Agency 2013) has a large concentration of motorized vehicles. In 2017, registered vehicles in the country have reached 831,235. Among these about 65 percent of the total vehicles operate in the city. This number is expected to be growing at a rapid pace with the development of the socio-economic activity of the city and growing purchasing power of its people. (EthiopianTranportAuthority 2017)

A considerable amount of literature has been published on how traffic operations and pedestrian safety on road ways could be improved. (Serge 2001), Documented that the complexity of traffic flow system and pedestrian behavior may not allow analytical approaches to provide the desired result. So, models designed to characterize the behavior of the complex traffic flow system have become an essential tool in analysis and experimentation.

A. Motivation

The pedestrian mode of travel is a vital but neglected component of urban transport and urban activity system. Traffic research on roadways has always been on vehicles. Although pedestrians are vulnerable road users, concerns for the safety, comfort, and convenience of them have often come secondly while designing roadways. One of the major reasons is the complexity involved in modeling pedestrian behavior. At signalized intersections, pedestrian travel is very high in highly populous cities with least amount of safety measures provided to them. Pedestrian crossing behavior is an important and relevant objective when designing roadways and signalized intersection to meet the aims such as safety, comfort and convenience. By studying and describing the real situation of crossing behavior of pedestrians by statistical model and methods will better help in achieving and improving the safety, comfort and convenience of pedestrians

B. Statement of the problem

Traffic accident on pedestrians is increasingly becoming the country's major problem especially for Addis Ababa city. Report obtained from Addis Ababa police commission department of traffic police shows traffic accident on pedestrians is increasing with high rate and level of severity (see fig.1**)**. In addition, congestion is coming to be a vital problem on pedestrian walkways and crossing facilities. However, studies concerned on traffic engineering and safety mainly focus on vehicular traffic flow and pedestrian flow behavior or pedestrian crossing behavior is not studied yet.

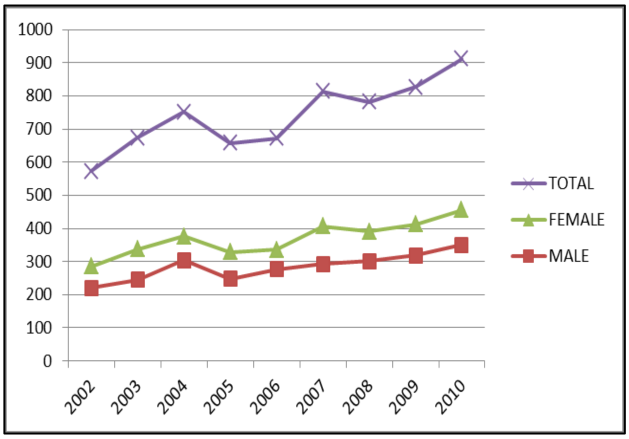

Figure 1 Fatal pedestrian crash data from 2002 up to 2010
SOURCE: Addis Ababa Police Commission: Department of Traffic police

C. Objectives of the Study

The key is to study the crossing behavior of pedestrian at signalized intersection.

- *Specific Objectives*
- To identify the factors that dominantly affect pedestrian crossing speed in the cross walk at signalized intersections
- To determine most significant factors affecting signal and crosswalk compliance & identify the reasons for pedestrian noncompliance of signal timing and crosswalk location.
- To examine the effects of pedestrian characteristics and traffic characteristics on pedestrian crossing behavior.template, modified in MS Word 2007 and saved as a "Word

II. RESEARCH METHODOLOGY

A. Description of the study area

In order to conduct this study, the researcher selected 4 intersections in Addis Ababa namely, Estifanos intersection, Legehar intersection, Banko de roma intersection and Post office intersection. Among these the first 3 are used for the major study and the last one for validation.

Location

The following are the description of location for the intersections:-

- Estifanos intersection-located around Estifanos church
- Legehar intersection- located around Legehar bus station
- Banko de roma intersection- located on Churchill avenue
- Post office intersection- located around post office on the way to Piazza from Biherawi

B. Methodology

- Data collection will be done by video recording technique and extraction will be carried manually.
- Pedestrians will be interviewed to know their attitude about the crossing environment
- ANOVA test, student t-test, Mann Whitney U test and Kruskal Wallis test will be applied.
- Multiple linear regression & Logistic regression model will be used to develop best correlating model
- The developed model will be checked if it is valid
- Result gained will be interpreted, discussed & potential recommendations will be given

The study design is based on the research objectives. After reviewing several researches in raw to see what parameters shall be captured in order to study the crossing behavior of pedestrians. In order to identify the factors that dominantly affect pedestrian crossing speed pedestrian characteristics such as gender, age group and pedestrian group size as well as traffic characteristic i.e. crossing signal phase are tested statistically for significant difference in their group mean. Here crosswalk length and crossing time are measured to compute crossing speed of each pedestrian observed in the video data.

Similarly, to determine the most significant factors affecting pedestrian signal compliance and crosswalk compliance pedestrian characteristics are captured and tested statistically. In order to identify reason of pedestrians' noncompliance of signal and crosswalk, pedestrian attitude survey is done.

To examine the effects of pedestrian characteristic and traffic characteristic on pedestrian crossing behavior, a model is developed using pedestrian characteristic and traffic characteristic parameters.

C. *Sample size analysis & determination*

- *Intersection sample determination*

There are 30 signalized intersections in the city of which 24 were found on function during the early stage of site visit. The list & details of the intersections is found at the appendix section. According to (P. Srinivas , et al. 1994)land use is a strong explanatory variable of pedestrian arrival process. Also, land use is a factor exogenous to behavior and would thus allow unbiased estimation of behavior study and model. Furthermore, pedestrian crossing behavior is partially dependent on trip purpose and land use around the intersection in general is a good determinant of pedestrians' trip purpose. The following are the figures showing the image of the selected intersections

Figure 2 Camera set up location for Estifanos intersection

Figure 3. Camera set-up location for Legehar intersection

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

Figure 4 Camera set-up location for Banko de roma intersection

Figure 5. Camera set-up location for post office intersection

- *Pedestrian sample size*

(Ann, Asha and Kevin 2012) Developed minimum sample size data needed for pedestrian and bicycle researches depending on confidence interval (CI), margin of error (E) and population size as tabulated below. The authors also stated that the values in the table were obtained by inputting values to the online calculator available http://www.surveysystem.com/sscalc.htm (Kreizek, et al. n.d.)

- *Data collection*

Primary data collection

Lenovo tablets were used to record video data of pedestrian movement at the intersections. The tablets were placed at different convenient positions so that clear view of pedestrians' movement could be provided. Data collection was conducted at the specified four intersections (Estifanos, Legehar, Banko de roma& Post office) from 29-3-2018 up to 3-5-2018 on working days, except for Monday and Friday because there exists an exaggerated traffic movement on these days, from 07:00 AM to 10:00 AM in order to capture the morning peak hour at the study locations. Video recordings were taken for each intersection for each leg. Pedestrian attitude survey data was also taken parallel to video data collection. Pedestrians using the intersections frequently (pedestrians who own shops, pedestrians that were refreshing in café and restaurants) were interviewed about the crossing environment.

- Secondary data collection

Geometric data (number of lane, length of crosswalk and crosswalk width) and signal related data (Green, Red and Yellow time) were collected at the selected study sites using the following specific methodology.

- No of lane: counting divided lane on each approach
- Crosswalk width: measuring with roller meter
- Length of crosswalk: measuring with roller meter
- Green, red & yellow time: Recording at the intersections

C. Research materials

- Data

Data required for the research is categorized under three main categories i.e.

- Pedestrian characteristics data
- Traffic characteristics data &
- Attitude survey data

Pedestrian characteristics data

Pedestrian characteristics data includes:

a) Pedestrian volume

-Total number of pedestrian as determined in article http://www.surveysystem.com/sscalc.htm (Kreizek, et al. n.d.)

b) Gender

Gender of a crossing pedestrian is identified and expressed as male or female.

c) Age group

Here the pedestrians' age groups that were crossing at the intersection are assessed. According to (Daamen and Hoogendoorn 2003) three classes are adequate within the peak hour of traffic study. i.e :< 20 years of age (juveniles), 20 to 55 years of age (middle aged) and > 55 years of age (seniors).

d) Pedestrian group size

Pedestrian group size is categorized based on the concept of level of service explained in HCM 2000. According to HCM 2000 space & flow rate (this can be determined with speed, density & volume concepts) are used to determine the quality of pedestrians movement. Category of pedestrian group size is presented as follows assuming pedestrians in the same categories will experience closer quality of movement &speed while crossing the road.

TABLE 1: CATEGORY OF PEDESTRIAN GROUP SIZE

No. of pedestrians	Code	Description
1	0	Pedestrian is expected to achieve LOS A moving in a desired path without altering their movement in response to other pedestrian
2-3	1	Pedestrian is expected to achieve LOS B or LOS C There is sufficient area for pedestrians for normal walking & by passing other pedestrians
4-5	2	Pedestrian is expected to achieve LOS D or LOS E There is restricted walking speed Movements face conflicts
>5	3	Pedestrian is expected to be in LOS F Walking speed is severely restricted Movements will be by shuffling

e) Crossing time

Duration that takes a pedestrian to cross the crosswalk is measured using stopwatch.

f) Crossing speed

Speed of a pedestrian is calculated as V=S/ t where V-crossing speed, S- Crossing distance and t-crossing time

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

III. ANALYSIS OF DATA

The research findings are presented and discussed analytically, statistically, graphically and in tabular form. In the action to do so, the chapter is divided into five parts: Identifying factors related to pedestrian crossing speed, Identifying factors related to signal and crosswalk compliance, Effect of pedestrian characteristic and traffic characteristic on pedestrian crossing behavior, Model validation and pedestrian attitude survey

A) Identifying Factors Relating to Pedestrian Crossing Speed

Before analyzing the differences in behavior the distribution of the dependent variable i.e. crossing speed is tested in order to choose an appropriate test. The descriptive statistics of the crossing speed is shown in table 2 below.

TABLE 2 DESCRIPTIVE STATISTICS FOR CROSSING SPEED

No.	Valid	3783
	Missing	0
Mean Speed		1.668
95% Confidence Interval	lower	1.646
	upper	1.691
Median		1.620
Variance		0.361
Standard Deviation		0.6009
Minimum		0.3
Maximum		3.3
Range		3

To check the speed data for normal distribution SPSS software is used. Three familiar and reliable methods were applied to the data and the following result was obtained.

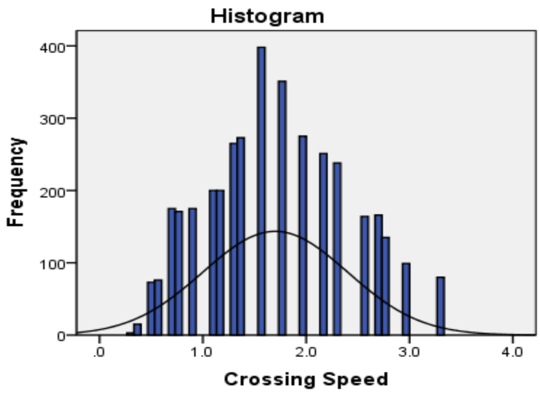

Figure 6 Histogram of crossing speed

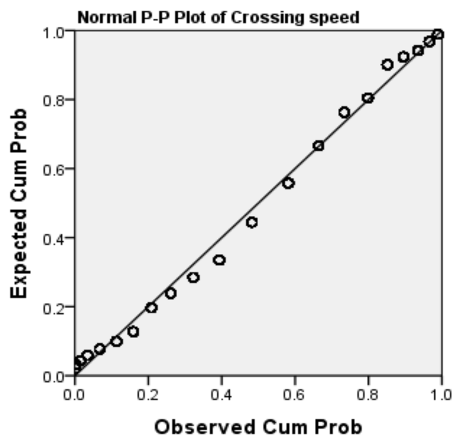

Figure 7: Normal p-p plot of crossing speed

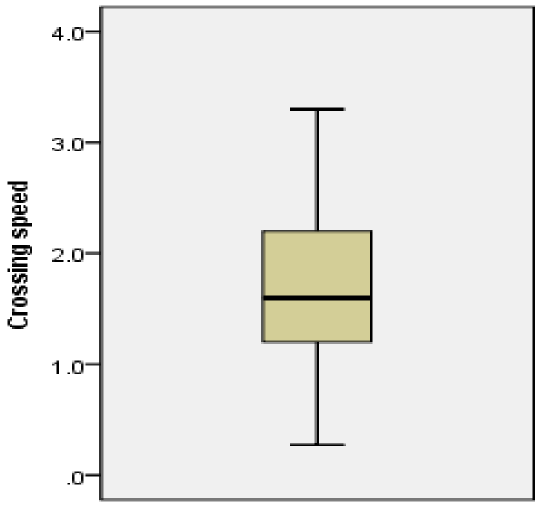

Figure 8: box plot of crossing speed

- Figure 6 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 data is from a normal distribution.
- The normal probability plot graph in figure 7 shows the pattern of the data approximately following the straight line strengthening the conclusion that the data is from normal distribution.
- Additionally, in figure- 8 the box plots shows the distribution of the data about the median has lower range and appears to be found approximately at the middle of the box that shows a balanced concentration of the data. The whiskers also show no significance different in length which again emphasizes no significant skewness.

a) Gender differences

To test if a significant difference in crossing speed exists between men and women the following hypothesis was drawn and an independent sample mean test was performed. The result is presented in table 3 below.

Ho: no significant difference in crossing speed exists between men and women

H1: significant difference in crossing speed exists between men and women

TABLE 2 INDEPENDENCE SAMPLE T-TEST FOR GENDER DIFFERENCES IN CROSSING SPEED

Independet sample t-test									
		equality of variances		T-test for equality of means					
		F	Sig.	t	df	Sig. (2- tailed)	Mean difference	95% confidence interval of the difference: Lower	Upper
Crossing speed	Equal variances assumed	0.49	0.484	3.107	3780	0.002	0.81	0.0018	0.1321
	Equal variances not assumed			3.101	1676	0.002	0.81	0.0017	0.1322

The P value expressed as "Sig. for the Levene's test (0.484) is greater than the level of significance i.e. 0.05, so that Ho is failed to be rejected and concluded that the assumption of equal variances is met.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

Moving on, the P value for the t-test expressed as Sig (2tailed)=0.002 is less than the significance level indicating that Ho can be rejected for the group differences and conclusion is drawn as, there exists a significant difference in crossing speed between men and women.

b) Age group differences

The hypothesis formulated to test if a significant difference exists in crossing speed among the age groups is:

Ho: no significant difference in crossing speed exists among age groups

H1: significant difference in crossing speed exists among age groups

The result of one way ANOVA test from SPSS is presented in table 3 below

The p-value in the table shows that it is much lower than p=0.05 hence, Ho is rejected and concluded that there is a significant difference in crossing speed among the age groups,

TABLE 3 ONE WAY ANOVA TEST FOR THE AGE GROUP DIFFERENCES

One way ANOVA test					
	Sum of squares	Df	Mean square	F	Sig
Between groups	13.042	2	6.521	13.369	0.000
Within groups	1843.251	3781	0.488		
Total	1856.256	3783			

The p-value in the table shows that it is much lower than p=0.05 hence, Ho is rejected and concluded that there is a significant difference in crossing speed among the age groups.

c) Difference in crossing pedestrians group

The following hypothesis is tested in order to know if a significant difference in crossing speed exists among pedestrian group sizes.

Ho: no significant difference in crossing speed exists among pedestrian group sizes.

H1: significant difference in crossing speed exists among pedestrian group sizes.

TABLE 4: ONE WAY ANOVA TEST RESULT FOR PEDESTRIAN GROUP SIZE DIFFERENCES

One way ANOVA test					
	Sum of squares	Df	Mean square	F	Sig
Between groups	1070.56	3	356.85	1715.84	0.000
Within groups	785.73	3780	0.208		
Total	1856.29	3783			

The P-value 0.000 which is much less than the level of significance 0.05, leads to the rejection of the null hypothesis. Therefore, at α=0.05 level of significance there exist enough evidence to conclude that a significant difference in crossing speed exists among the pedestrian group sizes.

B. Identifying Factors Relating to Signal and Crosswalk Compliance

Dependent variables crossing signal phase and crossing locations are dichotomous (discrete) variables with nominal level of measurement. As discussed in section 3.7.1.4 measurement of crossing signal was coded as 0 and 1 for pedestrians crossing on green phase and for pedestrians crossing on red phase respectively. Similarly, measurement of crossing location was coded as 0 and 1 for pedestrians crossing away from designed crosswalk and for pedestrians crossing on designed crosswalks respectively. Since, it cannot be observed continuity of a discrete data, it is impossible to test the normality (normal distribution) of it. For this reason rather than using the T-test and ANOVA it is recommended to use non parametric test which does not require an assumption of normal distribution of a data

a) Gender

To test mean differences in gender the Mann Whitney U-test in SPSS is used instead of T-test. The hypothesis formulated to test if a significant difference exists between the gender groups is stated as follows:

Ho: no significant difference in signal compliance exists between men and women

H1: significant difference in signal compliance exists between men and women

As shown in table 5 the Mann-Whitney U independence sample mean test gives a P value of 0.432 which is greater than the significance level. Depending on this result, the null hypothesis is failed to be rejected and concluded that there doesn't exist a significant difference in gender group means

TABLE 5: GENDER RANKS FOR CROSSING SIGNAL

Ranks				
	Gender	N	Mean rank	Sum of ranks
Crossing signal	Male	2814	1896.7	5335415.50
	Female	969	1876.41	1818237.50
	Total	3783		

TABLE 6 MANN-WHITNEY U T-TEST FOR DIFFERENCE IN GENDER IN CROSSING SIGNAL

Test statistics: grouping variable -Gender	
	Crossing signal
Mann Withney U	1348272.500
Wilcoxon W	1818237.500
Z	-0.785
Asymp.sig.(2tailed)	0.432

b) Age group differences

To test age group differences in crossing signal, an extension of Mann-Whitney U test i.e. Kruskal Wallis one way ANOVA test is used. The result of test of the following hypothesis is presented in table 7 and table 8

TABLE 7: AGE GROUP RANK FOR CROSSING SIGNAL

Ranks			
	Age group	N	Mean rank
Crossing signal	<20	402	1954
	20-55	2756	1888.78
	>55	625	1864.71
	Total	3783	

TABLE 8: KRUSKAL WALLIS TEST FOR DIFFERENCE IN AGE GROUP IN CROSSING SIGNAL

Test Statistics-Age group	
	Crossing signal
Chi.square	27.312
df	2
Asymp.sig.	0.006

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

c) Difference in crossing pedestrians' group size

Difference in pedestrians' group size is also tested using Kruskal Wallis one way ANOVA test. The hypothesis tested is:

Ho: no significant difference in signal compliance exists among pedestrian group size groups

H1: significant difference in signal compliance exists among pedestrian group size groups

The test result in table 9 shows the P value of 0.000 which is less than the significant level. This leads to the rejection of Ho. As a result it is concluded that, a significant difference in signal compliance exists among pedestrian group size groups.

TABLE 9: PEDESTRIAN GROUP SIZE RANKS FOR CROSSING SIGNAL

Ranks			
	Pedestrian group size	N	Mean rank
Crossing signal	1 (single)	1222	1846.32
	2–3	825	1795.67
	4–5	768	2041.84
	>5	968	1910.95
	Total	3783	

TABLE 10 KRUSKAL WALLIS TEST FOR DIFFERENCE IN PEDESTRIAN GROUP SIZE IN CROSSING SIGNAL

Test Statistics-Pedestrian group size	
	Crossing signal
Chi.square	57.75
df	3
Asymp.sig.	0.000

C. Testing Multiple Linear Regression assumptions

Before conducting the regression analysis, the six major assumptions of multiple linear regressions needs to be checked for violation. In order to get an appropriate result these assumptions of MLR have to be met.

Assumption 1: Continuous dependent variable

The assumption in regression is that the response variable is continuous that it can take on any value within a range of values. The dependent variable in the regression model is crossing speed, which is a continuous variable taking values in between 0.3 and 3.3 as shown in the descriptive statistics above.

Assumption 2: Linear relationship between dependent and each of independent variables.

The relationship between the dependent and each of independent variables should be linear. In the fig.9, fig.10, fig.11 and fig 12 below 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.

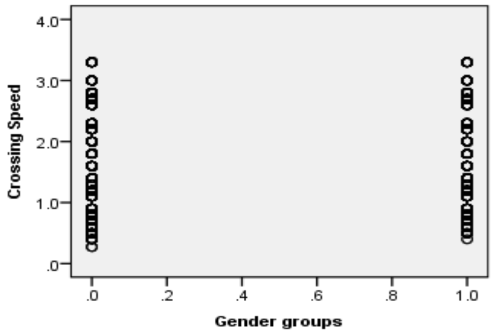

Fig 9 Scatter plot crossing speed vs. gender groups

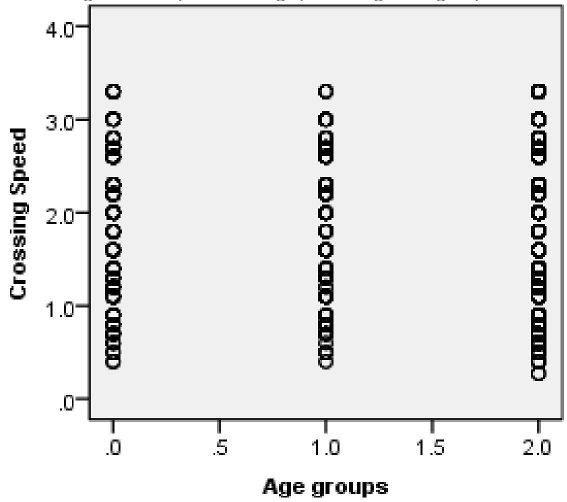

Fig 10 Scatter plot crossing speed vs. age groups

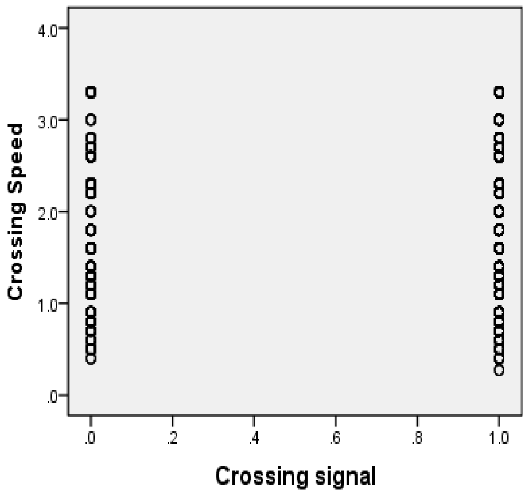

Fig 11 Scatter plot crossing speed vs. crossing signal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

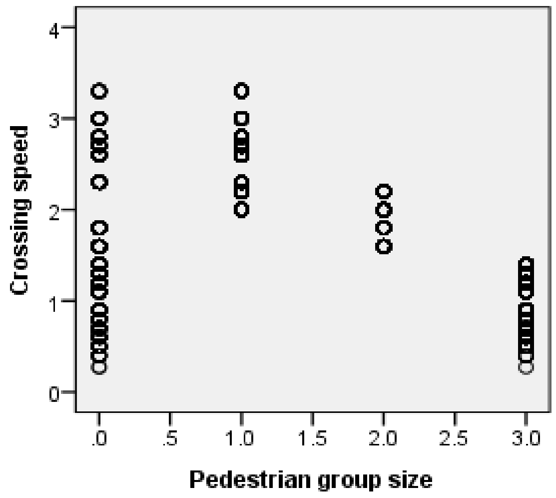

Fig 12. Scatter plot: Crossing speed vs pedestrian group size

Assumption 3: Multicollinearity (the predictor variables should not be highly correlated with each other*).*

Computing a correlation analysis in SPSS gives the following correlation matrix presented in table 12 below

TABLE 11 CORRELATION MATRIX BETWEEN INDEPENDENT VARIABLES

	Gender	Age group	Pedestrian group size	Signal phase
Gender	1			
Age group	-0.030	1		
Pedestrian group size	-0.080	0.240	1	
Signal phase	-0.130	0.004	-0.26	1

TABLE 12 VARIANCE INFLATION FACTOR FOR PREDICTOR VARIABLES

Variable	Gender	Age group	pedestrian group size	Signal phase
VIF	1.009	1.001	1.013	1.005

Assumption 4: 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.

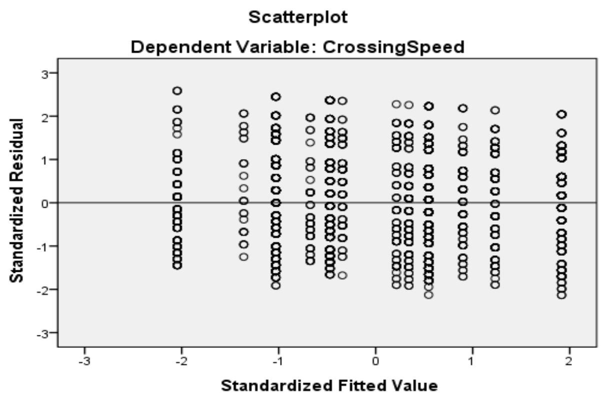

Figure 13 Scatter plot of standardized residuals vs. standardized fitted value.

Assumption 5: Residuals are normally distributed

The errors should be approximately normally distributed. The normal p-p plot if the residuals presented below shows the spread of the residuals approximately lying on the straight line (close to the line) which shows that the residuals are normally distributed thus, assumption 6 is satisfied.

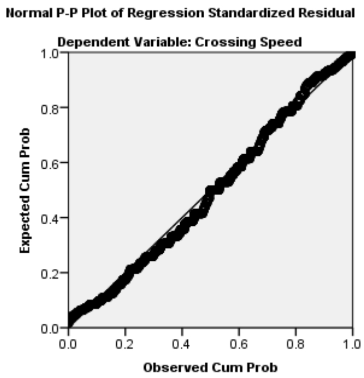

Figure 14 Normal p-p plot of residuals

D. Regression Analysis

Among the tested parameters, significant explanatory variables were selected and multiple linear regression analysis was performed to model pedestrian crossing speed at signalized intersections.

- Dependent variable

-crossing speed (Cs)

- Independent variables

-Gender (Male (0), Female (1))

-Age group (<20(1), 20-55 (2), >55(0))

- Pedestrian group size (single (0), 2-3 (1), 4-5 (2), >5 (3))

- Crossing signal phase (green (0), red (1)).

Enter' variable selection method is used to enter independent variable into the analysis. In this method all variables in a block are entered in a single step. Table 14 summarizes the analysis result.

TABLE 14 ML REGRESSION OUTPUT

Model (Crossing speed)	Coefficients β	Coefficients Std. error	t	Sig.(Pvalue)	95% CI for β	
Constant	1.751	0.026	75.618	0.000	1.695	1.842
Gender	-0.102	0.016	-3.931	0.015	-0.153	-0.051
Age group	0.063	0.015	4.374	0.000	0.042	0.095
Ped. Group size	-0.174	0.009	-18.825	0.000	-0.192	-0.156
Crossing signal phase	-0.002	0.031	-0.058	0.954	-0.063	0.052
R^2	0.8701					
Adjusted R^2	0.8598					

As shown in the table4.20 P value for the constant estimator and the first three explanatory variables is less than the significance level. This implies, except for crossing signal phase all predictors variables significantly predicted pedestrian crossing speed with R2adj =0.8598. This is interpreted as; about 85.98% of the variation in crossing speed is explained by gender, age groups and pedestrian group size.

The final prediction model with only the significant explanatory variables is written as:

Cs=1.751 - 0.102G +0.063A0.174Gr ....*Eqn 1*

Where:

Cs= Crossing speed

G= Gender

A= Age group

Gr=Pedestrian group size

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 06, June-2019

Effect of Pedestrian Characteristics on Signal Compliance

A binary logistic model is developed to predict the effect of pedestrian characteristics on signal compliance since, crossing signal takes two measurement level either 0 or 1. The final predicting model is described in table 15 below.

- Dependent variable

-Probability of signal compliance (Ps)

- Independent variables

-Gender (male (0), female (1)

-Age group (<20 (1), 20-55(2), >55(0))

- Pedestrian group size (single (0), 2-3 (1), 4-5 (2), >5 (3))

TABLE 15 BL REGRESSION OUTPUT FOR CROSSING SIGNAL COMPLIANCE

	variables	β	SE	Df	Sig.	Odds Ratio (Expβ)
Model (Probability of signal compliance)	Constant	1.662	0.102	1	0.000	5.269
	Gender	-0.075	0.101	1	0.360	0.927
	A0 (>55)	0.017	0.058	1	0.007	1.017
	A1(<20)	-1.160	0.024	1	0.007	0.314
	G1 (2-3)	-0.978	0.032	2	0.005	0.450
	G2 (4-5)	-0.510	0.041	2	0.015	0.600
	G3 (>5)	-0.346	0.011	2	0.003	0.707

TABLE 16 HOSMER-LAMESHOW GOODNESS OF FIT TEST FOR CROSSING SIGNAL COMPLIANCE MODEL

Step	Chi square	df	sig
1	6.096	1	0.645

The final prediction model with only the significant explanatory variables is expressed as equation 2 below.

Logit (Ps) = ln (p/(1-p)) =1.662+0.017A0-1.160A1-0.978G1-0.51G2-0.346G3…*Eqn. 2*

Where:

Ps –Probability of signal compliance

A0 – Age group (seniors)

A1– Age group (juveniles)

G1 – (2-3) Pedestrian group size

G2 – (4-5) Pedestrian group size &

G3 - >5Pedestrian group size

Before interpreting the estimators and odds of the model it is better to look at the goodness of fit test result. As it is shown in table 16 the HL goodness of fit test gives a P-value of 0.645 which is greater than the significant level. This shows that the model fit the data well.

For the age group categorical predictor having three level of measurement SPSS creates 2 dummy variables to represent these 3 groups. Here middle age group encoded 2 is considered as a reference group to which the other two groups should be compared. Similarly for the pedestrian group size categorical predictor that has four level of measurement SPSS creates 3 dummy variables. Single pedestrian group size category encoded 0 is considered as a reference group.

E. Model Validation

a) Multiple linear regression model validation

External model validation technique is used to validate the regression model developed for pedestrian crossing speed 1.744 randomly selected test data set, from Post office intersection, is used for the validation study. The test data set is independent from the training data set and there is no any prior relationship between them. Predictive squared correlation coefficient Q2 value estimated from external test data set is used to measure the predictive ability of the regression model.

$$Q_{F3}^2 = 1 - \frac{\sum_{i=1}^{n_{EXT}} (\hat{y}_i - y_i)^2 / n_{EXT}}{\sum_{i=1}^{n_{TR}} (y_i - \bar{y}_{TR})^2 / n_{TR}} \ldots\ldots\ldots\ldots\ldots\ldots . eqn\ 3$$

Where

$\hat{y}$-is Predicted value

Yi-Observed values

$\bar{y}$-TRis Response mean of training data set

$\bar{y}$-EXTis Response mean of external data set next is sample size of external data set

n_{tr}-is sample size of training data set.

The predictive squared correlation coefficient Q2 calculated using eqn 3 is 0.8494. Q2 value compares the predictive residual sum of squares (PRESS) with the total sum of squares (SS). The calculated model predictive ability measure, Q2, value is very close to the model's R2Adj value i.e.0.8598 which implies that the model's ability to predict future responses is as good as its performance on the training data. The validation study suggests that the developed model is appropriate for predicting future responses.

b) Binary logistic regression model validation

The same amount of validation data that is used for multiple linear regression model validation is used for the two binary logistic regressions developed in sections 4.4.2 and 4.4.3. The data is taken also from the same intersection i.e. Post office intersection.

Hosmer-Lameshow goodness of fit test is used to obtain the summary measure of test statistic for the validation sample.

$$C_v = \sum_{j=1}^{g} \frac{(O_j - E_j)^2}{n_j \bar{\pi}_j (1 - \bar{\pi}_j)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \textbf{\textit{eqn 4}}$$

Where:-

Cv –Hosmer-Lameshow goodness of fit test statistic for validation sample

Oj -Observed response

Ej –Expected response

nj –Total observation

$\bar{\pi}$_j= $\sum \frac{m_j \hat{\pi}_j}{n_j}$

mj- Number of subjects

$\hat{\pi}_j$–Logistic probability for the covariate pattern

Cv is distributed as χ2 (chi-square) with the null hypothesis H0: the model is correct. Using the validation data χ2 statistic result obtained for the H-L goodness of fit test statistic is 5.10 and 5.61 for probability of signal compliance logistic regression and probability of crosswalk compliance logistic regression models respectively. The respective P values are 0.784 and 0.850.This leads to the acceptance of the null hypothesis. As a conclusion, the binary logistic models developed for both signal and crosswalk compliance is found to be best predictive models.

IV CONCLUSIONS

This study set out to determine the significant factors affecting pedestrian crossing behavior expressed as crossing speed, signal and crosswalk compliance. In addition it is designed to determine the effect of dominantly affecting factors on crossing behavior the following conclusions and recommendations are drawn.

- Pedestrian characteristics i.e. gender, age group & group size and traffic characteristics i.e. crossing signal were tested if each of the factors significantly affect pedestrian crossing speed. As a result gender, age group and pedestrian group size were found as a dominant factors affecting pedestrian crossing speed.
- Age group and pedestrian group size are found as significant factors affecting pedestrian signal compliance and crosswalk compliance among the tested pedestrian characteristic.
- To study the effect of each predictor variable on pedestrian crossing speed multiple linear regression model is developed. Except crossing signal phase, all variables were found statistically significant to predict pedestrian crossing speed at 95% confidence level. The calculated R2adj for the regression model is 0.8598.
- Male pedestrians have greater crossing speed than female. Middle aged pedestrians have the greatest speed while seniors have the lowest crossing speed. Single pedestrians cross quicker than grouped pedestrians and crossing speed decreases as the number of pedestrians in the group increases.
- To study the effect of each predictor variable on signal compliance and crosswalk compliance binary logistic regression model is developed independently. Pedestrian age group and group size were found statistically significant to predict pedestrian probability of signal compliance and crosswalk compliance.
- At 95% confidence level. Hosmer-Lameshow goodness of fit test significance P value for signal compliance and crosswalk compliance are 0.645 and 1.000 respectively. This leads to the acceptance of the null hypothesis i.e. the developed model is correct and predicts the data well.
- External validation was performed to validate the multiple linear regression model and the model's predictive ability, Q2, is 0.8494. The regression model is validated and is applicable for signalized intersections of Addis Ababa.
- Similarly, to test the predictive ability of the developed binary logistic models external validation was performed using Hosmer-Lameshow goodness of fit test. H-L goodness of fit test statistic, Cv is 5.10 and 5.61 for signal compliance logistic regression and crosswalk compliance logistic regression models respectively. The respective P values are 0.784 and 0.850. This leads to the acceptance of the null hypothesis that states the model is correct. As a conclusion, the binary logistic models developed for both signal and crosswalk compliance is found to be best predictive models.
- The most dominating reasons of pedestrians for noncompliance with crossing locations are drivers' improper usage of roads during red phase and inadequacy of width of cross walk. Reasons for noncompliance with traffic light most likely related to the low level awareness of pedestrians and gap in enforcement and implementation of law.

A) Recommendations

Depending on the final results of this study the researcher recommends the following points so that it adds its own positive impact on the improvement of road users' life especially of pedestrians.

- Education & enforcement of traffic law especially for juveniles and middle aged pedestrians is essential for the improvement of pedestrians' compliance with signal as well as with crosswalk.
- Design of crosswalk width should be revised since the number of pedestrians in the city is increasing with high rate. So that, the designed crosswalk could accommodate pedestrians crossing with mass and pedestrians could comply with crosswalk.
- Enforcing laws have to be implemented for drivers on usage of roads during red signal phase so that pedestrians can use crosswalks properly.

In addition to the above recommendations the following future research areas related with this study are also recommended.

- Pedestrians delay model at signalized intersection
- Disabled pedestrians crossing behavior at signalized intersections
- Pedestrian crossing behavior at un-signalized intersections and at midblock.

REFERENCES

[1] Akash, Jain, Ankit Gupta , and Rajat Rastogi. "pedestrian crossing behavior analysis at intersections." International journal for traffic and transport engineering, 2014: 103-116.
[2] Alhajyaseen, W, and Asano M Iryo. "studying critical pedestrian behavioral changes for safety assesment at signalized crosswalk." saftey science, 2017: 351-360.
[3] Allen T Mense. Introduction to regression techniques. Principal engineering fellow, n.d.
[4] Ann, Forsyth, Weinstein Agrawal Asha, and J Krizek Kevin. "Simple inexpensive approach to sampling for pedestrian and bicycle surveys." Journal of transportation research board, 2012: 22-30.
[5] Arboretti, R Giancristofaro, and L Salmaso. "Modeling performance analysis and model validation in logistic regression." Journal of royal statistical society, 2003.
[6] Asaithambi, Gowri, O Kuttan Manu, and Sarath Chandra. "Pedestrian crossing behavior under mixed traffic conditions;a comparative study of an intersection before and after implementating control measures." Transportation in developing economies, 2016: 1-12.
[7] Athnatios, Galanis, and Eilous Nicholas. "pedestrian crossing behavior in signalized crossing in middle size cities in Greece." Tagungsbad: REAL CORP, 2012.
[8] Central Statistical Agency. 2013.
[9] Consonni, V, D Ballabio, and R Todeschini. "Evaluation of model predictive abalility by external validation technique." Journal of chemometrics, 2010: 194-201.

[10] Daamen, W, and Sp Hoogendoorn. Experimental research of pedestrian walking behavior. Annual meeting, Washington DC USA: Transportation Research Board, 2003.
[11] Deepty, Muley, Karbache Mohammed , Alhagyaseen Wael, and Al-salem Mohammed. "Pedestrian crossing behavior at marked crosswalks on channalized right turn lane at intersection." Science direct, 2017: 233-240.
[12] Ethiopian Transport Authority. 2017.
[13] Gang, Ren, Zhu Zhuping, Wang Wei, Zhang Yong , and Wang Weijie. "crossing behavior of pedestrain at signalized intersection." Transport research board , 2011.
[14] Hosmer, D W, and S Lemeshow. Applied logistic regression 2nd Edition. Newyork: John wiley & sons, 2000.
[15] Jiang , Xiaobie, Wang Wuhong, Bengler Klause, and Guo Weiwei. "Analysis of pedestrian behavior on mid block unsignalised crosswalk comparing Chineese and German cases." Advance in mechanical engineer, 2015: 1-7.
[16] King, Mark j, W Soule David, and Ghaoufarian Ameneh. "Illegal pedestrian crossing at signalised intersection incidence and relative risk." accident analysis and prevention ,41 no 3, 2009: 485-490.
[17] Koh, P.P. "Safety evaluation of pedestrian behavior and violations at signalized pedestrian crossings." Safety science 70 (2014): 143-152.
[18] Kreizek, J K, A Forsyth, and w Agrawal A. Pedestrian and bicycling survey (PABS): User's Manual.http//www.designforhealth.net/resources/. n.d. http://www.survey system.com/sscalc.htm. (accessed 12 30, 2018).
[19] Krysto, Lipovac, Vujanic Milan, Maric Bojan, and Nesic Miladin. "pedestrian behavior at signalized pedestrian crossings." Journal of Transport engineering, 2013: 165-172.
[20] Laura, Pathric Nichlas. Population 2018. 2018. http://.poplulation2018.com.
[21] Marisamynathan, Vendagri Perumal. "Study on pedestrian crossing behavior at signalized intersections." Journal of traffic and transportation engineering, no. 1(2) (2014): 103-110.
[22] Maristamynathan, S, and P Vendagri. "A statistical analysis of pedestrian behavior at signalized intersection." European transport, 2015: 1825-3997.
[23] Montogomery, C Douglas, and C George Runger. Applied statistics and probability for engineers. USA: John Wiley and sons, 2010.
[24] Nesic, Miladin, Krysto Lipovac, and Miroslav Rosic. "Pedestrian behavior at pedestrian crossing Regulated with traffic light." Road accident prevention. Belgrade serbia: XII International sympossium, 2014.
[25] Nicholas N Ferenchak, and Matin Kitirai. "Pedestrian crossing behavior in relation to grouping and gender in a developing country context." Global epidemology and environmental health , 2017: 37-45.
[26] Rosenbloom, Tova. "Crossing at red light : Behavior of individuals and groups." Transportation research F 12 (2009): 389-394.
[27] Scott, H Simpson. "Creating a data analysis plan :what to consider when choosing statistics for a study." Research primer, 2015: vol 68
[28] Sohel, Rana, Midi Habshah, and S k Sarkar. "Validation and Performance analysis of binary logistic regression model." WSEAS Internationl conference on environmental medicine and health science. Serdag, Malasia, 2015.
[29] Srinivas , Palamarthy, S Hani, Mahmassani, and B Machemhl Randi. Models of pedestrian crosssing behavior at signalized intersections. texas: center for transportation research, 1994.
[30] Srinivas, Palamarty, S Mahmassini Hani, and B Randi Machmehl. "Model of pedestrian crossing behavior at signalized intersection." Cetre for transportation research, 1994.
[31] Sun , Z. "study on pedestrian characteristics in crosswalks at signalized intersections." Msc thesis, Beijing, bBeijing university of technology, 2004.
[32] Tasnima , Abedin, Z.I Chowuduri Mohammed, Afzal Arfen, Turin Tanvir, and F Yeasmin. "Application of binary logistic regression in clinical research." JNHFB, 2016: 8-11.
[33] Tulu, Getu Segni. "why are pedestrian crash so different in developing countries? a review of relevant factors in relation to their impact in Ethiopia." Birsbane, Australia: Australasian transport research forum, 2013. page 4.
[34] Wall. "Road markings to improve pedestrians safety at crossings." Traffic engineering and control, 2000: 136-140
[35] Washington, Simon P, G Kralaftis Mattew, and L Mannering Fred. *Statistical and econometric methods for transportation data analysis.* 3. London: Newyork; Washington DC: CRC press, 2010.