THE ASSOCIATION OF RAINFALL AND GEOMETRIC CHARACTERISTICS ON TRAFFIC CRASHES

Ciro Caliendo
Associate Professor – University of Salerno, Department of Civil Engineering – ccaliendo@unisa.it
Maurizio Guida
Full Professor – University of Salerno, Department of Information and Electrical Engineering – mguida@unisa.it
Alessandra Parisi
PhD – University of Salerno, Department of Civil Engineering – alparisi@unisa.it

ABSTRACT

Many studies have quantified the effects of traffic and geometric factors on the expected number of road crashes. However, crash prediction models that include also rainfall and hazardous points such as junctions or tunnels have rarely been developed. In addition, most research has paid more attention to two-lane roads rather than to multi-lane roads. Finally, as far as the authors are aware, few researchers have investigated the relationships in Italy between crashes occurring on multilane roads and the combined impact of all variables mentioned above. Thus, in this paper prediction models for estimating traffic crashes on Italian multilane roads as a function of infrastructure geometric characteristics, pavement surface conditions (wet or dry), and hazardous points (junctions or tunnels) were set up. Accident data were observed on a specific four-lane median-divided Italian motorway during an 8-year monitoring period extending between 1999 and 2006. Negative Binomial Distribution, applied separately to tangents and curves, was used to model the random variation of the number of crashes. Model parameters were estimated by Maximum Likelihood Method, and the Generalised Likelihood Ratio Test was applied to detect the significant variables to be included in the model equation. Goodness-of-fit was measured by means of both the explained fraction of total variation and the explained fraction of systematic variation. The candidate set of explanatory variables was: length (L), curvature (1/R) and the presence of point hazards such as junctions (J) or tunnels (T). Separate prediction models for total and severe crashes only were proposed. For curves it is found that the most significant variables are L, 1/R and J, whereas for tangents they are L and T. The effect of rain precipitation, examined on the basis of hourly rainfall data and assumptions about drying time, shows that with a wet pavement significant increases in the number of crashes are expected. In particular, rain considerably increases the number of accidents on curves than on tangents.

Keywords: Rural road; Negative Binomial distribution; Road geometry; Weather

1. INTRODUCTION

During the last few years several studies have been carried out to investigate the influence of road geometric characteristics and traffic flow on accident frequency. However, the combined effect of these variables with rainfall and the presence of both junctions and tunnels have been even less investigated.

There is a general perception that inclement weather is associated with more hazardous driving conditions. Various papers show that precipitation in the form of rain causes more accidents as compared to dry conditions (Brodsky and Hakkert, 1988; Fridstrøm et al., 1995; Shankar et al., 1995 and 1996; Andreescu and Frost, 1998; Edwards 1998 and 2002; Golob and Recker, 2003; Eisenberg 2004; Keay and Simmonds, 2005 and 2006; Caliendo et al., 2007). This is probably due to a reduction in the pavement-tyre friction and a restricted visibility during wet weather. But at times the findings are difficult to compare so that a negative relationship between this variable and road accidents might also be found. The reasons for this could be due to: a reduction in traffic flow and/or in operating speeds during wet weather, an analysis based on different accident type (damage-only accidents or severe accidents) and different time periods and data source. Moreover the relationship between rain and road accidents is neither simple nor obvious. In addition, the impact of weather conditions has been rarely investigated in Italy, above all on multilane roads compared to two-lane roads.

Traffic accidents occurring along roadway segments at hazardous points also do not prove to be well focused. Although a continuing series of studies has established relationships between accidents and the geometric characteristics of various forms of intersections, these works have concentrated only on accidents occurring at intersections rather than on the effects they have on roadway segments. In addition, they have examined accident frequencies at signalized and un-signalized intersections and at roundabouts and mini-roundabouts referring prevalently to two-lane roads (Maycock and Hall, 1984; Pickering and al., 1986; Hall and al., 1986; Arndt and Troutbeck, 1995; Taylor and al., 1996; Layfield et al., 1996; Kennedy at al., 1997; Vogt and Bared, 1998; Greibe, 2003; Oh et al., 2004; Washington et al., 2005; Wang, 2006), whilst other forms of intersections such as junctions of multilane roads have rarely been considered (Hadi et al., 1995).

With regard to tunnels, there is also evidence that the effect of their presence has scarcely been investigated (Amundsen and Ranes, 2000).

Junctions and tunnels could represent high-accident locations due to weaving traffic or lighting conditions, respectively, so that it is high time these elements received greater attention in accident analysis.

Consequently, there are at least three main reasons for justifying this paper. The first is motivated by the need to quantify the expected number of crashes on multilane roads as a function of road geometry, rain precipitation and the presence of junctions or tunnels.

A further reason is to identify the significant elements that tend to increase or decrease accident frequency. Such knowledge may be utilised by engineers in adjusting or designing multilane roads.

Finally, given the gap of crash predictive models for Italian multilane roads, relationships that were not known before should be developed.

Such then, is the context wherein the present work is set. The objective of the paper is to identify a specific prediction model to estimate traffic crashes on multilane roads as a function of infrastructure geometric characteristics, pavement surface conditions (wet or dry), and hazardous points (junctions or tunnels).

For this purpose, the data-base of an 8-year monitoring period on a four-lane median-divided Italian motorway was utilized. Each carriageway was divided into segments with constant horizontal curvature and longitudinal slope. The total number of crashes occurring on these homogeneous segments during the period monitored was assumed as dependent variable. The analysis was conducted separately for total and severe (fatal and injury) accidents.

The Negative Binomial distribution, applied separately to tangents and curves, was used in order to account for "overdispersion" observed in accident counts. The likelihood function was maximized to obtain the estimations of model parameters. A stepwise forward procedure based on the Generalized Likelihood Ratio Test (GLRT) was used to decide which subset of the full set of potentially explanatory variables should be included in the final model. Subsequently the goodness-of-fit of the regression model was tested by means of the explained fraction of total variation and the explained fraction of systematic variation.

2. LITERATURE REVIEW

As far as the authors are aware the point of departure of studies on the effects of wet weather on road accidents is due to Brodsky and Hakkert (1988). Their paper showed that accident risk during wet weather is greater than in dry weather. Furthermore, their results confirmed the general intuition that a higher accident risk during the occasional rains of the transitional months rather than during the persistent rains of winter is expected. The reason for this is due to the fact that during dry periods grime tends to accumulate on road surface which may be slippery when the pavement becomes wet by the first rain. Therefore, the amount of rain precipitation and the time from last precipitation are the major factors which affect crashes.

Fridstrøm et al. (1995) related road accidents to four variables, namely traffic flow, speed limits, weather and lighting conditions, for the Nordic countries. With reference to environmental conditions they also showed that weather had a significant impact on accident counts. Yet, the effect of precipitation that follows a prolonged dry period was pointed out to be particularly dangerous.

Shankar et al. (1995) explored accident frequency as a function of geometric characteristics and weather. With regard to rain, in order to evaluate the effects of adverse weather conditions the authors considered two main variables, namely maximum daily rainfall in the month and number of rainy days in the month. Maximum rainfall on any given day in the month was found to play a significant positive role in accident occurrence. This variable appeared to consider not only the effects both of rainfall intensity and potential hydroplaning but also the influence of exposure. For example, the pavement surface probably remains wet during the night or early morning when daily maximum rainfall exceeds a fixed value. The number of rainy days played a significant positive role, too. This variable seemed to represent exposure to wet pavement and lower visibility. In a subsequent paper, Shankar et al. (1996) showed that a greater probability of possible injury relating to property damage only occurred in conditions of wet pavement.

Andreescu and Frost (1998) analysed the effect of rain, mean temperature and snow on automobile accidents in Montreal. This study also confirmed that rain is associated with an increase in the number of accidents compared to dry days.

Edwards (1998) investigated the relationships between weather and road accidents in England and Wales. A surprising result was found, namely that the risk of being involved in a fatal or serious road accident was reduced when the accident occurred in wet weather compared with fine weather. However, during rainfall a greater number of vehicles were involved in minor collisions. In a further study, Edwards (2002) examined motorists' speeds in wet weather compared with non-hazardous dry conditions on the M4 motorway in South Wales. The aim of the study was to establish whether drivers compensated for the additional risks associated with rain by reducing their speed. Findings showed only a marginal reduction in speeds during wet weather on both conventional and porous asphalt surfaces. As a result more drastic action should be taken in order to achieve the desired speed reduction more especially during wet weather.

Golob and Recker (2003) studied the types of accidents that occur on freeways in Southern California. The results indicated that the type of accident was influenced by different weather and lighting conditions. In fact off-road crashes from the right lane were related to wet pavement and night conditions, whereas left-lane crashes were more associated with dry pavement and daylight.

Eisenberg (2004) investigated the relationship between precipitation and traffic crashes in the US. Interesting in this work is the fact that the analysis was conducted for two different time units (daily and monthly). Results showed that daily precipitations had a positive relationship with accident occurrence for all types of severity. A surprising negative relationship between monthly precipitation and monthly fatal crashes was instead found. The reason for this is probably due to the lagged effects of precipitation days within a month. For example, if it rained a lot yesterday, then on average, today a reduction in fatal crashes is expected. However as the severity of crashes falls (from fatal to injury and to property damage only), the relationships between crashes and monthly precipitation becomes more positive.

Keay and Simmonds (2005) investigated the effect of weather variables on traffic flow in Melbourne. They found a reduction in traffic volume in spring and winter during daytime rainfall; whilst the reduction in traffic at night-time rainfall was significant over all seasons. In a later study Keay and Simmonds (2006) showed that rainfall had an enhanced effect on the accident count as the dry spell duration increases.

Caliendo et al. (2007) studied the effects of traffic flow, infrastructure characteristics and rainfall on road accidents occurring on a specified motorway, located in the South of Italy, that connects the city of Naples and Salerno. Rain was found to be a highly significant variable affecting the expected number of both total and severe accidents. However a wet pavement increased the number of severe crashes more than damage-only crashes. Finally, with rain a higher number of crashes on curves rather than on tangents was shown.

Finally, a substantial body of research shows that rain precipitations increase traffic crashes considerably, while others point out more moderate increases and only few find no significant change when compared to dry conditions.

With regard to relationships between accidents and road intersections, the early studies are generally attributed to the U.K. Transport Research Report (Maycok and Hall, 1984; Pichering et al.1986; Hall 1986).

Maycock and Hall (1984) reported the findings of a study of accidents at four-arm roundabouts on main roads in the UK. The proposed predictive models related the accident frequency by crash type to traffic flow and roundabout geometry. With reference to entering-circulating crashes, they found that this accident type increased with traffic and decreased by increasing the curvature of path deflection.

Pickering (1986) investigated accidents at rural T-intersections. As a result, crash frequencies were related to a range of explanatory variables that included traffic flow and the effect of the geometric and other features.

Hall (1986) described the results of a study of accidents at four-armed single carriageway urban traffic signals. In the final relationships accident frequencies by crash type were related to traffic and pedestrian flows and the geometric and control characteristic of signalized intersections.

Arndt and Troutbeck (1995) investigated, in contrast, accidents occurring at roundabouts in Australia. The major point of interest of this work consists in the fact that accident frequency by crash type, besides traffic flow was also related to the 85th speed percentile rather than to roundabout geometry.

Layfield et al. (1996) produced the findings of a study of accidents at urban priority crossroads and staggered junction on urban single carriageway roads in Great Britain. Relationships were developed between accident frequency and traffic and pedestrian flows, geometry and road features.

Taylor et al. (1996) presented the findings of a study on accidents occurring at three-arm traffic signals on urban single carriageway roads. The proposed accident predictive models related accident frequency to traffic and pedestrian flows and to the characteristics both of road and traffic signal control.

Kennedy et al. (1997) developed accident predictive models for urban mini-roundabouts. With reference to entering-circulating crashes, they found that this accident type increased with traffic flow and decreased by increasing the curvature of path deflection. This result agreed with the above-mentioned study on roundabouts due to Maycock and Hall (1984).

Vogt and Bared (1998) found that accidents at three- and four-lagged intersections of two-lane rural roads increased with road traffic and average speed on major roads. Accidents also depended on the presence of right- turn lanes on major roads.

Greibe (2003) proposed accident prediction models which would describe the expected number of accidents at intersections and on road links in urban areas. Single and rear-end accidents were significantly higher in signalised intersections than in non-signalised intersections. The number of crossing accidents was, however, lower in signalised intersections.

Oh et al. (2004) developed prediction models that could be used to identify countermeasures for improving signalised highway intersections and multi-lane stop controlled highways in rural areas.

Washington et al. (2005) described the validation and calibration of motor crash models for rural intersection. The objective of this study was to make marginal improvements to a set of existing models for predicting crashes at two- and four-lane intersections. The final intent was to use the findings for the Interactive Highway Safety Design Module (IHSDM).

Wang et al. (2006) investigated accidents at four-legged signalized intersections in the Central Florida area. The results showed that crash frequencies increased with heavy traffic, larger total number of lanes, large number of phases per cycle, high speed limits and high population areas. Lower crash risks occurred at intersections with more exclusive right-turn lanes and having a partial left-turn protection phase.

The papers cited above have generally focused on accidents that occurred at different intersection types with reference to two-lane roads. But other forms of intersections such as the junctions of multilane roads have hardly been investigated. In addition, very few works have separated the accidents occurring on the ramps of junctions from those occurring on multilane road segments due to the effects of the presence of junctions. Among these, the study that appears more interesting for the objectives of our paper was produced by Hadi et al. (1995), who proposed several accident prediction models with regard both to two-lane roads and multilane roads in rural or urban areas. The dependent variables were total crash frequency, injury crash frequency or fatal crash frequency. In particular with reference to four-lane rural median divided roads, they showed that crash frequency (total or injury) increased with the section length, traffic flow and number of junctions. The model for fatal crash frequency included only the section length and traffic flow. That was probably due to the fact that the number of fatal crashes was too low to allow a correct estimation of the other variable too.

Another questionable point is traffic accidents at road tunnels. Also in this case, a review of literature has revealed that too few researchers have investigated the relationships between accidents and traffic flow, lighting conditions and tunnel geometry (single or dual tube, length, width, number of lanes, etc.). Among these, the study by Amundsen and Ranes (2000) appears more interesting for the purpose of our present work. The accident rates were found to be higher in the entrance zone of tunnels. They showed also that the severity of accidents was higher in tunnels than on the open road.

The present paper focuses on variables relating to multilane road geometry, rainfall and the effects of the presence of junctions or tunnels.

3. DATA DESCRIPTION

An 8-year monitoring period extending from 1999 to 2006 was considered during which a rural four-lane median-separated motorway was examined. The monitored motorway is located in the South of Italy, connecting the cities of Salerno and Avellino. This infrastructure is about 30.50 km long, and the horizontal alignment contains tangents and circular curves without any transition curves. Vertical alignment consists of gradients and circular curves. The traffic flow expressed as annual average daily traffic (AADT) is about 20,000 vehicles per day.

During the monitored period, crash data and rainfall data were collected. Crash data were based on police reported accidents. For each accident a series of information was recorded, including date and location of accident, horizontal alignment (tangent or curve), vertical alignment (upgrade or downgrade), weather and pavement surface conditions (dry or wet), type and severity of accident, and number of vehicles and persons involved. The database does not include accidents taking place on the ramps of junctions, on service areas or on shoulders, since such accidents are not due to motorway geometry and rainfall. Pedestrians and bicycles are forbidden to use this infrastructure. Thus no pedestrians or bicycles were involved in the accidents. Some 1,505 accidents were considered in this study, 551 of which were severe accidents (fatal and injury). A total of 588 accidents occurred on curves (215 of which were severe accidents) and 917 accidents occurred on tangents (336 of which were severe accidents). Table 1 gives accident count data observed during the 8-year monitoring period.

Table 1 Accident count data observed during the 8-year monitoring period

Year	Numbers of total (severe) crashes						
	North			South			Year's total
	Curves	Tangents	Total	Curves	Tangents	Total	
1999	21 (12)	38 (19)	59 (31)	24 (14)	30 (15)	54 (29)	113 (60)
2000	25 (15)	38 (17)	63 (32)	40 (15)	54 (20)	94 (35)	157 (67)
2001	23 (10)	28 (13)	51 (23)	57 (23)	56 (20)	113 (43)	164 (66)
2002	53 (17)	58 (21)	111 (38)	54 (18)	87 (35)	141 (53)	252 (91)
2003	22 (3)	67 (4)	89 (7)	39 (5)	45 (9)	84 (14)	173 (21)
2004	45 (16)	75 (28)	120 (44)	30 (9)	61 (18)	91 (27)	211 (71)
2005	28 (15)	80 (34)	108 (49)	42 (15)	55 (22)	97 (37)	205 (86)
2006	50 (20)	86 (42)	136 (62)	34 (8)	60 (19)	94 (27)	230 (89)
Total	268 (108)	469 (178)	737 (286)	320 (107)	448 (158)	768 (265)	1505 (551)

3.1 Horizontal and vertical alignment

Horizontal and vertical alignments were derived from a 3-D aerial survey. The Autodesk Inc. Autocad® 2000 software was applied in order to measure the geometric characteristics (e.g. length of tangents and curves, horizontal and vertical curvature, longitudinal slope). Tangents with length ranging from 0.004 to 2.251 km and horizontal curves with radii from 0.4 to 1.563 Km were computed. Furthermore gradients with longitudinal slope ranging between -4.6 % and + 4.6% were estimated and vertical curves of circular types were defined. A summary and the statistics of the above independent variables are given in Table 2.

Table 2 Summary statistics of independent variables – North and South direction carriageway

	Mean		Mode		Standard deviation		Minimum		Maximum	
	North	South	North	South	North	South	North	South	North	South
Length (km)	0.367	0.364	0.439	0.439	0.334	0.319	0.004	0.004	2.251	2.251
Longitudinal slope (%)	0.739	-0.831	3.050	-3.05	2.327	2.342	-3.95	-4.61	4.610	3.950
Curvature (km^{-1})	0.566	0.569	1.000	1.000	0.701	0.703	0	0	2.500	2.500

3.2 Junctions and tunnels

Ten junctions for each direction were located along the monitored motorway (six of which were on curves and four on tangents). Junctions had both acceleration and deceleration lanes. These elements were considered in the statistical analysis as dummy variables (J) assuming value equal to 1 if they were present in the examined trait, otherwise 0.

There were also two tunnels on the motorway, which were located only on tangents, and whose length was 0.220 km and 2.250 km, respectively. For the former tunnel a substantially higher number of accident was observed, which was probably due to the combination of certain hazardous factors such as: the tunnel being positioned on a downgrade with a great slope, the entrance zone of tunnel being preceded by a curve with a small radius of curvature, and the width of the tunnel being narrower if compared to that of an open road. Therefore a point hazard variable (T), which assumes value 1 or 0, was considered in the analysis in order to take the presence of this tunnel into account.

Another road segment characterized by a very high number of accidents was observed at the end of the studied motorway in a southern direction. That was due to the fact that this trait represented the entrance zone for both the A3 Salerno-Reggio Calabria motorway and the Salerno-Naples motorway. Therefore the weaving and maneuvering traffic, speed changes, probable queues and road factors (downgrade and small radius of curvature) caused a major concentration of accidents. As a result an additional point hazard variable (SF) assuming value 1 or 0 was also included in the analysis.

3.3 Weather conditions

The monitored motorway is located in Southern Italy where the temperature drops below freezing only very rarely, and snow and fog are but sporadically observed during the monitored period. As a result, weather conditions such as ice, snow and fog cannot be considered as causes of crashes in this study. In contrast rain precipitations are frequent. Rain precipitation data were derived from the Functional Hydro-Geological Centre of the Campania Region. They consisted of millimeters of rain per hour measured at seven weather stations located along the motorway route.

Of all the 1,505 accidents registered on the motorway in the 8-year monitoring period, 471 are reported as having occurred on wet pavement. Rain precipitation data had a maximum value of 66 mm/h and a mean value of 1.39 mm/h.

In order to evaluate a potential effect of a wet pavement on the number of crashes, the amount of time the pavement was wet during the observational period was estimated by using the hourly rainfall data and by making assumptions about drying time. For this purpose the approximate procedure developed by Caliendo et al. (2007) was used. The amount of time the pavement was wet in a year for each segment (curve or tangent) was estimated by summing up the hours of rainfall recorded by the nearest weather station. This amount of time expressed in hours was subsequently transformed into a time-equivalent number of "days with a wet pavement". Thus, accidents occurred when the pavement surface was "dry" ("wet") were associated with days with a conventional surface status "dry" ("wet"). All the remaining days, both "dry" or "wet", had zero accidents. Thus, a data set resulted which consisted of the daily number of accidents on each road section in the 8-year monitoring period, along with the conventional daily status ("dry" or "wet") of the pavement surface.

Table 3 presents summary statistics for rain data. Day counts for all road segments and the 8-year monitoring period are given as a function of two categorical variables, namely (day with no accidents; day with at least one accident) and (day with a dry pavement; day with a wet pavement). These summary statistics refer to all crashes.

Table 3 Summary statistics of rain data: day counts for all road segments from 1999 to 2006

	Wet pavement	Dry pavement	Row total
With 0 crashes	45945	449371	495316
With at least 1 crash	417	1005	1422
Column total	46362	450376	496738

3.4 Data matrix

As a consequence of the above analyses a data matrix was created containing the following column variables: the total number of accidents for the monitored period (eight years) and for each carriageway occurring on tangents or curves as dependent variable; section length (L), curvature (1/R), the presence of junctions (J) and of the above-mentioned tunnel (T) and of (SF) as independent variables.

In order to evaluate the effect of a wet pavement, a different matrix was subsequently created containing the number of accidents per day as the dependent variable and the same previous independent variables plus a dummy (0, 1) variable for "dry" and "wet" pavement conditions, respectively.

4. METHODOLOGY

The applied statistical methodology was based on the assumption that the fluctuation of accident counts, say Y_i, which occur on a road section i during the observational time interval, is a Negative Binomial (NB) random variable (r.v.) with $E\{Y_i\} = \lambda_i$ and $Var\{Y_i\} = \lambda_i \left(1 + \lambda_i / \varphi\right)$. Unlike the Poisson model, the NB model allows for the variance of accident counts to be greater than the mean, provided that that $1/\varphi > 0$. For this reason $1/\varphi$ (or φ itself) is often called the "overdispersion parameter". The objective of statistical road modelling is to estimate the expected number of accidents on a given section as a function of geometric characteristics and weather conditions. This implies defining a "regression model" where the explanatory variables (and possibly combinations thereof) act as "covariates".

Let $\mathbf{x}$ be a vector of *k* covariates and $\boldsymbol{\beta}$ a vector of *k* (unknown) coefficients. A regression model of the expected number of accidents is defined by $\lambda_i = g(\mathbf{x}_i; \boldsymbol{\beta})$ where $g(\cdot)$ denotes a certain function. Following Hauer (2004a, b), we used a regression model with both a multiplicative and an additive portion:

$$\lambda_i(\mathbf{x}_i;\boldsymbol{\beta}) = \exp(\beta_0 x_{0i})\left[\exp\left(\sum_{j=1}^{k-1} \beta_j x_{ji}\right) + \sum_{j=k}^{m} \beta_j x_{ji}\right] \qquad \text{(Eq. 1)}$$

since the effect of variables that influence the probability of accident occurrence along a significant portion of a segment is more effectively represented by multiplicative terms, whereas the influence of variables that behave as point hazards are more effectively represented by additive terms. As to the multiplicative component, the exponential choice appears to be a natural one, in that it ensures that the expected number of accidents is always a positive number. Note that model (1) is not in the class of Generalized Linear Models (GLM), since no link function exists which is linear in $\boldsymbol{\beta}$. Nevertheless, a model with only the multiplicative or the additive portion is a GLM.

In order to obtain estimates, $\hat{\boldsymbol{\beta}}$ and $\hat{\varphi}$ say, of the unknown parameters $\boldsymbol{\beta}$ and φ, the likelihood function under the governing model is maximized. In the light of the invariance property of maximum likelihood (ML) estimation, $\hat{\lambda}_i = \lambda_i\left(\mathbf{x}_i; \hat{\boldsymbol{\beta}}\right)$ is the ML estimate of the expected number of accidents per unit of time. In this paper a Fortran code was implemented which maximizes the likelihood function by using the double precision routine DBCONF of the IMSL® MATH/LIBRARY (1989).

In order to decide which subset of the full set of potentially explanatory variables should be included in the regression model, a stepwise forward procedure based on the Generalized Likelihood Ratio Test (GLRT) was used.

To measure the overall goodness-of-fit (g.o.f) of the regression model, the likelihood ratio g.o.f. statistic R_D^2 was used, which is defined by

$$R_D^2 = 1 - \frac{D^m/(N-m)}{D^0/(N-2)} \qquad \text{(Eq 2)}$$

where D^m is the "scaled deviance" of the regression model with *m* parameters, D^0 denotes the "scaled deviance" of the zero model, i.e. the model with only a constant term and an overdispersion parameter and *N* denotes the observation number. We also computed the fraction of the explained variation as compared to the "systematic" component of variation alone rather than to the total variation, which also contains the purely (inexplicable) random component. In fact, this allows computing g.o.f. measures that are equally informative, no matter how large the size of random variation may be (see Fridstrøm et al. (1995) for details).

5. RESULTS

5.1 Model for Curves

The data set for curves consists of the number of accidents registered in $n = 8$ years from 1999 to 2006 on $N = 78$ segments, 39 for each carriageway. The candidate set of explanatory variables for curves was: length (L), curvature (1/R) and presence of junctions (J). Moreover, an additional point hazard variable (SF) was associated to the above-mentioned specific section in the southern direction of the motorway. In order to take into account the presence of point hazards, the regression model with both a multiplicative and an additive component is assumed for the expected number of counts on section *i*.

5.1.1 All accident counts

The model estimated for curves has coefficients with the expected signs. The number of all accidents per carriageway occurring during the observational period increases with length, curvature, the presence of a junction (J) and in correspondence with the point hazard (SF). All of these variables were found to be highly statistically significant. The sequence of regression models and parameters developed by the stepwise procedure is shown in Table 4. A non-linear dependence between accidents count and curves length is also revealed. The fraction of total variation explained by the regression model is 27.4%. However, it is worth noting that the fraction of systematic variation explained amounts to 40.8%.

Table 4 Stepwise procedure: sequence of models and parameters for NB model (curves –all crashes)

step	φ	const	log-length	Curva-ture	J	SF	log-LKH	GLRT	P-Value
1	0.622	2.0200					983.52		
2	0.922	3.3636	1.0811				998.16	29.28	6.26E-08
3	1.352	3.0626	0.9975			4.8673	1009.10	21.88	2.90E-06
4	1.686	2.3418	1.1509	0.7089		9.8096	1014.22	10.24	1.37E-03
5	**1.937**	**1.9941**	**1.0345**	**0.7551**	**0.9254**	**13.1005**	**1017.30**	6.16	1.31E-02

The total accident-prediction model is:

$$\hat{\lambda} = \exp(1.9941)\begin{bmatrix}\exp(1.0345 * LogL + 0.7551 * 1/R) + \\ +0.9254 * J + 13.1005 * SF\end{bmatrix} \qquad \text{(Eq. 3)}$$

where $\hat{\lambda}$ is the predicted total crashes in eight years for the carriageway, L the curve length in kilometres, 1/R the curvature in km^{-1}, J the junction (1 if present, 0 if absent) and SF cited segment in the southern direction (1 or 0).

5.1.2 Severe accident counts

The previous analyses, relating to all crashes, were repeated for severe crashes only. The findings are quite similar to the former ones. However, the variables highly statistically significant are length, curvature and the point hazard variable SF.
The variable J has a P-value (0.134) which is somewhat higher. In this case J does not appear to be statistically significant at the 5% level. The sequence of regression models and parameters developed by the stepwise procedure is shown in Table 5 ($\hat{\lambda} = \exp(1.161)\left[\exp(1.4867 * LogL + 1.1131 * 1/R) + 11.3444 * SF\right]$). The fraction of total variation explained by the regression model with only the most significant variables is 36.9%, whereas the fraction of systematic variation explained amounts to 66.1%.

Table 5 Stepwise procedure: sequence of models and parameters for NB model (curves – severe crashes)

step	φ	const	Log-length	Curva-ture	J	SF	Log-LKH	GLRT	P-Value
1	0.539	1.0139					123.88		
2	0.871	2.6465	1.3417				138.61	29.46	5.71E-08
3	1.341	1.1505	1.6838	1.4210			149.54	25.86	3.67E-07
4	**2.481**	**1.1610**	**1.4867**	**1.1131**		**11.3444**	**157.95**	16.82	4.11E-05
5	2.790	0.9132	1.4302	1.1843	0.5658	14.0660	159.07	2.24	1.34E-01

5.2 Model for Tangents

The data set for tangents consists of the number of accidents registered in $n = 8$ years from 1999 to 2006 on $N = 92$ segments, 46 for each carriageway. The candidate set of explanatory variables for tangents was: length (L) and presence of junctions (J). Moreover, the additional "presence of a tunnel" (T) point hazard variable was considered, since there was observed to be a very high number of accidents concentrated in a short section of the motorway in the southern direction corresponding to a tunnel. In order to take into account the presence of point hazards, the regression model with both a multiplicative and an additive component is assumed for the expected number of counts on section *i*.

5.2.1 All accident counts

The model estimated for tangents has coefficients with the expected signs. The number of all accidents per carriageway occurring on tangents during the observational period increases with length and in the presence of a tunnel. Both these variables are highly statistically significant. The presence of a junction did not result in a significant variable, however. Note that a non-linear dependence between accidents count and tangents length is also revealed.

The sequence of regression models and parameters developed by the stepwise procedure is shown in Table 6: $\hat{\lambda} = \exp(3.142)\left[\exp(0.9401 * LogL) + 3.0824 * T\right]$ where T represents the mentioned tunnel in the southern direction with value 1 or 0. The fraction of total variation explained by the regression model is 20.7%, whereas the fraction of systematic variation explained amounts to 29.9%.

Table 6 Stepwise procedure: sequence of models and parameters for NB model (tangents – all crashes)

step	φ	constant	log-length	T	Log-LKH	GLRT	P-Value
1	0.701	2.2993			1643.53		
2	1.093	3.2242	0.8999		1665.76	44.46	2.60E-11
3	**1.407**	**3.1420**	**0.9401**	**3.0824**	**1675.84**	20.16	7.12E-06

5.2.2 Severe accident counts

The previous analyses, relating to all crashes, were repeated for severe crashes only. It was found that the number of severe accidents per carriageway occurring on curves during the observational period increases with length and in presence of a tunnel. Both these variables are highly statistically significant.

The presence of a junction did not result in a significant variable. The sequence of regression models and parameters developed by the stepwise procedure is shown in Table 7 ($\hat{\lambda} = \exp(2.015)\left[\exp(0.8254 * LogL) + 3.9763 * T\right]$). The fraction of total variation explained by the regression model is 24.9%, whereas the fraction of systematic variation explained amounts to 43.4%.

Table 7 Stepwise procedure: sequence of models and parameters for NB model (tangents – severe crashes)

step	φ	constant	log-lenght	T	log-LKH	GLRT	P-Value
1	0.810	1.2953			212.28		
2	1.254	2.0877	0.7666		229.57	34.58	4.09E-09
3	**1.933**	**2.0150**	**0.8254**	**3.9763**	**240.75**	22.36	2.26E-06

With regard to the above-mentioned fraction of total variation explained it is worth noting that the rather low observed value of R_D^2 both for curves and tangents is not surprising. This is due to the fact that other variables are not considered in the analysis. For example, unfortunately traffic data relating to each tangent or curve were not available. In addition, information for quantifying the effects of driving behaviour on accidents is still wanting.

In spite of these limitations, however, we believe that the results found here, which identify the more significant elements that tend to increase or decrease accidents, should be used by engineers in adjusting or designing multilane roads for improving safety.

5.3 Modelling rain effect

In order to evaluate the effect of rain on the number of crashes, a data set was constructed which consisted of the number of accidents occurring in each day of the 8-year monitoring period from 1999 to 2006 along with the road pavement status, as was previously illustrated. A regression analysis, based on the NB model, was then applied to curves and tangents. The candidate set of explanatory variables for curves and tangents was the same as in Sections 5.1 and 5.2 plus a dummy (0,1) variable which indicates whether in each day the conventional status of the pavement surface was “dry” or “wet”, respectively.

5.3.1 All accident counts

A total of 471 accidents occurred when road pavement surface was wet, 227 of which on curves and 244 on tangents. Thus, accidents on curves represented 48.2% of all accidents under wet conditions. A stepwise procedure was implemented and the same variables which were found to be significant on a 8-year scale were confirmed to be significant also on a daily scale. In addition, wet pavement was found to be a highly significant factor in increasing the number of crashes. In fact, when road pavement surface is wet, the expected number of crashes increases by a factor 3.72 for tangents and by a factor 6.11 for curves relative to dry surface conditions.

5.3.2 Severe accident counts

167 severe accidents occurred when the pavement surface was wet, 75 of which on curves and 92 on tangents. Thus, accidents on curves represented 44.9% of severe accidents under wet conditions. A stepwise procedure was implemented and the same variables which were found to be significant on an 8-year scale were confirmed to be significant also on a daily scale. In addition, wet pavement was found to be a highly significant factor in increasing the number of crashes. In fact, when road pavement surface is wet, the expected number of crashes increases by a factor 4.16 for tangents and by a factor 5.16 for curves relative to dry surface conditions.

6. SUMMARY AND CONCLUSIONS

This paper was first motivated by the need to quantify, for Italian multilane roads, the expected number of total and severe crashes as a function of the following variables: road geometric characteristics, rainfall and the presence of junctions or tunnels. A further interesting point was to detect the more significant elements that increase or decrease accident frequency, with a view to correcting possible deficiencies that are an important step toward improving road safety.

On the basis of the 8-year monitoring period extending from 1999 to 2006 carried out on a four-lane median-divided motorway it may be concluded that the number of both total and severe accidents per carriageway, occurring on curves increases with the length, the curvature, the presence of junctions, in correspondence with a specific road section in the southern direction of the motorway and when the road pavement surface was wet.

For tangents, the number of both total and severe accidents per carriageway, increases with length and when the road pavement surface was wet. The presence of junctions does not appear to be a statistically significant variable. A dummy variable was considered to take into account the presence of a tunnel as a point hazard. Results showed an increasing number of accidents due to the presence of a tunnel both for total and severe crashes.

It is interesting to note that a non-linear dependence between the number of accidents and segments length was found.

It is also worth noting that the presence of junctions or tunnels seems to be a critical factors affecting accident occurrence. These results suggest that engineers should pay more attention to road design when junctions and tunnels are included.

Junctions, for example, should be located on tangents rather than on curves. They should also always have acceleration and deceleration lanes. Furthermore, the length of these exclusive lanes should be sufficiently long enough to accommodate speed changes and to limit interactions among vehicles.

With regard to tunnels, these should be located on road segments having a very low longitudinal slope and should not be preceded by curves with a small radius of curvature. The width of tunnels should also be as that of an open road. Furthermore improvements consist in a good design of the entrance zone of tunnels, markings and lighting. In addition, they and should also be sufficiently distanced from junctions.

There is also some evidence that the road segment located in the southern direction at the end of the motorway studied, should also be improved. The adjustment and/or the design of new infrastructural measurements should be addressed above all both to accommodate the weaving and manoeuvring of traffic between the various motorways that are here highly concentrated, and also eliminate queues. For this purpose, road ramps should be introduced.

Rain was also found to be a highly significant variable, affecting the expected number of all (severe) accidents by a factor 6.11 (5.16) for curves and by a factor 3.72 (4.16) for tangents. These results clearly indicate that the added risk of an accident in wet pavement conditions is substantial. Besides, by comparing accident percentages it appears that rain increases the number of accidents on curves more than on tangents. That is probably due to the combined effect of rain and centrifugal force in curve. The higher number of accidents on a wet pavement suggests that appropriate countermeasures should be made. Pavement resurfacing with porous asphalt appears to be the most suitable solution for preventing many wet-skidding or hydroplaning accidents. Other measures might consist in lower speed limits in adverse weather and above all when the pavement is wet by the first rain.

However, the authors believe that the measures mentioned above should also be integrated with electronic equipment. The latter, which consist generally of variable message signs, located along roads or in vehicles could help drivers to adapt their driving behaviour to the variable conditions of traffic flow, road and weather by reducing speeds and increasing the gap distance between vehicles. Thus, research needs to be addressed to increasing actions that take into account both traffic-road-weather dynamic interaction and the development of electronic systems that in future could be the key factor in accident risk reduction. In addition, more stringent controls of the violations of Highway Code should be imposed.

REFERENCES

AMUNDESEN, F.H., RANES, G. (2000). "Studies on traffic accidents in Norwegian road tunnels". *Tunnelling and Underground Space Technology*. Vol.15, No. 1, pp. 3-11.

ANDREESCU, M.P., FROST, D.B. (1998). "Weather and traffic accidents in Montreal, Canada". *Climate Research*, Vol. 9, pp. 225-230.

ARNDT, O.K., TROUTBECH, R.J. (1995). "Relationships between roundabout geometry and accident rates". *1st International Symposium on Highway Geometric Design Practices.* Boston, Massachusetts. TRB, Washington, D C.

AUTODESK INC. AUTOCAD® 2000. Copyright© 1999 Autodesk, Inc.

BRODSKY, H., HAKKERT, A.S. (1988). "Risk of a road accident in rainy weather". *Accident Analysis and Prevention*, Vol. 20, No. 3, pp. 161-176.

CALIENDO, C., GUIDA, M., PARISI, A. (2007). "A crash-prediction model for multilane roads". *Accident Analysis and Prevention*, Vol. 39, No. 4, pp. 657-670.

EDWARDS, J.B. (1998). "The relationship between road accident severity and recorded weather". *Journal of Safety Research*, vol. 29, n.4, pp. 249-262.

EDWARDS, J.B. (2002). "Motorway speeds in wet weather: the comparative influence of porous and conventional asphalt surfacing". *Journal of Transport Geography*, vol. 10, pp. 303-311.

EISENBERG, D. (2004). "The mixed effect of precipitation on traffic crashes". *Accident Analysis and Prevention* , vol. 36, pp. 637-647.

FRIDSTRØM, L., IFVER, J., INGEBRIGTSEN, S., KUMALA, R., KROGSGÅRD THOMSEN, L. (1995). "Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts". *Accident Analysis and Prevention*, vol. 27, pp. 1-20.

GOLOB, T.F., RECKER, W.W. (2003)."Relationships Among Urban Freeway Accidents, Traffic Flow, Weather, and Lighting Conditions". *Journal of Transportation Engineering*, Vol. 4, pp. 342-353.

GREIBE, P. (2003). "Accident prediction models for urban roads". *Accident Analysis and Prevention*, vol. 35, pp. 273-285.

HADI, M.A., ARULDHAS, J., CHOW, L.F., WATTLEWORTH, J.A. (1995). "Estimating safety effects of cross-section design for various highway types using Negative Binomial regression". *Transportation Research Record*, 1500, pp. 169-177.

HALL, R.D., (1986). "Accident at four-arm single carriageway urban traffic signals". *Transportation Research Laboratory Contractor Report CR65*. Crowthorne, Berks, U.K.

HAUER, E. (2004a). "Statistical road safety modelling" – *Proceedings of the 83rd TRB Annual Meeting*, Washington, DC, USA.

HAUER, E. (2004b). "Safety models for urban four-lane undivided road segments" – *Proceedings of the 83rd TRB Annual Meeting*, Washington, DC, USA.

IMSL® MATH/LIBRARY (1989). "Fortran Subroutines for Mathematical Applications". IMSL.

KEAY, K., SIMMONDS, I. (2005). "The association of rainfall and other weather variables with road traffic volume in Melbourne, Australia". *Accident Analysis and Prevention*, vol. 37, pp. 109-124.

KEAY, K., SIMMONDS, I. (2006). "Road accident and rainfall in a large Australian city". *Accident Analysis and Prevention*, vol. 38, pp. 445-454.
KENNEDY, J.V., HALL, R.D., BARNARD S.R. (1997). "Accidents at urban mini-roundabouts". *Proceedings of the 25th PTRC Summer Annual Meeting, Seminar K.* September, pp. 117-129.
LAYFIELD, R.E., SUMMERSGILL, I., HALL, R.D., CHATTERJEE, K. (1996). "Accident at urban priority crossroads and staggered junctions". *Transportation Research Laboratory Report*, No. 185. Crowthorne, Berks, U.K.
LORD, D., PERSAUD, B.N. (2004) "Estimating the safety performance of urban road transportation networks". *Accident Analysis and Prevention*, vol. 36, pp. 609-620.
MAYCOCK, G., HALL, R.D., (1984). "Accident at 4-arm roundabouts". *Transportation Research Laboratory Report LR 1120*. Crowthorne, Berks, U.K.
OH, J., WASHINGTON, S., CHOI, K. (2004). "Development of accident prediction models for rural highway intersections". *Proceedings of the 83rd TRB Annual Meeting*, Washington, DC, USA.
PICKERING, D., HALL, R.D., GRIMMER, M., (1986). "Accident at rural T-junctions". *Transportation Research Laboratory Report*, No. 1120. Crowthorne, Berks, U.K.
SHANKAR, V., MANNERING, F., BARFIELD, W. (1995). "Effect of roadway geometric and environmental factors on rural freeway accident frequencies". *Accident Analysis and Prevention*, Vol. 27, No. 3, pp. 371-389.
SHANKAR, V., MANNERING, F., BARFIELD, W. (1996). "Statistical analysis of accident severity on rural freeways". *Accident Analysis and Prevention*, Vol. 28, No. 3, pp. 391-401.
TAYLOR, M.C., HALL, R.D., CHATTERJEE, K. (1996). "Accidents at 3-arm traffic signals on urban single-carriageway roads". *Transportation Research Laboratory Report*, No. 135. Crowthorne, Berks, U.K.
VOGT, A., BARED, J.G. (1998). "Accident models for two-lane rural roads: segments and intersections". FHWA-RD-98-133.
WANG, X., ABDEL-ATY, M., BRADY, P. (2006). "Crash estimation at signalized intersections: significant factors and temporal effect". *Proceedings of the 85th TRB Annual Meeting*, Washington, DC, USA.
WASHINGTON, S., PERSAUD, B., LYON, C., OH, J. (2005). "Validation of Accident Models for Intersections". FHWA-RD-03-037.

ACKNOWLEDGMENTS

The authors would like to thank the Italian Road Police for providing data relating to accidents occurred on the Salerno-Avellino Motorway, and the Functional Hydro-Geological Centre of the Campania Region for providing rain precipitation data.

The paper benefited from the input of three anonymous referees. Their comments proved invaluable.