Measurement of Traffic Accident Size Using Structural Equation Modeling

Pankaj Prajapati
Associate Professor,
Dept. of Civil Engineering,
Faculty of Technology & Engineering,
The Maharaja Sayajirao University of Baroda,
Vadodara, Gujrat, India

Abstract

Traffic Accident size can be expressed as the number of deaths and/or the number of injured. Accident size is the one of the important indices to measure the level of safety of transportation facilities. Factors such as road geometric condition, driver characteristic and vehicle type may be related to traffic accident size. However, all these factors interact in complicate ways so that the interrelationships among the variables are not easily identified. A structural equation model (SEM) is adopted to capture the complex relationships among variables because the model can handle complex relationships among endogenous and exogenous variables simultaneously and furthermore it can include latent variables in the model. In this study, we use accident data having 1867 traffic accident rewarded in Vadodara city since 2013 to 2017 and estimate relationship among exogenous factors and traffic accident size. The model suggests that road factors, vehicle factors, victim factors and time factors are strongly related to the accident size.

Keywords: Traffic Crash; Accident Size Analysis; Factor Analysis; Structural Equation Modeling.

Introduction

Traffic accident forecasting models have been developed to understand factors affecting traffic accidents and eventually to reduce traffic accidents by controlling and/or improve factors. Accident statistics most often used to quantify and describe three principal informational elements: accident occurrence, accident involvements and accident severity. Accident occurrence relates to the numbers and types of accidents, accident involvements concerns the numbers and types of vehicles and drivers involved in accidents, and accident severity is generally expressed as the numbers of deaths and/or injuries occurring. While each statistic provides a meaningful information, an integrate information of accidents is also very useful. A new statistic “accident size” is adopted in this study, which can be described in terms of the number of deaths and injured persons. From the previous researches, factors such as road geometric conditions, driver characteristics and vehicle types can be related to accidents.

However, all those factors interact in complicated way so that the inter relationships among the variables are not easily identified. In this study, we use 1867 accident data occurred in Vadodara city of Gujarat, India and estimate relationships among exogenous factors and traffic accident size. In modeling process, we create exogenous latent variables such as “road factors”, “vehicle factors”, “victim factor” and “time factors” to identify latent relationships to an endogenous variable “accident size”.

Aim of the Study

A structural equations model (SEM) is adopted to capture the complex relationships among variables because the model can handle complex relationship among endogenous and exogenous variables (responsible for road accidents) simultaneously and furthermore, it can include latent variables in the models. This will help the administrators to take appropriate actions to control road accidents.

Regional Road Accident Scenario

Gujarat is one of the most industrially developed and agriculturally advanced fertile state of India. So as the road length in Gujarat has increased from 47,426 km in 1981 to 67,065 km in 1991 to 79,619 km in 2011. With increased in road length, the total number of registered vehicles in Gujarat has increased from 10,28,90,560 in 2007-2008 to 23,28,64,180 in 2017-2018. Due to that the rate of accidents in Gujarat is 12.4 accidents per 10000 vehicles. Gujarat has 38 fatality rate of 100 accident average and it is increased since 2007. The fatality rate of 2007 is 21 of 100 accident average. (RTO, 2017)

In 2016, total number of registered accident is 1046 in Vadodara city of Gujarat, India. In that injury accident is 654 and fatal accident is 203 recorded. In 2016, 878 person is injured and 214 person are killed in road accident in Vadodara. The severity index is 20.5 per number of person killed per 100 accidents in Vadodara city. (MoRT&H, 2016)

Review of Literature

Extensive effort has been made by many researchers in various transportation fields to explain traffic accident occurrence and factors affecting accidents. They attempted to develop different types of model, which can explain severity of accidents and eventually understand and predict accidents. Several selected studies related to ours are summarized among numerous studies, which are accident analysis and SEM. The description of SEM will be presented in next section.

Zhou et al. (2015) have observed in their study that the action of the family members and friends was more influential on pedestrians than whether or not they approve of the behaviour. It was also supported by other studies that most of human behaviour was learned by observing. By observing family and friends violate traffic laws, an individual would imitate the behaviour.

Al-Mahameed et al. (2019) used the structural equation model technique in their study for pedestrian and bicyclist crash frequency and more than 60 explanatory variables for 200 highway corridors in Wisconsin were collected. The interrelationships between observed “manifest” variables and unobserved “latent” variables were tested. The results suggested that the most important latent variables influencing the crash frequency of VRUs are bicycle/pedestrian-oriented roadway design (e.g., paved shoulders, sidewalks, and bike lanes), exposure (e.g., walking and biking activity, and employment density), and low social status (e.g., educational level, and wage percentage). The benefits of this study might help community planners, transportation researchers, and policymakers with a better understanding of the intricate interrelationship of the influential factors contributing to VRUs road crashes.

For Hong Kong, the effects of factors influencing on the severity of injury from an accident was examined for factors such as human, vehicle, safety, environment and site. Risk factors associated with each of the vehicle types were identified by means of step wise logistic regression models. For private vehicles, district board, gender of driver, age of vehicle, time of the accident and street light conditions were significant factors determining the severity of injury (KelvinandYau,2004). Miltonetal.(2008) proposed a mixed logit model using highway-injury data from Washington State. Findings in the study indicated that volume-related variables such as average daily traffic per lane, average daily truck traffic percentage, number of interchanges per mile and weather effects such as snowfall are best modelled as random-parameters, while roadway characteristics such as the number of horizontal curves, number of grade breaks per mile and pavement friction are best modelled as fixed parameters.

Kim et al. (2007) conducted research for the factors contributing to the injury severity of bicyclists in bicycle–motor vehicle accidents using a multinomial logit model. The model predicted the probability of four injury severity outcomes: fatal, incapacitation, non-incapacitation, and possible or no injury. The results showed several factors, which more than double the probability of bicyclist suffering a fatal injury in an accident, all other things being kept constant. Notably, inclement weather, darkness with no streetlights, a.m. peak (06:00a.m. to 09:59a.m.), head-on collision, speeding-involved, vehicle speeds above 48.3 km/h, truck involved, in toxicated driver, bicyclist age 55 or over, and intoxicated bicyclist. Many researches applying the SEMs can be found in transportation fields. These researches try to understand the complex relationships among the variable using SEM. Hamdar et al. (2008) developed a quantitative intersection aggressiveness propensity index (API). The index was intended to capture the overall propensity for aggressive driving to be experienced at a given signalized intersection. The index was a latent quantity that could be estimated from observed environmental, situational and driving behavior variables using SEM techniques. The exogenous variables were number of heavy vehicles, number of pedestrians, traffic volume, average queue length, percent grade, number of lanes, number of left turn lanes and so forth.

Choo (2007) analyzed telecommunications impacts on travel in a comprehensive system considering demand, supply, costs, and land use, using SEM. The model results suggested that as telecommunications demand increases, travel demand increases, and vice versa. Additionally, transportation infrastructure and land use significantly affect travel demands. In addition, SEM is frequently adopted intravel value and behavior field. Chung and Ahn (2002) developed SEM that presented relationships among socio-demographics, activity participation (i.e., time use), and travel behavior for each day during a week in a developing country. It was tentatively concluded that there were similar relationships between socio-demographics and travel behaviors in developing and developed countries. It was also confirmed that activity patterns were significantly different on weekdays and weekends. Furthermore, during weekdays there were day-to-day variations in the patterns of activity participation and travel behavior. Choi and Chung (2003) adopted multivariate SEM to handle the hierarchical nature of the data and explain complex relationship among socioeconomic factors of individuals and household, activity participation, and travel behavior using Puget Sound Transportation Panel data. Chung and Lee (2002) constructed an SEM to estimate aggregated automobile demand with data from Korea. The results indicated that both the number of driver’s license holders and total road length had a statistically significant effect on automobile demands. In addition, several other determinants of the endogenous variables were found such as average household size, economically active population, personal transportation expenditure, urbanized area, and population density. Lu and Pas (1999) described the development, estimation andi nterpretation of a model relating socio-demographics, activity participation (time use) and travel behavior. Activity participation (time allocated to a number of activity types) and travel behavior were endogenous to the model. They reported the relationships betweenin-homeandout-of-home activity participation and travel behavior.

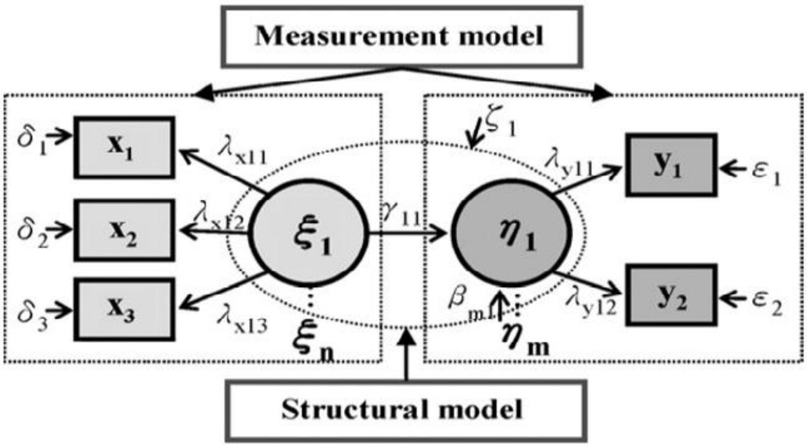
Research Methodology

Structural Equation Modeling is a technique that can handle a large number of endogenous and exogenous observed variables simultaneously. Since SEM consists of a set of equations that are specified by direct links between variables, it can be called "the simultaneous equations" from the perspective. However, in SEM, we can introduce 'latent variables' which are the unobserved variables and represent uni-dimensional concepts in their purest form. Other terms for these are unobserved or unmeasured variables and factors. The observed variables of a latent variable contain random or systematic measurement errors, but the latent variable is free of these. Since all latent variables corresponding to concepts, they are hypothetical variables. Latent variables specified as linear combinations of the observed variables. The linear combinations are weighted averages. Hence, regression, path analysis, factor analysis and canonical correlation analysis are all special cases of SEM. In SEM we can separate errors in measurement from errors in equations. (Bollen, 1989)

Elements of SEM

A SEM with latent variables has at most three components as shown in Figure 1: (a) a measurement model for the endogenous variables (Y measurement model), (b) a measurement model for the exogenous variable (X measurement model), and (c) a structural model.

Figure 1 an example of SEM



The basic equation of the latent variable model is the following (Bollen, 1989):

$$\eta = B\eta + \Gamma\xi + \zeta \quad \text{.....(1)}$$

where:

η (eta)is an (*m* x 1) vector of the endogenous latent variables

ξ (xi)is an (*n* x 1) vector of the exogenous latent variables

ζ (zeta)is an (*m* x 1) vector of random variables

B and Γ are structural coefficient of the model.

The structural parameters are the elements of the three matrices. B is the matrix (m×m) of direct effects among endogenous latent variables and Γ is the matrix (m×n) of regression effects for exogenous latent variables to endogenous latent variables. Λ is linking matrix between the latent and observed variables. The elements of SEM are explained in Table 1.

Table 1 Elements of Structural Equation Model

Measurement model	X	q×1 column vector of observed exogenous variable
	Y	p×1 column vector of observed endogenous variable
	ξ	n×1 column vector of latent exogenous variable
	η	m×1 column vector of latent endogenous variable
	δ	q×1 column vector of measurement error terms for observed variable x
	ε	p×1 column vector of measurement error terms for observed variable y
	Λx	The matrix (q×n) of structural coefficient for latent exogenous variables to their observed indicator variables
	Λy	the matrix (p×m) of structural coefficient for latent endogenous variables to their observed indicator variables
Structural model	Γ	the matrix (m×n) of regression effects for exogenous latent variables to

		endogenous latent variables
	B	the coefficient matrix (m×m) of direct effects between endogenous latent variables.
	ς	m×1 column vector of the error terms
Covariance matrix	Θ_ε	The covariance matrix (p×p) of ε
	Θ_δ	The covariance matrix (q×q) of δ
	Φ	The covariance matrix (n×n) of ξ
	ψ	The covariance matrix of (m×m) of ξ

The β coefficients (components of B matrix) and the γ coefficients (components of Γ matrix) are magnitudes of expected changes after a unit increases in η or ξ. Similarly, coefficients (components of Λ matrix) are expected changes of observed variables with respect to a unit change in the latent variable.

Data description

Growth of road transport in Vadodara city is very fast. There is also considerable increase in vehicle ownership and population in the city. The accident data has been collected from various police station of Vadodara city of the year 2013-2017 with permission of police department. This data has been complied and analysed in this paper.

A total accident data was collected for years 2013-2017. This data included 2854 accidents between all types of vehicles and pedestrian. 1867 cases of conflict between two vehicles were noted. The proportion of this type of conflict was higher as in the city most of the accidents occurred between two vehicles.

Data encoding of conflict between two vehicles

Coding of data is crucial and important part of data analysis. Before any analysis procedure conducted data should be converted in to proper format so it can be easily applicable to any relevant software and examined into it.

Table ***2*** shows the coding adopted for conflict between two vehicles.

Table 2 Coding of Accident Data of Two Conflicting Vehicles

No.	Description	Measurement scale
1	Divider	0. No 1.Yes
2	Collision spot	0. Road junction 1. On straight road
3	Collision type	0. vehicle hit from side 1. vehicle head on 2. vehicle hit from back
4	Impacting vehicle	0. Bike 1. Auto Rickshaw 2. Car 3. Light Commercial Vehicle (LCV) 4. Heavy Commercial Vehicle (HCV)
5	Impacting driver gender	0. Man 1. Woman
6	Impacting vehicle manoeuver	0. Proceeding straight 1. Turning 2. Overtaking 3. Parked 4. Wrong side
7	Victim Vehicle	0. Bicycle 1. Bike 2. Auto Rickshaw 3. Car 4. L.C.V. 5. H.C.V.
8	Victim Vehicle driver gender	0. Man 1. Woman
9	Victim type	0. Passenger 1. Driver
10	Victim age group	0. Child 1. Minor 2. Adult-1 3. Adult-2 4. Senior citizen
11	Weekday	0. Sunday 1. Monday 2. Tuesday 3. Wednesday 4. Thursday 5. Friday 6. Saturday
12	Day or night	0. Night 1. Day
13	Season	0. Winter 1. Summer 2. Monsoon
14	No. of fatalities	0. No fatalities 1. One fatalities 2. Two or more fatalities
15	No. of injuries	0. No injury 1. One injury 2. Two injury 3. Three or more injury

SEM model for two vehicles conflict

The data used in this study are 1867 complete accident records during the year 2013-2016 between two conflicting vehicles, Each accident record has various and rich information such as the divider (whether the road has divider has or not), collision spot, Day or Night (when accident occurred at that time the scenario of time was day or night), Season, weekday, holiday, Collision type, impacting vehicle type, impacting vehicle manoeuvring, victim vehicle type, impacting vehicle driver gender, victim vehicle driver gender, victim age group, victim type, number of injured person, number of deaths.

Table 3 Descriptive Statistics of Conflict Between Two Vehicles

Observed variables		Frequency	Percentage	Mean No of Fatalities	Mean No of Injury
Divider	No	557	29.83%	0.075	1.443
	Yes	1310	70.17%	0.156	1.302
Collision Spot	Road Junction	622	33.32%	0.068	1.457
	On Straight Road	1245	66.68%	0.164	1.288
Day or Night	Night	863	46.22%	0.153	1.340
	Day	1004	53.78%	0.114	1.349
Season	Winter	588	31.49%	0.155	1.304
	Summer	692	37.06%	0.133	1.368
	Monsoon	587	31.44%	0.107	1.356
Weekday	Sunday	260	13.93%	0.165	1.454
	Monday	256	13.71%	0.129	1.285
	Tuesday	237	12.69%	0.122	1.312
	Wednesday	278	14.89%	0.151	1.335
	Thursday	261	13.98%	0.115	1.276
	Friday	295	15.80%	0.119	1.366
	Saturday	280	15.00%	0.121	1.375
Holiday	No	1536	82.27%	0.134	1.334
	Yes	331	17.73%	0.121	1.393
Collision Type	Vehicle Hit From Side	509	27.26%	0.073	1.458
	Vehicle Head On	424	22.71%	0.071	1.422
	Vehicle Hit From Back	934	50.03%	0.192	1.247
Impacting vehicle type	Bike	765	40.97%	0.064	1.400
	Auto Rickshaw	114	6.11%	0.035	1.518
	Car	594	31.82%	0.126	1.311
	LCV	86	4.61%	0.128	1.244
	HCV	308	16.50%	0.347	1.234
Impacting vehicle Manoeuvre	Proceeding Straight	1513	81.04%	0.132	1.350
	Turning	196	10.50%	0.066	1.464
	Overtaking	27	1.45%	0.111	1.185
	Parked/Stopped	34	1.82%	0.088	1.441
	Wrong Side	97	5.20%	0.289	1.031
Victim vehicle Type	Cycle	118	6.32%	0.178	1.000
	Bike	1118	59.88%	0.134	1.345
	Auto	109	5.84%	0.073	1.679
	Car	306	16.39%	0.039	1.402
	LCV	54	2.89%	0.222	1.130
	HCV	162	8.68%	0.265	1.327
Impacting vehicle driver Gender	Man	1786	95.66%	0.136	1.348
	Woman	81	4.34%	0.037	1.272
Victim vehicle driver Gender	Man	1778	95.23%	0.134	1.345
	Woman	89	4.77%	0.090	1.337
Victim Group	Child	3	0.16%	0.333	1.000
	Minor	79	4.23%	0.190	1.253
	Adult-1	678	36.31%	0.131	1.397
	Adult-2	955	51.15%	0.127	1.332
	Senior Citizen	152	8.14%	0.132	1.243
Victim Type	Passenger	200	10.71%	0.190	1.430
	Driver	1667	89.29%	0.125	1.334

The influence of road factors in the occurrence of an accident is significant.

Table *3* considers divider as a part of road and displays the proportion of accidents that occurred in presence and absence of a divider on the carriageway. It can be inferred that divider has a significant impact on the occurrence of accidents as 70% of accident had occurred in presence of divider.

P: ISSN No. 2231-0045 RNI No. UPBIL/2012/55438 VOL.-7, ISSUE-4 (Part-1) May-2019

E: ISSN No. 2349-9435 Periodic Research

Table *3* shows the proportion of accidents with reference to the location of the occurrence of the accidents. It is clearly visible that there are more chances (67%) of the occurrence of an accident on a straight road as compared to a junction. The reason behind this may be that people tend to reduce the speed of vehicle at junctions being precautious to accident.

Table *3* portrays the comparison of accident rates with respect to the time viz. day or night, seasonal variation, weekdays and holidays. But obvious the number of accidents will be more at night time as vision during night time is less as compared to day time and lights also play a significant role in the accident. It can be seen that there is not much difference in the accident rates with the variation of seasons.

Table *3* depicts the number of accidents taking place on weekday and a holiday. It was observed that more number of accidents occurred on weekdays. This result can be attributed to the trip purpose as more trips are made on weekdays for work, school, colleges, etc. while on holidays the trip purpose may be solely refreshment. Further the number of weekdays are more as compared to holidays, so the chances of occurrence of accidents on weekdays are more.

It can be clearly seen that the rear end collisions are higher as compared to head on collision and collision from side. The reason behind this may be that there must not be sufficient braking distance available before collision.

It can be inferred that male dominate the proportion, the reason being that male tend more to make trips as compared to female due to work purpose and various other factors. The proportion of victims can also be classified on the basis of age group. As shown, the percentage of adults i.e., between 19-60 years of age dominate this classification. This may be due the fact that adults are more among commuters with personalized vehicles as compared to kids and aged people, as they prefer public transport or are dependent on others for making a trip. The proportion of drivers (90%) among victims is higher as to those of passengers. This can be explained by the fact that in most of the type of accidents drivers are the first one being affected due to collision between two vehicles.

Development of SEM

The final model specification is derived using a two-stage development process. At the first stage, we conduct factor analysis to classify observed variables into several groups. Factor analysis is often used to analyse the correlations among several variables in order to estimate and to describe the number of fundamental dimensions that underlie the observed data. Those fundamental dimensions (factors) can be latent variables in SEM. At the second stage, we estimate the correlations matrix of observed variables and finally develop a SEM having the best-fit statistic. (Lee, Chung, & Son, 2008)

Factor analysis

Factor analysis is performed on 13 X observed variables, based on the result of which exogenous latent variables is determined. The results of the factor analysis for two vehicle conflict and pedestrian conflict with orthogonally rotated are shown in

Table *3*. Which are the correlations between each variable (rows) and each factor (columns). Loadings nearby 0.6 are usually considered 'high' and those below 0.4 are 'low.'

Table 4 Rotated Component Matrix of two vehicles conflict

No.	Observed variables	Component						
		1	2	3	4	5	6	7
1	Divider	.470	-.122	-.288	.154	.076	.003	.401
2	Collision spot	.819	.026	.037	-.048	.062	-.043	-.054
3	Day night	-.019	.014	-.018	.151	.197	.816	-.073
4	season	.065	.137	.170	.140	.698	.080	-.109
5	weekday	-.103	.052	.086	-.065	-.041	.020	.827
6	Collision type	.796	.061	-.052	-.026	-.056	.057	-.069
7	Impacting vehicle type	.144	.304	-.657	.090	.025	-.097	.216
8	Impacting vehicle manoeuver	-.043	-.234	-.187	-.304	.653	-.072	.149
9	Victim vehicle type	.064	.681	.180	-.342	-.031	-.086	-.059
10	Impacting vehicle driver gender	.028	.106	.726	.169	.076	-.071	.255
11	Victim vehicle driver gender	-.024	-.066	.064	.824	-.021	-.012	-.020
12	Victim age group	.052	-.057	.036	-.276	-.312	.595	.142
13	Victim type	.010	-.737	.222	-.126	-.023	-.042	-.073

The relationship of each variable to the underlying factors expressed by the so-called factor loading. For example, the first factor can be called 'road factor' because it is seems like divider presence and collision spot load highly on it. The second factor can be called 'vehicle factor' because Impacting vehicle and victim vehicle have high loadings for the factor and also we put impacting vehicle manoeuvring in that because it is most suitable in that factor. The third factor called 'victim factor' is associated with impacting vehicle driver gender and victim vehicle driver gender and the victim age group also merged with that group because it has similar properties. Other observed variables are classified into fourth, fifth and sixth factors. Generally, varimax rotation is a common technique, which attempts to minimize the complexity of the factors by making the large loadings larger and the small loadings smaller within each factor.

Finally, four factors are used with exogenous latent variables in the model. Sixteen observed variables (14 X observed variables and two Y observed variables) into five latent variables (four exogenous and one endogenous variables) for SEM are classified based on the result of factor analysis. Exogenous latent variables are factor1 (road factor), factor2 (vehicle factor), factor 3 (victim factor) and factor 4 (time factor) Endogenous latent variable is “accident size factor”.

Abbreviations of parameters

The following Table 5 explains the abbreviations used in the model.

Table 5 - Abbreviations of Variables

No.	Abbreviations	Full form
1	Divider	Divider
2	Collisionspot	Collision spot
3	daynight	Day night
4	season	season
5	weekday	weekday
6	V1type	Impacting vehicle type
7	V1manoeuvering	Impacting vehicle manoeuver
8	V2type	Victim vehicle type
9	V1drivergender	Impacting vehicle driver gender
10	V2drivergender	Victim vehicle driver gender
11	Victimagegroup	Victim age group
12	noofinjury	No. of injury
13	nooffatalities	No. of fatality

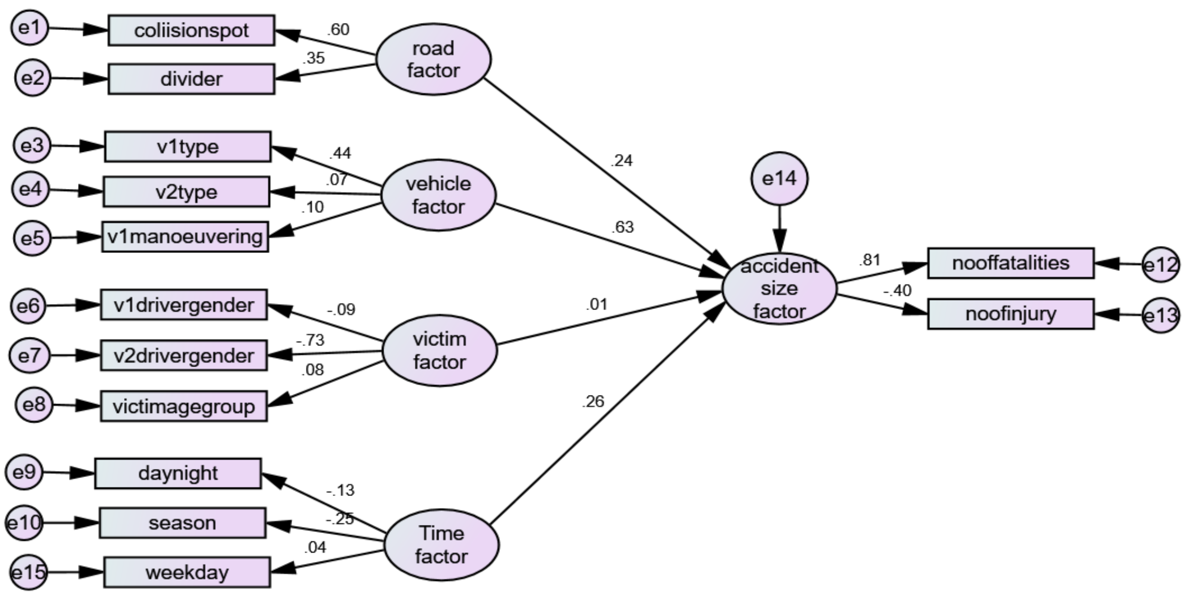
Results

Path Diagram for Two Vehicle Conflict

Path diagram shows a conceptual relationship of observed variables and latent variables. For present study a concept diagram is shown in *Figure 2*

In our case ML (Maximum likelihood) estimation method is employed. In above figure the regression weights are all standardized regression weights are given. The proportion of 2W and cars was higher in the accident affected vehicles as compared to other categories. Major proportion of victims consisted of those who preferred car, 2w and 3w as a mode of transport as compared to LCV and HCV, but higher rate of fatalities were observed for the victims of LCV and HCV as compared to the former categories. More fatalities were noted when accident occurred as a result of straight moving vehicle and vehicles travelling on wrong side.

Figure 2 Final Path Diagram



The probability of fatal accidents was higher while vehicle was travelling straight as the proportion of traffic moving straight is higher, but it was investigated that fatal accidents dominated in the category of wrong moving vehicles. The probability of injuries in accident during turning movement was similar to that during moving forward, although the less cases of turning movements were conformed as compared to the later. Accounting to the above mentioned reasons it can be concluded that vehicle factor had a significant impact on accident size Figure 2

Further, time factor is the second dominating factor on accident size Figure 2. Under this, the proportion of fatal accidents during night time was higher and higher injuries were noted during day time. This can be attributed to the night vision that comes into the picture during night time. No significant variation was observed in the proportion of accidents on weekdays.

Road factor is the third one influencing the accident size. The number of injuries were higher in the absence of divider but in presence of it, the proportion of fatalities in accident affected people dominated. Further, higher rate of injuries were observed in the accidents that happened at road junctions, but the probability of fatalities during accidents on straight road was higher.

On modeling it was deduced that the regression weights were negative for impacting vehicle drivers and victim vehicle drivers affected in accidents. This was due to the fact that the proportion of female drivers was less as compared to the male drivers. Similarly, the observed regression weights of male drivers were positive for the effect of gender on accident size, as the proportion of male drivers was more in the commuters, so they are more likely to meet an accident or be a cause of accident. While deriving the correlation between the age group of the commuters and the accident size, the regression weight for the adult group of 19-60 was the highest in accident size and the fatalities and injuries noted were also significant. No matter the female proportion was less in victim driver but can't be ignored it too has an effect on the accident size.

In addition to above findings, correlation between fatalities and injuries was evaluated viz. how much proportion of change was observed in injuries when there was a unit change in fatalities. A combined effect of these two variables on accident size was then evaluated.

From Table ***6*** the goodness of fit is quite satisfactory, these indices determine how well a our model fits the sample data and allow model with superior fit to be chosen. (Byrne, 2016)

Table 6 - Indices of Goodness of fit (two vehicles conflict)

No.	Description	Observed Value	Permissible Value
1	Chi-Square/Degree of freedom	2.897	<=3.00
2	Goodness of fit (GFI)	0.980	>0.90
3	Adjusted Goodness of fit (AGFI)	0.971	>0.90
4	Comparative fit index (CFI)	0.808	>0.90
5	Root mean square residual (RMR)	0.026	<0.10
6	Root mean square Error (RMSEA)	0.040	<0.06 or <0.08

Conclusions

In this research, we postulated that road factors, vehicle factors, victim factor and time factors are exogenous latent variables and accident size factor is an endogenous latent variable for SEM to analyse traffic accidents size. The observed variables for latent variables are collision spot, divider, impacting vehicle type, victim vehicle type, impacting vehicle manoeuvring, impacting vehicle driver gender, victim vehicle driver gender, victim age group, day or night, season and weekday. Using factor analysis, the 11 variables are grouped into five latent variables (four exogenous and one endogenous variables) for SEM.

The SEM illustrates positive or negative effects of each variable on the accident size. According to the SEM model, the total effect of vehicle factors on accident size is 0.63, so that accident size tends to increase when vehicle factors have higher values. Vehicle factors increase in case of proportion of 2W, car are increases in impacting vehicle type. The estimated coefficient of road factors is a positive value 0.24. This result indicates that road junction and undivided road are tends to decreases accident size. In case of time factors, the estimated coefficient is 0.26, which means that day or night and seasons are tends to decrease accident size.

The estimated coefficients are all standardized solutions, so we can conclude that the major factors influencing on the accident size is vehicle factor. Among four exogenous latent variables (road, vehicle, victim and time factors), the effect of vehicle factor on accident size is highest. In order to decrease the traffic accident size handling the road factor is more effective than handling vehicle, victim and time factors. It can be a positive result to traffic engineers because as they can handle 'road factors', they hardly manage 'vehicle, victim and time factors'. The findings in this research offer information about the relationships between accident size and various factors and they can contribute to reduce traffic accident size. Although there are countless factors having relation to "accident size", obtainable information from fields is very limited. Hence, while some aspects are not properly described and explained by models.

References

Bollen, A. K. (1989). Structural equations with latent variables.

Choi, Y. S., & Chung, J. H. (2003). Multilevel and multivariate structural equation models for activity participation and travel behavior. Journal of Korean Society of Transportation 21 (4), 145-154.

Chung, J. H., & Ahn., Y. (2002). Structural equation models of day-to-day activity participation and travel behaviour in a developing country. Transportation Research Record 1807, 109-118.

Chung, J. H., & Lee, D. (2002). Structural model of automobile demand in Korea. Transportation Research Record 1807, 87-91.

Farah J. Al-Mahameed, Xiao Qin, Robert J. Schneider, Mohammad Razaur Rahman Shaon , (2019), Analyzing Pedestrian and Bicyclist Crashes at the Corridor Level: Structural Equation Modeling Approach, Transportation Research Record, https://doi.org/10.1177/0361198119845353

Hamdar, S. H., Mahmassani, H. S., & Chen, R. B. (2008). Aggressiveness propensity index for driving behavior at signalized intersections. Accident Analysis and Prevention 40 (1), 315–326.

Kelvin, Y. K. (2004). Risk factors affecting the severity of single vehicle traffic accidents in Hong Kong. Accident Analysis and Prevention 36, 333–340.

Kim, J.-K., Kim, S., Ulfarsson, G. F., & Porrello, L. A. (2007). Bicyclist injury severities in bicycle–motor vehicle accidents. Accident Analysis and Prevention 39, 238–251.

Lee, J.-Y., Chung, J.-H., & Son, B. (2008). Analysis of traffic accident size for korean highway using structural equation models. Accident analysis and Prevention 40, 1955-1963.

Lu, X., & Pas, E. I. (1999). Socio-demographics activity participation and travel behavior. Transportation Research Part A 33, 1-18.

Milton, J. C., Shankar, V., & Mannering, F. L. (2008). Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accident Analysis and Prevention 40 (1), 260–266.

Zhou, H., et al., An extension ofthe theory of planned behavior to predict pedestrians' violating crossing behavior using structural equation modeling. Accid. Anal. Prev. (2015), http://dx.doi.org/10.1016/j.aap.2015.09.009