CHAPTER 2

ROAD TRAFFIC ACCIDENTS IN DIFFERENT COUNTRIES

2.1 INTRODUCTION

This chapter presents the literature reviewed on the factors influencing highway accidents and their relationship with the accidents. A brief review on the traffic variables and road environment causing highway accidents is presented. Then they are followed by road inventory detail survey and field traffic studies as described in literature.

2.2 MODELS BASED ON GEOMETRIC FEATURES

Shankar et al (1994) studied the effect of roadway geometrics and environmental factors on rural freeway accidents frequencies. They explored the frequency of multivariate analysis of roadway geometrics (e.g. horizontal and vertical alignments), weather and seasonal effects, using negative binomial model. They found that maximum rainfall and the number of rainy days played a significant positive role in accident occurrences.

Vogt and Bared (1998) developed accident models for two-lane rural segments and intersections using Poisson and negative binomial model. They used data form states of Minnesota and Wasington. Lane width, shoulder width, degree of horizontal curve, driveway density and road sign hazard rating were taken as segment variables. Average Daily

Traffic (ADT) from main line, minor line, and posted speed limit were taken as intersection variables. They concluded that all variables made contributions to the accident rate and right turn lanes on the main line increase the likelihood of accidents of three-legged intersections.

Karlaftis and Golias (2002) analyzed the relationship between rural road geometric characteristics, accident rates and their prediction using hierarchical tree based regression model for both rural two-lane and multilane roadways. They found that 'geometric' variables and pavement conditions variables were the two important factors affecting accident rates and concluded that lane width, pavement type and friction factors were most important variables affecting crash rates for two-lane case. Lane width seemed to increase with higher AADT and pavement condition factors seemed to increase with lower AADT due to higher speeds. For rural multilane roads they found that access control and median width became more important in reduction of accidents.

Seunglim Kang et al (2005) developed a traffic accident analysis method based on the accident risk index according to the combination of alignment elements and tested applying Geographic Information Systems (GIS). They studied the crash experiment on curves of four-lane expressway in Korea from 1996-2000. 345 numbers of curves were selected in yeongdong expressway in Korea. Curve radius, Curve length, grade and super elevation were negatively correlated with accident rate. Grade was with positive correlation.

Traffic Accident risk model

$$Y = 3.638+6.163CLR+3.74RR+2.566GR$$

where, Y = traffic accident rate, accidents/10^8 vehicle-km

CLR = accident risk for the combination of radius and curve length (CL>0)

RR = accident risk for the combination of radius, tangent length and curvature change (RR>0)

CR = accident for the combination of radius and vertical grade (GR>0)

Basyonny and Sayed (2006) made comparison of two negative Binomial Regression Techniques in developing accident prediction model (the traditional negative binomial and the modified negative binomial). They investigated the difference between the two approaches in terms of goodness of fit as well as identification and ranking of accident prone location. They used a sample of accident, traffic volume and geometric data corresponding to 58 arterials (392 segments) in British Columbia. It was concluded that modified NB approach seemed to fit the data better than TNB approach.

Yannis et al (2007) developed multilevel negative binomial models to investigate the effects of the intensification of police enforcement on the number of road accidents at national and regional level in Greece, during 1998-2002. They found that the Alcohol Enforcement improved the road safety in Greece.(Vehicle ownership, speed, persons killed).

Wong et al (2007) used Poisson and negative binomial regression to quantify the influence of traffic flow, geometric design, road environment and traffic control on the incidence of killed and severe injury crashes at 262 signalized intersections in Hong Kong during 2002 and 2003. They found that the road environment, degree of curve and presence of tram stops were significant factors for incidence and traffic volume had diminishing effect on the crash risk.

Haneen Farah (2009) developed probit regression model explaining minimum time to collision. Model developed based on drivers' passing decisions on two-lane rural highways using an interactive driving simulator. Questionnaire surveys were also conducted. They concluded that the factors related to traffic, geometric design and driver characters had significant contribution to accidents risk.

Nikiforos Stamatiadis et al (2010) focused on developing crash prediction models and Accident modification Factors (AMF) for multilane rural roads regarding lane width, shoulder width and median width and type. It was found that increasing shoulder width had a positive effect on crashes. For divided highways he concluded that the increase of shoulder width caused reduction of crashes and the wider medians reduced crashes

Quddus et al (2010) explored the relationship between the severities of road crashes and the level of traffic congestion, using ordered response models. Crash data, traffic characteristics (e.g., congestion flow and speed), and road geometry (e.g., curvature and gradient) were collected from M25 London orbital motorway between 2003 and 2006. They concluded that the level of traffic congestion did not affect the severity of road crashes on the M25 motorway. Table 2.1 Summaries the research works on models or studies based on the geometric features.

Table 2.1 Summary of Models Based on Geometric Features

Sl.No	Model / Study	Variables considered
1.	Poisson and negative binomial model (Vogt and Bared, 1998).	ADT, lane width, driveway density and degree of horizontal curve
2.	Effect of roadway geometrics and environmental factors on freeway accidents. (Shankar et al 1994).	Horizontal and vertical alignment, weather and seasonal effects.
3.	Hierachical tree based regression model. (Karlaftis and Golias, 2002).	AADT, lane width, shoulder width and pavement condition.
4.	Accident risk index using GIS. (Seunglim Kang et al. 2005)	Tangent length, radius and degree of horizontal curve.
5.	Ordered response models. (Quddus et al. 2010)	Gradient, degree of curve and congestion flow.

Most of the models based on geometric features considered ADT and concluded that ADT was positively correlated with accidents. But the study conducted by Seunglim et al and Shankar et al excluded ADT and their work considered geometric variables, weather and seasonal effects.

2.3 MODELS BASED ON TRAFFIC FLOW, SEGMENT LENGTH AND ACCESS ROAD

Sayed et al (1995) studied the method a method to identify accident prone locations (APL) based on assessment of factors that contributed to accidents with the help of British Columbia Ministry of Transportation and Highways. They argued that current methods to identify APL made no distinction between accidents resulted from road and non road related factors. They presented a computerized procedure using fuzzy pattern which excluded the accidents which were not related with road environment and concluded that the fund could be used for safety improvements effectively and thereby avoiding misallocation of funds by road authority.

Sayed and Walid Abdelwahab (1997) analyzed accident prone locations, using traditional and modified black spot programs based on empirical Bayes technique. They expressed that the problem associated with these traditional methods in that often locations would be identified as accident prone, even if the accidents occurring at a location were not correctable by road-improvement strategies and modified technique considered only road related accidents and modified procedure eliminated the locations that may not be correctable from a road- authority perspective. They concluded that modified programme identified locations with road-related accidents while traditional programme identified locations that were not truly hazardous from a road-safety authority perspective and consequently caused a misapplication of safety improvement funding.

Milton and Mannering (1998) found that average annual daily traffic led to an increase in accidents and also found that peak percent of this traffic caused to an decrease in accidents. (i.e. congestion leads to reduced accidents).They examined and found that increasing the number of lanes on a given road segment, led to more accidents and that in Eastern Washington narrower “substandard” lane widths (of less than 3.5m) reduced accidents frequency. They concluded that horizontal curve did not by itself increase accidents but were dependent upon whether large straight sections preceded the curves.

Saccomanno et al (2001) developed log linear Poisson regression and Empirical Bayesian (EB) Models for identifying Black Spots (BS) along 27 km stretch of Highway SS107 (divided into four sections, two sections of positive and negative grade straight sections and positive and negative grade curve sections) in Southern Italy. They considered section length, radius of curvature, lane width reduction, 85^{th} percentile operating speed, number of private driveways and intersections as independent variables. They worked and compared the results of the two models in identification of black spots and also they worked out the cost of taking counter measures for controlling occurrence of accidents. They found that on an average only 87% of the route length identified as a BS by the Poisson model was identified as such by the EB model. For the positive grade straight sections, the EB model yielded only about 18.6% of the BS suggested by the Poisson model. For the more dangerous negative grade curves sections, the EB model yielded about 90% of the BS section lengths identified by the Poisson model. On curve sections, more BS was identified than on straight sections, especially for negative grade sections. For these sections the EB model yielded about 17 BS sections as compared to 20 BS sections in the Poisson model. They concluded that the EB model could produce significant savings over Poisson model (up to 5% for the 25 km length of SS 107) if all BS locations would be subjected to some kind of safety measures.

Ossenbruggen et al (2001) developed logistic regression model to identify statistically significant factors that predicted the probabilities of crashes and injury crashes. Land use activity, road side design, use of traffic control devices and traffic exposure were explanatory variables. They found that village sites were less hazardous than residential and shopping sites.

Ashur (2003) evaluated the National strategic plan for Traffic safety in the United Arab Emirates. He found that the number of fatalities due to vehicle crashes had increased in the past few years. The study focused on evaluating traffic safety in the UAE. The data showed that the number of crashes was decreasing per 1, 00,000 populations and the severity of accidents were stable in the past few years. Pedestrian crashes increased in 1995-2000 period.

Xiao Qin et al (2003) used zero-inflated-Poisson (zip) to estimate models for prediction of counts for crash types (single-vehicle, multi-vehicle same direction, multi-vehicle opposite direction) as a function of the daily volume, segment length, speed limit and road way width. They found that the relationship between crashes and the daily volume (AADT) was non-linear and varies by crashes type, and is significantly different from the relationship between crashes and segment length for all crash types.

Girma Berhanu (2004) used Poisson and negative binomial regression methods to relate the discrete accident data with the road and traffic flow explanatory variables in Addis Ababa in Ethiopia. They found that the provision of median was found significant in both multiple-vehicle and pedestrian accident models. The derived relationship indicating that multiple-vehicle accidents increase by 30% as a result of high on-street parking. Total and multiple-vehicle accident risks on divided roads increased with the increase of ADT. Also they found increase of accident rates with the density of access to urban divided highways.

Jutaek oh et al (2004) developed accident prediction models for rural highway intersections. Poisson and negative binomial regression model were fitted to intersection crash data from Georgia, California and Michigan. They found that crash frequency increased with higher log (AADT) on major and minor roads for the three intersections and two accident types investigated. Their results revealed that the turning volume percentages were significant for 3-legged or T intersections for injury crashes. The number of commercial driveways on the major within 250 feet of the intersections was associated with higher crash frequencies for 3-legged stop controlled multilane intersections and signalized intersection. For 4-legged stop controlled multilane intersections, sight distance was a significant effect.

Mohamadreza Banihashemi et al (2005) built a linear optimization model to identify the best combination of improvements for different sections of a case study highway of 4.2 km two lane rural highway pre-defined set of improvement alternatives, given cost constraints. A ‘C’ code program was developed to build the test problems based on this model. A 4.2 km two-lane rural highway was selected as a case study, with three alternative designs containing five different types of improvements. Unit costs associated with these improvements were estimated. For the test problem 200 m was chosen as a minimum length, i.e., the minimum length for which the improvements are maintained. For a 3, 00,000 investment, crashes were expected to decrease to 10.0 /yr from 14.4 crashes/year, for a reduction of 4.4/yr.

Ns = $(\sum_{i=1}^{H} f\ (ADi, Li)(\prod_{si=1}^{a} AM\ Fsi)$

base form of crash prediction model

where,	Ns	=	Expected number of crashes for all highway segments
	H	=	number of homogenous segments
	ADTi	=	ADT for homogenous segment
	Li	=	Length of homogenous segment
	f (ADTi, Li)	=	A function of ADTi and Li taking into account the effect of these two parameters in predicting the expected number of crashes for segment
	a	=	Number of highway features for which these are AMFs in the model
	AMFsi	=	AMF number‘s’ for homogenous segment i
	Si	=	Index associating the AMFs to the highway features (values from perpendicular to ‘a’)

Minwook Kang et al (2005) collected accident data for four years (1993-1996) on interstate highway No. 5 and 90 in the state of Washington. They developed Negative binomial model to analyze impact of length of highway horizontal section, on accidents based on decision sight distance. They found that accidents decreased as the curve radius became larger.

Perez (2005) evaluated safety impacts of four (highway upgrading, traffic signing, pavement markings, and pavement resurfacing) engineering treatments implemented in Spain, by conducting before and after studies using empirical Bayes method. He found that highway upgrading had a positive and significant safety impact value of 33% , improvement of traffic signing had a positive but non-significant safety impact value of 11%, repainting of pavement markings in a bad state of maintenance had a positive but non-significant value of 14% and pavement resurfacing on sections with dry -weather accidents had non-significant value of -2% and on sections with wet-weather accidents had a non-significant value of 36%.

Jun_seok oh et al (2005) attempted to identify traffic conditions that might lead to a traffic accident from real time freeway traffic data. They used loop detector data in measuring the likelihood of an accident from real time traffic conditions. The likelihood that the given traffic measures, described by spaced variation, belong traffic conditions leading to an accident was estimated, by non parametric Bayesian model.

Dominique Lord and Bonneson (2006) proposed the procedure for calculating accidents per year, for a particular highway segments connecting two major intersections. Traffic volume, segment length and shoulder width were taken as explanatory variables and they found all variables were significant.

Pardillo et al (2006) refined the negative binomial model using random reduction of the sample and stratification of data in the model techniques. In random reduction of the sample technique they made models using 2500, 2000 and 1500 of 3297 original accident data of two-lane rural roads in Spain. They found that in the model fitted with the 1500 sample data effect due to redundant data was very significant. For sub samples 2500 and 2000 records there was no significant variation in the measures of goodness of fit. In the model stratification they considered four ranges of stratification of AADT < 6000 and AADT > 6000, AADT < 8000 and AADT > 8000, AADT < 6000; 6000 < AADT < 8000 and AADT>8000 and AADT < 6000; 6000 < AADT < 8000; 8000 < AADT < 12,000 and AADT > 12,000. They found that magnitude of over estimation or under estimation was considerably reduced by this technique. Variables were total traffic flow, access density (access points per km), minimum sight distance within one segments and minimum design speed of alignment elements included in the one km segment (km/hour), maximum longitudinal grade (%) and reduction in design speed. On the whole they concluded that the random reduction of the original sample resulted in a significant improvement in the fit of the predicted values to the observed values to the observed data. And in the stratified model there was a considerable decrease in the variation of the residual errors.

Akgungor and Osman yildiz (2007) did the sensitivity analysis of an accident prediction model by the fractional factorial method. The evaluation of sensitivity analysis indicated that average daily traffic, lane width, width of paved shoulder, median and their interactions had significant effects on the number of the accidents. They concluded that ADT was directly related to the number of accidents.

Ciro Caliendo et al (2007) developed crash prediction models for multilane roads in Italy. 5 years accident data from 1999-2003, were collected for 46.61m four lane median divided motorway. They considered a total of 1916 accidents 21 of which were fatal and 594 were injuring accidents. Length, longitudinal slope, pavement surface condition, sight distance curvature & AADT were considered as independent variables. The Poisson, negative binomial and negative multinomial regression models were applied separately to tangents, and curves. They found that the number of accidents per year increased with curvature & AADT and also that wet pavement was found to be highly significant factor in increasing the number of crashes.

Model equations, for curve

- Total crashes,

  $\hat{\lambda} = \exp[-0.07130 + 0.80311\mathrm{InL} + 0.27017 \times 1/R + 0.32660 \times AADT \times 10^{-4}]$

- Severe crashes,

  $\hat{\lambda} = \exp[-1.45703 + 0.86881\mathrm{InL} + 0.33793 \times 1/R + 0.40863 \times AADT \times 10^{-4}]$

where, $\hat{\lambda}$ =predicted crashes/ year and carriageway

L = curve length in km

1/R = the curvature in km^{-1}

AADT = annual average daily traffic in vehicles/day

For tangents

- Total crashes

  $\hat{\lambda} = \exp\ (0.50347)\ [\exp(0.85729\mathrm{InL} + 0.23960 \times AADT \times 10^{-4}) + 0.22848 \times AADT \times 10^{-4} \times J]$

- Severe crashes

$$\hat{\lambda} = \exp(-1.40044) \quad [\exp(0.76232 \ln L + 0.42575 \times AADT \times 10^{-4}) + 0.505628 \times AADT \times 10^{-4} \times J]$$

where, $\hat{\lambda}$ = predicted crashes/ year and carriageway

L = the tangent length in km

AADT = annual average daily traffic in vehicles/day

J = junction (1 if present, 0 if absent)

Haneen Farah et al (2007) developed two crash-prediction models based on the infrastructure coefficients and crash records. Infrastructure coefficients that reflected road alignment, road side environment, sight distance along the highway, presence of guard rails, number of percentage of access points, roadway consistency alignment, lane and shoulder width as independent variables. Two infrastructure coefficients were developed and calibrated by two statistical methods. They suggested that these models could be used to evaluate the safety level of existing or planned highways.

Yi (Grace) Qi et al (2007) formulated free way accident likelihood prediction using a panel data Analysis Approach (i.e. from different individuals, groups etc). The empirical analysis showed that factors from three categories, i.e. traffic flow characteristics, weather, and geometry, had statistically significant association with traffic accidents.

Kay Fitzpatrick et al (2010) developed a horizontal curve accident modification factor (AMF) for rural four-lane divided and undivided highways. Variables considered were driveway density, lane width, shoulder width, median width, and median type, degree of curve, segment length, and average daily traffic. Crash data were collected for 5 years (1997-2001). Models were developed for curve, tangent and both curve and tangent together. They concluded that when both driveway and non-intersection crashes were considered ADT, driveway density, median width and outside shoulder width along with degree of curve were significant. The effect of driveway density on segment crashes was similar for horizontal curve and tangent curves.

Salvatore Cafiso et al (2010) developed comprehensive accident models for two-lane rural highways using exposure geometry, consistency and context variables. They collected 279 accident data over 5-year period, with a total of 640 injured persons and 16 fatalities, in Italy.19 models were developed and they recommended 3 models. Model 1 included only the exposure variables, length and traffic volume. Model 2 included length, traffic volume (AADT) driveway density, curvature ratio and the standard deviation of the operating speed profile. Model 3 included length, traffic volume (AADT), driveway density, road side hazard rating, curvature ratio and speed differential higher than 10 kmph. They found that accident-AADT was non-linear. It was concluded that with the increasing standard deviation of speed, accidents increased and accidents decreased with increasing curvature ratio. Table 2.2 Summaries the research works on models or studies based on Traffic Flow, Segment Length and Access Road.

Table 2.2 Summary of Models Based on Traffic Flow, Segment Length and Access Road

Sl.No	Model / Study	Variables considered
1.	Accident prone locations using fuzzy pattern (Syed et al. 1995)	Road environment
2.	Log linear Poisson regression and empirical Bayesian model. (Saccomanno et al. 2001)	Lane width reduction, speed, length and radius of curve.
3.	Logistic regression model. (Ossenbruggen. 2001)	Traffic exposure, traffic control device, road side sign and land use activity.
4.	Zero-inflated-Poisson model. (Xiao Quin et al, 2003)	ADT, segment length and roadway width.
5.	Poisson and negative binomial regression model. (Girma Berhanu, 2004)	ADT, density of access roads and provision of median.
6.	Poisson and negative binomial regression model. (Jutaek et al. 2004)	AADT and commercial driveways.
7.	Safety impact of four engineering treatments. (Perez, 2005)	Highway up gradation, improvement of traffic signing, resurfacing and painting of pavement markings.
8.	Procedure for calculation of number of accidents. (Dominique Lord, 2006)	Traffic volume, segment length and shoulder width.
9.	Sensitivity analysis by fractional factorial method. (Alip Osman, 2007).	ADT, width of shoulder, lane and median.
10.	Poisson and negative binomial model. (Ciro Caliendo et al. 2007).	AADT, segment distance, sight distance, longitudinal slope and pavement surface conditions
11.	Horizontal curve accident modification factors. (Kay Fitzpatrick et al. 2010).	ADT, degree of curve, segment distance, width of lane, shoulder and median.
12.	Comprehensive accident model for two-lane road. (Salvatore Cafiso et al. 2010).	AADT, driveway density and curve ratio.

Many studies focussed on traffic flow, segment length and access roads used Poisson and negative binomial regression models which included traffic flow, road environment and segment distance except a study on accident prone locations by Sayed et al using fuzzy pattern which considered only road environment. The model developed by Ciro Caliendo et al had additionally considered sight distance while the comprehensive model developed by Salvatore Cafiso et al considered curve ratio.

2.4 MODELS BASED ON SPEED

Solomon and Cirillo (1974) used accident involved vehicles on two-lane and which four-lane rural highways and interstates as their unit of analyses. Their studies estimated the incremental deviation from the mean speed of the accident involved vehicle’s speed. They found that the lowest accident rate occurred within a speed range of 15 to 20 percent higher than the mean speed. They concluded that as deviation increased above this range, accident involvement rates increased for vehicles speeds either higher or lower than the mean speed, besides speeds were not free-flow.

Garber Gadiraju (1988) studied the relationship between speed variance and accident experience. They examined 36 roadway segments in Varginia. They analyzed and compared design speed, posted speed, mean and variance of operating speeds. Their results compared the effects of traffic and geometry on vehicle speeds, indicating that the accident rates increased with increasing speed variance for all road classes and also speed variance on highway segment tended to be minimum, when the difference between design and posted speed was between 8 and 18 kmph.

John collins et al (1998) studied the speed variability on rural two-lane highways. They found that speed variance measures cannot be used to evaluate geometric design consistency and they also found that there was low correlation between geometric features and speed variance. Further they concluded that speed variance was not an appropriate measure of design consistency for horizontal curves on rural two-lane highways.

Xuedong et al (2005) used multiple logistic regression model to study the multi vehicle rear end accident characteristics, by extracting 2001 accident database from Florida Department of Highway safety and motor vehicles (DHSMV). They found that risk of rear – end accidents for 6 – lane highways was higher than 2 – lane and 4 – lane highways and also it was found that the relative accident involvement ratio for night is apparently lower than daytime ; compared to dry road surface, the wet and slippery road surface could contribute to rear-end accidents. And also concluded that when the speed limit increased, the risk of rear-end accidents increased, especially when the speed limit was higher than 65kmph.

Letty Aarts and Ingrid Van Schagen (2006) investigated the impact of driving speed on the risk of road crashes. They found that the crash rate increased faster with a particular increase in speed on minor/urban roads than on major/rural roads. Also they concluded that lane width, junction density and traffic flow were found to interact with the speed-crash rate relation.

Goldenbeld and Schagen (2007) studied the credibility of 80 Kmph limit for different rural roads and assessed the effects of characteristics of roads and its environment as well as the effects of person and personality characteristics. They enquired about 600 Dutch car drivers. They found that large differences in both preferred speed and safe speed limit between the road scenes, both below and above the limit of 80 Kmph. They concluded that drivers preferred to drive 4-5 Kmph faster than the speed limit they considered to be safe

Chanyu Kong and Jikuang Yang (2010) investigated the association between the impact speed and risk of pedestrian casualties in passenger vehicle collision based on real world accident data in China. They developed a multiple logistic regression model considering impact speed and age and they found that risk of pedestrian fatality is 26% at 50kmph, 50% at 58 kmph and 82% at 70 kmph. At an impact speed of 80 kmph, the pedestrian rarely survived.

Surenchen and Fengchen (2010) focused on characterizing the transient process of accidents, introducing new critical variables on assessing the accident risk under more comprehensive hazardous driving conditions and establishing more realistic accident criteria. It was used to define appropriate safe driving speed limits for vulnerable vehicles under normal extreme conditions and predict potential crash and injury risk of vulnerable drivers. Table 2.3 Summaries the research works on models or studies based on Speed.

Table 2.3 Summary of Models Based on Speed

Sl.No	Model / Study	Variables considered
1.	Accidents on two-lane and four-lane. (Solomon and Cirilo, 1974)	Speed range
2.	Relationship between speed variance and accident experience. (Garber Gadiraju, 1988).	Design speed, posted speed, mean and variance of operating speed.
3.	Speed variability. (John Collins et al. 1998).	Speed variance and geometric design.
4.	Credibility of 80 kmph limit. (Goldenbeld and Schagen,2007).	Speed limit, preferred speed and drivers.
5.	Multiple logistic regression model. (Chanyu Kang and Jilkey, 2010).	Impact speed, pedestrian and age.
6.	Characterizing transient process of accidents. (Surenchen and Fengchen, 2010).	Speed limit and driving conditions.

Most of the studies with respect to models based on speed considered speed, speed variance, divers and pedestrians age except a study which considered only speed range conducted by Solomon and Cirillo.

2.5 MODELS BASED ON HEAVY VEHICLES

Wael and Bruce (1998) developed prediction models for truck accidents at free way ramps in Washington State using regression and AI techniques. They compared the results by regression model area Adaptive-networks-Based Fuzzy inference system (ANPIS) and concluded that the relationship between the criterion variable (in terms of truck accident frequency) and the explanatory variable (in terms of ramp ADT, ramp length, surface condition, weather condition etc.) was complex and could not be explained by simple relationship represented by a regression model. AI techniques were more capable of explaining such complexity compared to regression procedure.

Lynn Meuleners (2006) undertook to estimate the number of fatalities and serious injuries for heavy vehicles drivers involved in a crash in Western Australia from police records and hospitals. Capture and Recapture method was used to assess the difference and similarities in characteristics of heavy vehicle drivers from both sources. Each heavy vehicle driver involved in a crash from the police report was matched against heavy vehicle driver's hospitalization record with name, age, date of crash, vehicle type as matching fields. The estimated number of fatalities and series injuries to heavy vehicle drivers from 1st July 1999 – 31st December 2000 was 5 and 59 respectively, which was 25 and 31 % higher based on the C-R methodology than the aggregated (non-overlapping) total officially reported to the police and hospitals. Number of significant age difference ($p>0.05$) was found for drivers in a heavy vehicle crash between the two sources (37years Versus 40 years of age). However female heavy vehicle drivers were over represented in hospital records (11%) compared to the police records.

Landge et al (2006) developed Poisson regression model for traffic accidents on two-lane rural highway under mixed traffic conditions. Variables spot speed, shoulder width, traffic volume, percentage of heavy vehicles, lane width and intersection density were collected from NH-58. They found that traffic volume had positive sign and shoulder width had negative sign showing that with more shoulder width, less would be the accidents. And also they concluded that the fatality rate reduced as intersection density reduced.

Joon-Kikim et al (2007) developed a microscopic model of freeway rear-end crash risk, based on Poisson regression and modified negative binomial regression model and using Washington data. Rear-end crash was considered to be based on (1) That a lead vehicle becomes an obstacle to following vehicle and (2) that a following vehicle fails to avoid a collision given the obstacle vehicle. The variables selected were AADT, truck percentage, VMT per lane, shoulder width, horizontal curve, urban and rural area (nature of area) and posted speed limit. It was found that these variables were statistically significant at 90 % level in the model.

Janine Duke (2010) attempted to review age related safety and identify other factors that contribute to accidents experienced by heavy vehicle drivers. They used Canadian centre for occupational Health and safety and found heavy vehicle drivers younger than 27 years of age demonstrated higher rates of accident/ fatality involvement which decline and plateau until the age of 63 years where increased rates were again observed. Other contributing factors to heavy vehicles accident included: long hours and subsequent sleepiness and fatigue, employer safety culture, vehicle configuration particularly multiple trailers, urbanization and road classification. They concluded that drivers of heavy vehicles were over-involved until age 27 years, however a higher risk of accident involvement for both younger and older drivers.

Sravani Vadlamani et al (2011) identified high- risk-sites(hot-spot) for large truck crashes in Arizona and examined potential risk factors related to the design and operation of the high risk site. High-risk sites were identified using both state of the practice methods (accident reduction potential using negative binomial regression with long crash histories) and a method using property damage only equivalents (PDOE). He used crash data for 6 years (2001-2006) from Arizona. Table 2.4 Summaries the research works on models or studies based on heavy vehicles.

Table 2.4 Summary of Models Based on Heavy Vehicles

Sl.No	Model / Study	Variables considered
1.	Adaptive networks-based on fuzzy inference system. (Wael and Bruce, 1998).	Traffic flow, shoulder paving and cost / benefit ratio.
2.	Poisson regression model. (Landge et al. 2006).	ADT, percentage of heavy vehicles, spot speed, lane width and intersection density.
3.	Poisson and modified negative binomial regression. (Joon-Kikim et al. 2007).	AADT, truck percentage, VMT per lane, shoulder width and nature of area.
4.	Negative binomial regression model. (Sravani Vadlaman et al. 2006	Heavy vehicles and cross-sectional elements.

Mostly the research works in model building based on heavy vehicles considered ADT, percentage of heavy vehicles and lane width. But the model developed by Joon-Kikim considered heavy vehicles and cross sectional elements.

2.6 MODELS BASED ECONOMETRIC AND SOCIAL VARIABLES

Fridstrom and Ingebrigsten (1991) estimated a model using monthly data on traffic accidents for 18 countries in Norway. They found that extensions and improvements to the national road network do not have the expected effect of improving safety. They also found more congested roads led to fewer casualties. This study controlled for many different casual factors that also contributed to crashes.

They explored the effectiveness of key local area traffic management (LATM) treatment types in reducing traffic speeds and hence in reducing the risk of causality crashes on local roads. He provided updated data on operating speed reductions and crash reductions at various types of LATM treatments. Typical vehicle speed profiles through each treatment type were also developed. Finally a speed-based design tool prototype was developed to estimate operating speed profiles and the expected crash savings for proposed LATM schemes.

Ogden (1997) studied the safety effects paved shoulders on accidents on Rural Highways in Australia. Data were obtained on the location, condition and cost of shoulder paving projects on two lane-two way roads. Shoulder width was typically between 600 and 1200mm, with 600 or 800mm being most common. The results for this type of treatment indicated that shoulder paving was associated with a statistically significant reduction in casualty accident frequencies at sites where it was installed on two- lane two-way rural highways in Victoria. Overall, casualty accidents were reduced by 41 percent on a per vehicle kilometre basis. The main accident reduction were for accidents involving near end, overtaking-out of control, off carriage way to left, and off carriageway to right into fixed object. He estimated benefit/cost ratio shoulder paving as 2.8 times the AADT, by forming relationship between benefit/cost ratio and traffic flow.

John Collins et al (1998) studied the speed variability on rural two-lane highways. They found that there was low correlation between geometric features and speed variance. Further they concluded that speed variance was not an appropriate measure of design consistency for horizontal curves on rural two-lane highways. While an increase in speed variance might be an indicator of potential safety problems for some geometric design features or traffic situations, it was not useful in explaining safety differences between tangents and horizontal curves on two-lane highways.

Ivan et al (1999) estimated Poisson regression model for predicting both single and multivehicle highway crash rates as a function of traffic density and land use as well as ambient light conditions and time of day. Their study focused on seven rural, two lane highway segments, each one-half mile in length with varying land use patterns and where actual hourly exposure values were available in the form of observed traffic counts. Land use effects were represented by the number of drive ways of various types on each segment. Hourly exposure was represented for single vehicle crashes as the total vehicle miles travelled and volume/capacity ratio; for multi-vehicle crashes it was the product of the hourly volumes on the main highway and the roads intersection it along the study segment. For with a positive or negative effect were daytime (06:00 -19.00h, negative effect), the natural log of the segment volume/capacity ration (negative), percent of the segment with no passing zones (positive), shoulder width (positive) number of intersections (negative), and driveways (mixed effects by type). Variables for multi-vehicle crash prediction models were, daylight conditions from 10:00 – 15:00h and 15:00 -19:00h (positive), number of intersections (negative) and driveways (positive for all types). The found that the traffic intensity explained differences in crash rates even when controlling for time of day and light conditions, and that these effects were different for single and multi-vehicle crashes.

Hassan et al (2000) developed 3D(mark c software) operating speed model to evaluate highway consistency and safety by analyzing 8 km segment of highway predicted operated speed to the profile of allowable speed based on 3D sight distance and vehicle dynamics, the consistency or deficiency in highway segments was identified. They found that positive value of the difference indicated a safe design, while negative values indicated design deficiencies as the drivers travelled faster than the maximum allowable speed. They concluded that as the difference between 2D and 3D model results was significant (difference can range from +100% to -30%) the 2D analyzes was proven to be inadequate for evaluation and adopting 3D models in the design guides and commercial computer software would greatly benefit the highway designer.

Feiyuan et al (2000) studied safety benefits of intersection approach realignment on rural two-lane highways at 12 sites using empirical Bayesian Methods. They observed that multi-vehicle intersection crashes other than head-on and rear-on crashes increased at site 2012 with traffic signal and average daily traffic of more than 20,000 vehicles entering the intersection. They suggested that these factors might reduce the effectiveness of treatment for crash reduction and also they found suburban sites tended to have lower crash reduction factors than did the rural sites. They concluded that the treatment did not necessarily reduce all types of crashes but instead might have increased some. For the curve realignment improvement, run-off road crashes and head-on and rear-end crashes at the intersection were more greatly reduced than other types of crashes. Run-off crashes increased at some sites.

Hassan et al (2001) developed artificial neural network models to predict driver injury severity in traffic accidents at signalized intersection, using 1997 accident data for central Florida. The relationship between driver injury severity and driver, vehicle, roadway and environment characteristics was examined. They focused on two –vehicle accidents that occurred at signalized intersections. They used multilayer perception (MLP) and fuzzy adaptive resonance (ART) theory neural networks. They found that rural intersections were more dangerous in terms of driver injury severity than urban intersections. Also female drivers were more likely to experience sever injury than male drivers and speed ratio increased the likelihood of injury severity and they concluded that wearing a seat belt decreased the chance of sustaining severe injuries and vehicle type played a role in driver injury severity.

Noland (2001) analyzed the relationship between road infrastructure and safety a cross sectional time-series, using negative binomial fixed effect model and number of models were developed and no significant effect was found for increases in average amount of interstate lanes on fatalities. They concluded that charges in highway infrastructure had not reduced traffic fatalities and injuries and even had the effect of increasing total fatalities and injuries.

Jean-Louis Martin and Robert Quincy (2001) studied the cross over crashes at Median Strips equipped with barriers on a French Motor way network. They found 15,943 out of 44,696 accidents led to an exit from roadway into the median barrier. In 206 numbers of accidents, the median barrier was totally crossed; that is the barrier was crossed in 1.3% of accidents in which vehicles collided with the median barrier. They concluded that median barrier crossing the penetration onto the oncoming roadway was a rare events occurring in only 0.5 percent of all vehicles.

Christo Bester (2003) studied the effect of Road roughness on safety. He developed two regression models for total accident rate and single vehicle accident rate, considering shoulder width, topography, present serviceability index and total paved width as independent variables. He found that the accident rate increased with improved riding quality in terms of present serviceability Index and also the cross sectional elements, alignment (topography) and riding quality of a road section played a significant role in the safety thereof. He concluded that wider cross-section elements usually led to lower accident rates, more rugged terrain invariably led to higher accident rates, and bad riding quality led to an increase in accident rate.

Clark and Brad Cushing (2004) in US studied the effect of population density on rates of motor vehicle murtalil mortality in rural and urban areas while controlling for VMT. Urban data for traffic mortality, vehicle miles (VMT) are obtained from Federal Highway Administration for 1998-2000. Traffic mortality data were obtained from National Highway Traffic Safety Administration. Linear regression was used to estimate the effect of population density, VMT per capita southern location and presence of trauma system on mortality It was found that variation in rural mortality rate (per 100,000 population) was proportional to rural VMT per capita , but population predictors, together accounting for 91% of this variation. Variation in urban mortality rates was not affected by population density but urban rates here also higher in South. The exposure-based rural mortality rate (death per 100 million VMT) was inversely proportional to density, which along with southern location explained 41% of the variation for state to state. The presence of a state trauma system did not affect mortality after controlling for VMT and southern location, state population density was moderately strong predictior of rural but not urban traffic mortality rate.

Moore et al (2004) used 84, accident records from California Highway Patrol's First Incident Response Services Tracking System. They defined secondary accidents and developed a method for identifying secondary accidents occurring by traffic queue. They estimated the number of secondary accidents per accident ranges from 0.015 to 0.03 and that the number of accidents per incident ranged from 0.007 to 0.013.

Casaer Filip et al (2004) investigated road traffic accident in Belgium. They concluded that for both sexes risk increased significantly with age. Upto 60 % of the younger drivers categories were involved in single-vehicle crashes, while these only applied to 20% of the 65+age category. While the middle age road users are the one strongly present within the rear end collisions, the elderly(65+)were more involved in lateral (40% of them) and frontal (19% of them) collisions.

Vivian Robert and Veeraragavan (2004) conducted road safety auditing on state highway-17 in Karnataka in India linking Bangalore and Mysore. Based on rural safety audit, they attempted to develop hazard rating scores indicating the safety performance of individual Kilometre for a stretch of 23 km of the highway. They arrived at hazard ratio to be used for prioritizing the sections within the stretch for safety improvement.

Filip Vanden Bossche et al (2004) investigated the frequency and severity of road traffic accidents, in Belgium by a regression model with ARMA errors. (Auto-Regressive Moving Average). They found that introduction of the sealt belt law resulted in a 6.7% reduction of the number of accidents. Effects were revealed for 3-legged multilane intersections and signalized intersections.

Pei Liu and Hsien-Guo Young (2004) investigated 1593 traffic accidents reported at 62 signalized intersections in Taiwan during year 2000 and 2001. The results indicated that the effect of specific variable on number of intersection revealed crashes was not identical for various intersections. Most sensitive characteristics on intersection-related crashes were found to be width of fast traffic lanes, number of fast traffic lanes, median type between fast and slow traffic lanes, left-turn signal seconds, and central median type.

Kumara and Chin (2004) studied the fatal traffic accidents in Asia Pacific countries. They found that the size of road network, per capita gross national product, population and number of registered vehicles were important factors that increased accident occurrence.

Khaled Abbas (2004) developed statistical models predicting expected number of accidents, injuries, fatalities and casualties on the rural roads in Egypt. Time series data of traffic and accidents over a 10 years period for the considered roads was utilized in the calibration of these predictive models. They considered six main categories contributing accidents, namely driver related, pedestrian related, vehicle related, road related, environmental related causes and other causes. Most of highly contributing causes provided to be driver related. These included loss of control of driving wheel, over speed, sudden slowing/ stoppage. Vehicle related causes were tire burst and vehicle turn over or vehicle turn off the road. Together, these six causes contributed around 83% of all accident causes on the five roads. Driver related causes contributed around 59-73% followed by vehicle related causes contributing in the range of 23 %. Pedestrian related causes also contributed around 4% while road related causes are also in the range of 3.5%.

Rune Elvik (2006) proposed new approach to the analysis of accidents at hazardous road locations. Eight accidents were recorded at hazardous road locations; the analysis showed that five of eight accidents were pedestrian accidents. He found that the normal probability of a pedestrian accident is 0.125 (probability of non-pedestrian accident is 0.875) based on binomial trial. On the whole, the pedestrian accidents occurring at night on a wet road surface the number on accidents normally expected to bear according to all logically possible combination if values for road user group (Pedestrian (p=0.125) or other (p=0.875)), road surface condition (wet (p=0.25) or dry (p=0.75)) and presume of alcohol (yes=0.125 & no=0.875) had been estimated. It was concluded that combination of pedestrian, wet road and alcohol involved occured much more frequently than one would expect in a random sample of accidents.

Karl Kim et al (2006) explored the relationship between land use, population size, and employment by sector, economic output and motor vehicle accidents. They found positive statistically significant relationships between population size, job counts, economic output and accidents. They found majority of traffic related accidents occurred in urban areas. They developed negative binomial and Poisson regression linear model. They found that these were more than ten times the total number of accidents occurring in urban areas compared to agricultural areas. They found a positive relationship between population and accidents, population explained approximately 22.30 percentage of total variation in accidents, but 18.80 percentage of the variation in vehicle to vehicle accidents and even less for pedestrian (13.60 %) and bicycle accidents (9.20%).

Qadeer (2006) developed Poisson and binomial road accident regression prediction models for Great Britain to identify important variables for occurrence of accidents. He used accident data from 1991-2002. The variables used were day, month, year, holidays, Christmas, time, New Year, population, length of road, population density and police force. The most dangerous day and month were found to be respectively Friday and November. Christmas, New Year and holidays had fewer accidents then other days. The day of week, population, population density, length of roads and police force variables were found to explain most of the variation in daily accident occurrence.

Richard Tay et al (2009) applied logistic regression model to identifying the factors affecting the occurrence of hit-and-and run fatal crashes in California, USA. They concluded that roadway functional class, routes, traffic flow, types of roadway section, speed limit, traffic control device, lighting condition and roadway alignment contributing crashes.

Yuanchang Xie et al (2009) analysed crash injury severity using Bayesian ordered probit models. They found that senior drivers or drivers in old vehicles tended to have slightly more severe injuries. The chance for male drivers to suffer the most severe category of injuries was less than female drivers under same crash circumstances. They concluded that drunk driving was very dangerous and significantly increased the possibility of more severe injuries. Junction, icy or wet surfaces and adverse weather actually led to lower probability of suffering the most severe category of injuries.

Lovegrove et al (2010) developed community-based, micro level Collision Prediction Model (CPM) use with a regional transportation plan. The data were extracted from over 400 Greater Vancouver in British, Columbia, Canada including output from the regional transportation model. They found that VKT and Volume Capacity were the dominant variables.

Schultz et al (2010) investigated the relationship between access management and other physical roadway characteristics and safety, by linear regression analysis. Physical characteristics of road are signal spacing, unsignalized access spacing and median openings. They found that crash rates were related to signal spacing such that every signal per 1.6 km (one mile) corresponded to 0.48 crashes per MVKT (0.92 crashes per MVMT).

Additionally, road segments with adjacent commercial land use had on an average 0.77 crashes per MVKT (1.23 additional crashes per MVMT) than did segments with adjacent residential land use. Posted speed limit was found to be negatively related to crash rate, with 16 kph (10 mph) increase corresponding to a reduction of less than one (0.44) additional crash per MVKT (0.71 per MVMT). Finally the presence of a raised median corresponded to a reduction of 0.77 crashes per MVKT (1.23 crashes per MVMT).

Yannis et al (2010) did a statistical analysis for using Greek drivers attitude data collected within the scope of an extensive recent survey in 23 European Countries. A statistical analysis of perception usefulness and acceptance of new technologies by older drivers was presented. The developed ordered logit models provided insight into the human factors aspect of the introduction of advanced technologies with respect to these more sensitive segments of the driver population.

Clarke et al (2010) A sample of 2000 crashes involving drivers aged 60 years, or over was considered, from force U.K midland police forces, from the years 1994-2007 inclusive. Each case was summarized on a database including the main objective features (such as time and place), a summary narrative, a sketch plan and a list of explanatory factors. The main finding were that older drivers had significant problems with intersection collisions and failing to give right of way; these formed the largest single class of crashes in the sample possible behavioural explanations

Domonique Lord and Mannering (2010) discussed the key issues associated with crash-frequency data as well as the strength and weakness of the various methodological approaches that researchers had used to address these problems.

Deogratias et al (2010) used Ohio, US crash data for 2003 to 2007 data to investigate the odds of a motorcyclist being fatally injured in a crash and the risk factors involved. It was found that the fatality rate was highest for those who were drug impaired (15.7 %), then alcohol use (13.8%), speeding (6.2%) and no helmet use (4%). Road bends and grades had higher rates of 5.3% and 4.8% respectively. Night time crashes tended to result into a higher than average fatality rate of 4.8%. Other factors that showed higher than average fatality rate were bad weather condition (4.1%), week end crashes (3.7 %), and summer season (3.6%).

Xiugang Li et al (2011) developed accident modification factors for rural frontage road segments in Texas, using generalized additive models (GAM). They collected 123 segments on rural frontage roads in Texas and found that GAMs indicated a non-linear relationship between crash risk and changes in lane and shoulder widths for frontage roads in Texas. They extracted crash data for 1997-2001 years from Texas department of public safety. Further they concluded that lane and shoulder widths were found to be not completely independent. Table 2.5 Summaries the research works on models or studies based on Econometric and Social Variables.

Table 2.5 Summary of Models Based on Econometric and Social Variables

Sl.No	Model / Study	Variables considered
1.	Artificial neural network. (Hassan Abdel Wahab, 2001)	Male driver and female driver.
2.	Negative binomial model. (Robert Noland, 2001).	Driver and changes in cross-sectional elements.
3.	Ashur (2003)	Pedestrians
4.	Effect of population density on motor vehicle mortality. (David Clark and Cushing, 2004).	VMT and population.
5.	Poisson regression model. (Khaled Abbas, 2004).	Driver, over speed, sudden slowing/stoppage and pedestrian.
6.	Accident involved with motor-cyclists. (Deogratias, 2010).	Alcohol use, speeding, road bends and grades.

Mostly the research work involved with econometric and social variables considered population and pedestrians, while a study conducted by Ashur considered only pedestrians.

2.7 SUMMARY

The models described above have taken into account one or few causative variables. Anyone who is familiar with traffic condition and road environment in Indian context would agree that mostly the factors involved with accidents are traffic flow, segment distance, heavy vehicles, traffic speed and road environment and specifically none of the above studies included the 85^{th} percentile speed of LCVs and horizontal curves in highway. In this work the combined effect of 85^{th} percentile speed of LCVs and horizontal curves with above variables, on highway accident using multiple linear regression which could be taken as a comprehensive model for estimating highway accident scene is studied.