ACCIDENT PREDICTION MODELS FOR ROAD NETWORKS

Gianluca DellôAcqua
Assistant Professor, Ph.D., P.Eng.
Department of Transportation Engineering ñLuigi Tocchettiò
University of Naples ñFederico IIò
Via Claudio 21, I-80125 Naples, Italy
Phone: +39 0817683934
Fax: +39 0817683946
E-mail: gianluca.dellacqua@unina.it

Francesca Russo (Corresponding Author)
Ph.D., P.Eng.
Department of Transportation Engineering ñLuigi Tocchettiò
University of Naples ñFederico IIò
Via Claudio 21, I-80125 Naples, Italy
Phone: +39 0817683372
Fax: +39 0817683946
E-mail: francesca.russo2@unina.it

ABSTRACT

This paper illustrates road safety statistical models to predict injury accidents. Since 2003 the Department of Transportation Engineering at the University of Naples has been conducting a large ï scale research program based on the accident data collection in Southern Italy. The Italian analyzed roadways in the Salerno Province are composed of multilane roadways for 242 kilometers and Major and Minor two-lane rural roads for 3,101 kilometers.

Two accident prediction models were calibrated: one is associated with two-lane rural roads and the other with multilane roadways. Explanatory variables were used including traffic flow, lane width, vertical slope, curvature change rate, roadway segments length.

Several procedures exist in the scientific literature to predict the number of accidents per kilometer per year, and a lot of relationships between accidents and explanatory variables exist basing on the multiple-variable non linear regression analyses. The accident data, presented in this manuscript, were analyzed using this procedure based on least squares method. The predicted values obtained by calibration procedure were then compared to several models presented in the scientific literature to analyze the residuals by using the t-test.

Keywords: road safety analysis, calibration and validation injurious accidents prediction models

INTRODUCTION

Data published by the *World Health Organization and the European Union* demonstrated that the main cause of death is attributed to roadway accidents which numbered 1.2 million fatalities per year round the world.

To examine the phenomenon, and to restrict its consequences, national and international research on road safety is being conducted to assess the relationship between vehicles, users and the environment

Regression equations in the scientific literature we had studied attempt to predict the number of accidents per year per kilometer and particularly significant employed variables were chosen. These were the Average Daily Traffic (ADT) and roadway length.

The accident prediction models presented here were subsequently validated using the data set which had not been included in the calibration phase. In fact prior to calibration, 10% of the total number of accidents leading to injury were excluded from the sample data in order to validate the procedure at the end.

LITERATURE REVIEW

The first crash prediction model for multilane roads was defined by Persaud and Dzbik (1). The model presented relationships between the number of accidents and the traffic flow expressed as average daily traffic (ADT) and hourly volume. The analysis was based on generalized linear models. The results demonstrate that the accident rate increases with traffic flow.

Knuiman et al. (2) examined the effect of the mean value of roadway width for four-lane roads on accident prediction models. A negative binominal function was used for data distribution. The results indicated that the accident rate decreases when roadway width increases, reaching minimal crash-rate values when the cross-section dimensions are high.

Fridstrøm et al. (3) correlated the number of crashes with four variables: traffic flow, speed limits, weather and light conditions.

Hadi et al. (4) put forward numerous crash prediction models for multilane roads and rural or urban two-lane roads.

Persaud et al. (5) presented accident prediction models which developed different regression equations for circular curves and tangent elements. All the analyses were then completely focused on two-lane roads. Crash frequency was predicted by using traffic flow and roadway features.

Persaud and Lord (6) applied the generalized regression equation (GEE) to accident data from the state of Toronto (Canada)

Some models in the scientific literature (7) can be considered as a source for other accident regression equations. To predict crash frequency, for example, some roadway variables have been used in the literature such as the following(8) (9):

$$E(m)_i = \left(\frac{ADTL}{1000}\right)^{1.242} \times LEN^{0.696} \times e^{(0.1955\times LN - 0.1775\times SHW + 0.2716\times MT2 + 0.5669\times TS - 0.1208\times PTC - 0.0918\times Y91)} \qquad (1)$$

where

- $E(m)_i$: expected accident frequencies on road section *I*,
- ADTL: annual average daily traffic (AADT) per lane,
- LEN: roadway segment length in Km,
- LN: number of roadway lanes in the analyzed section *i*
- SHW: roadway width in meters,
- MT2: binomial indicator reflecting the type of center line (0 = painted, 1 = physical barrier),
- TS: binomial indicator reflecting the presence of signs (0 = signs not present, 1= signs present),
- PTC: binomial indicator reflecting the pattern type (0 = combined, 1 = commuter),
- Y91: year 1991 (0=92, 1=91).

Martin (10), studying French interurban motorways, described the relationship between the crash rate and volume of hourly traffic (VH), and the impact of traffic on fatal crashes.

Golob and Recker (11) used linear and non-linear multi-variable regression analysis to compare the correlations between crashes and traffic flow, lighting and weather conditions. In a subsequent study, Golob et al. (12) evaluated the effects of changes in traffic flow on road safety.

E. Hauer of the University of Toronto developed statistical regression equations to predict the number of crashes per year in relation to geometric roadway features and traffic flow. The crash data were analysed using binominal negative distribution (13). An innovative aspect of this study is the introduction of an alternative instrument to measure the adequacy of accident prediction models, called the Cu.Re method (Cumulative Residuals).

Hauer (14) subsequently calibrated other models to predict crash frequency (number of crashes per year) on multilane urban roads by using variables listed as AADT (Annual Average Daily Traffic), the percentage of trucks, slope, horizontal curve length, roadway width, type and width of clear zones, danger levels of road shoulders, speed limits, points of access, and the presence and nature of parking areas. The results demonstrated that AADT, the point of roadway access, and the speed limits were the significant variables for predicting crash frequency. Statistical tests were developed by the same author to test the statistical significance of the obtained results (15) (16) (17).

Caliendo et al. (18) showed a strong correlation between the number of crashes, traffic flow, and infrastructure features. The equation-form of the accident prediction models for horizontal curves is the following:

$$\lambda = e^{\left(-1.45703+0.86881\times\ln L+0.33793\times\frac{1}{R}+0.40863\times AADT\times 10^{-4}\right)} \quad (2)$$

where

λ: predicted number of severe crashes per year
L: circular curve length in kilometers
1/R: curvature radius in km^{-1}
AADT: annual average daily vehicle traffic per day

Tarko suggested some guidelines to examine road safety (19) (20) and to realize all the operations needed to improve it.

DATA COLLECTION

The collected data of the number of accidents covered a period of three years from 2003 to 2005 and relate to the road network of the Province of Salerno in Southern Italy; all geometric features and crash rates for each roadway segment were initially analyzed using a Geographic Information System (G.I.S.).

The data were given to the Department of Transportation Engineering at the University of Naples by the Administration of the Province of Salerno. Table 1 refers to the complete analyzed database while tables 2 and 3 refer to the database used to calibrate the models.

TABLE 1 Roadway and Injurious Features of Analyzed Network

Type of road	Type of Analysis	Roadway segment length [Km]	Injurious crashes in 3 years (2003-2005)
Divided Roadways	Calibration	223	1,205
	Validation	19	51
	D.R. total	**242**	**1,256**
Undivided Roadway	Calibration	2,651	1,131
	Validation	450	510
	U.R. total	**3,101**	**1,641**
TOTAL		3,343	2,897

The initial database is filtered by removing accident data with an ADT of less than 200 vehicles/day and records with incomplete information.

The final database comprises approximately 700 records. Table 1 shows the geometric and harmful features of the Italian analyzed roadway.

Prior to calibrating injurious/fatal accident prediction models, 10% of these events were randomly extracted. These injurious/fatal accidents were subsequently used to validate the regression equations.

Tables 2 and 3 show the descriptive statistics for geometric features and I/F accidents per year per kilometer on the Italian roads with two-lane rural and urban roads (undivided roadways) and the multilane roadways (divided roadways) analyzed.

The broad variation in the ADT interval (minimum and maximum values) is due to the fact that ramps are also included in the database. The study conducted here illustrates a ñnetworkò approach in which it was deemed opportune not to exclude road sections that connect one functional sub-network and to another (now including the ramps).

TABLE 2 Descriptive Statistics of the Geometric and Injurious Features of rural and urban roads

	Length *[Km]*	***ADT*** Average Daily Traffic in vehicles/day	***Roadway width*** *[m]*	***Injurious events per year per Km***
Average	4.60	3,300.15	7.69	0.12
Standard error	0.18	148.11	0.07	0.01
Median	3.25	1,789.75	7.00	0.00
Mode	/	1,174.00	7.00	0.00
Standard deviation	4.36	3,554.75	1.75	0.28
Sample variation	18.98	12,636,264.81	3.05	0.08
Kurtosis	22.19	3.10	2.62	15.20
Asymmetry	3.20	1.78	1.38	3.49
Interval	50.02	18,855.50	11.83	2.13
Minimum	0.08	201.50	4.50	0.00
Maximum	50.11	19,057.00	16.30	2.13
Sum	2,651.21	1,900,889.20	4,427.47	71.36
Total	576.00	576.00	576.00	576.00

TABLE 3 Descriptive statistics of the geometric and injurious features of multilane roads

	Length *[Km]*	***ADT*** Average Daily Traffic in vehicles/day	***Roadway width*** *[m]*	***Injurious events per year per Km***
Average	3.82	18,904.78	7.44	1.69
Standard error	0.42	1,957.54	0.38	0.36
Median	2.44	12,696.88	7.25	0.00
Mode	/	41,675.00	11.25	0.00
Standard deviation	3.33	15,413.69	2.96	2.82
Sample variation	11.07	237,581,939.93	8.74	7.94
Kurtosis	1.82	-1.19	-0.72	12.20
Asymmetry	1.57	0.59	0.30	3.00
Interval	14.38	48,259.33	11.50	16.47
Minimum	0.52	1,247.67	3.50	0.00
Maximum	14.90	49,507.00	15.00	16.47
Sum	236.78	1,172,096.44	461.00	104.78
Total	62.00	62.00	62.00	62.00

CALIBRATION OF I/F ACCIDENT PREDICTION MODELS

The I/F crash roadway prediction models were calibrated using statistical software (Statistica ï Statsoft).

This system permitted the analysis, in an attempt to analyze regression, of the degree of correlation and some statistical parameters explaining the statistical significance of the employed variables.

Calibrating the Accident Model for Multilane Roadways

The descriptive statistics for the database employed to calibrate fatal crash prediction models for multilane roadways are briefly summarized in Table 4; they cover 223 kilometers of Italian roadways analyzed within the Salerno Province network.

The Gauss-Newton method based on the Taylor series was used to estimate the coefficients of employed variables. All the parameters included in the model are significant with a 95% confidence level.

TABLE 4 Descriptive Statistics of Variables Employed to Calibrate the Prediction Model

Variable	Average μ	Standard Deviation σ	Minimum	Maximum
ADT [vehicles/day]	17,372.24	15,746.20	166	49,507
Roadway Segment length [Km]	3.65	3.38	0.06	14.90
Injurious accidents per year [crashes/year]	6.58	11.17	0.00	56.00

The best specification of this ordinary-least-square model (OLS) was developed using 1,205 injurious crashes; the equation-form is the following:

$$y_1 = \left(\frac{ADT}{1000}\right)^{0.5761} \times 0.4087 \times Lu \qquad (3)$$

where

- y_1: number of fatal crashes per year observed on roadway segment length L_u
- *ADT*: average daily traffic in vehicles/day observed in three years
- L_u: length of the analyzed roadway segment

The adjusted coefficient of determination (ρ^2) of the model is equal to 67.3%.

Using Hauerôs procedure, a diagram of residuals was plotted based on the ADT values as shown in Figure 1. The residual is the value of the difference measured between the predicted value of fatal accidents using the model and the real value of the number of accidents surveyed on the same roadway segment. In the diagram, the residual values were placed on the y-axis, while the ADT/1,000 values are reported in the x-axis corresponding to the same roadway segments. The traffic variable was plotted on the diagram from the lowest to the highest value.

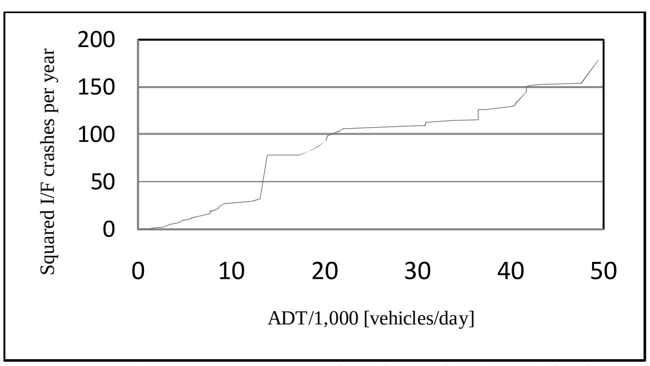

FIGURE 1 Diagram ADT-cumulated quadratic residuals.

Good predictions of accident models subsists up to an ADT value of 10,000 vehicles per day. Around this value, the residuals actually start to be significantly noticeable. Cumulated quadratic residuals corresponding to the ADT values greater than 10,000 vehicles a day show a vertical jump more correctly known as an ñoutlierò (22). The presence of an ñoutlierò indicates an observation very different from a sample data distribution and it appears when the real crash rate is very dissimilar to the predictive value using regression equations. In this case, it is necessary to carry out more investigations to decide whether to use these observations or not.

Observed residuals have a minimum value of zero and a maximum value of 6.50 injurious crashes per year per kilometer. The average value is 0.48 injurious crashes per year per kilometer, while the standard deviation is 1.66 injurious crashes per year per kilometer.

Calibrating the Accident Model for rural and urban Roadways

The descriptive statistics of some variables employed to calibrate injurious crash prediction models on these roadway types are briefly summarized in Table 5.

The database consists of 576 records and covers a total of 2,651 kilometers of Italian roads analyzed within the Salerno Province network.

TABLE 5 Descriptive Statistics of Employed Variables in the Prediction Model

Variable	Average μ	Standard Deviation σ	Minimum	Maximum
TGM [vehicles/day]	3,300.15	3,554.75	202	19,057
Section length [Km]	4.60	4.36	0.08	50.11
Injurious accidents per year [crashes/year]	0.65	2.03	0.00	25.00
V [Km/h]	58.54	10.04	30.00	83.30
Roadway width [m]	7.69	1.75	4.50	16

All the parameters included in the model are significant with a 95% confidence level. The best specification of the ordinary-least-square model (OLS) was produced using 1,131 injurious crashes. The equation-form is the following:

$$y_1 = \left(\frac{ADT}{1000}\right)^{0.6444} \times Lu \times e^{(-11.7399 + 0.1739\times V + 3.5583\times CP - 3.7087\times CT - 0.2514\times La)} \qquad (4)$$

where

- y_1: the number of fatal crashes per year observed on roadway segment length L_u
- *ADT*: average daily traffic in vehicles/day observed in three years
- L_u: length of the analyzed roadway segment
- *V*: mean value for speed in free flow conditions on a selected roadway segment in Km/h
- *CP*: slope coefficient equal to 0.8 for low slopes, 0.9 for high slopes and 1 for very high slopes
- *CT*: tortuosity coefficient of 0.8 for low tortuosity, 0.9 for high tortuosity and 1 for very high tortuosity
- L_a: roadway width in meters

The adjusted coefficient of determination (ρ^2) of the model is equal to 68.0%.

Observed residuals have a minimum value of zero and a maximum value of 1.89 injurious crashes per year per kilometer; the average value is 0.03 injurious crashes per year per kilometer, while the standard deviation is equal to 0.26 injurious crashes per year per kilometer.

Figure 2 is a CuRe (Cumulative Residuals) diagram. Figure 2 shows how the model is statistically significant until the ADT value is equal to 5,000 vehicles per day. In the diagram, the residual values were placed on the y-axis, and the x-axis gives the ADT values, which correspond to the same roadway segments. The traffic variable was plotted on the diagram from the smallest to the highest value.

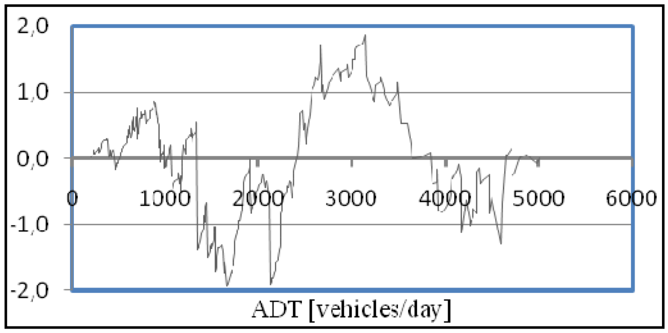

FIGURE 2 CuRe diagram .

VALIDATION OF ACCIDENT PREDICTION MODELS

This section presents the validation procedure for crash prediction models for roads with rural and urban roadways and for multilane roadways. This method evaluates the accuracy of two injurious accident prediction equations by analyzing the differences in observed and predicted values. The data used to validate the two regression equations have not been included in the calibration phase of some models. 10% of the total observed accident data was initially extracted from the entire database for subsequent use in the validation procedure.

60 roads with undivided roadway segments were used in this phase, giving a total length of 220 Kilometers and a total of 58 injurious crashes recorded in the three years from 2003 to 2005.

7 roads with divided roadway segments were then used in this phase for a total length of 18 Km and a total of 11 injurious crashes over the same three years.

The descriptive statistics of the features observed on roads with undivided roadway and roads with divided roadways which were used to validate the two prediction injurious accident models are shown in Table 6.

TABLE 6 Descriptive Statistics of Employed Variables to Validate Models

Variable	**μ**	**σ**	**Min**	**Max**	**Total**
ROADS WITH DIVIDED ROADWAYS:					
ADT[vehicles/day]	10,156.10	7,965.12	161,300	23,499.50	71,092.67
Section L [Km]	2.64	2.55	0.53	7.99	18.49
Injurious accidents per year	0.52	1.39	0.00	3.67	3.67
ROADS WITH UNDIVIDED ROADWAYS:					
ADT	3,038.11	3,120.35	230.25	14,972.00	194,439.15
Section L [Km]	3.54	3.23	0.12	16.79	226.56
Injurious accidents per year	0.30	0.82	0.00	5.00	19.33
V [Km/h]	59.09	9.39	30.00	74.30	-
Tortuosity Coefficient [-]	0.86	0.07	0.80	1.00	-
Slope Coefficient [-]	0.94	0.06	0.80	1.00	-

The validation procedure estimates the following synthetic statistical parameters:

- Residual (D_i) = value estimated from the difference between predicted fatal crash values $\underline{Y}_i$ and the observed fatal crash values Y_i: $D_i=\underline{Y}-Y_i$.
- MAD (Mean Absolute Deviation) = constant value equal to the sum of the absolute values of D_i divided by the total number (n) of observations

$$MAD = \sum_{i=1}^{n} |D_i| / n \qquad (5)$$

- MSE (Mean Squared Error) = constant value equal to the sum of D_i squared divided by n

$$MSE = \sum_{i=1}^{n} D_i^2 / n \tag{6}$$

- I = constant value equal to the square root of MSE divided by the mean of the predictive injurious crashes

$$I = \frac{\sqrt{\sum \frac{(Y_i - Y_i)^2}{n}}}{\sum \frac{Y_i}{n}} \leq 0.1 \tag{7}$$

The predictive accident model for roads with a divided roadway has a MAD value of 3.55 injurious crashes per year per kilometer and an MSE value of 17.93 squared injurious crashes per year per kilometer.

The predictive accident model for roads with undivided roadways presents a MAD of 0.48 injurious crashes per year per kilometer and an MSE of 0.72 squared injurious crashes per year per kilometer.

Thus, it can be concluded that two accident prediction models are statistically significant because the residual values are in a limited range around the mean. This was confirmed by a low value for MAD, MSE and I indicators. In particular, I reflects a good prediction when lower than 0.1. This holds for both models.

The accident prediction models for roads with undivided roadways were compared with Tarkoôs crash prediction model (19). Two models were applied to the sample data used to validate the regression equations. This procedure was carried out to verify that the residuals given by the two models may be judged statistically in different ways. A total of 64 fatal accidents were randomly extracted from the database used for the validation procedure because this is the sample size used by Tarko. The mean value of the residuals obtained from the accident prediction model reached from the Italian regressions was 0.50 injurious crashes per year per kilometer, while according to Tarkoôs model the result should have been 0.72. A t-test was then conducted to verify whether the mean value of residuals obtained from the application of our model is statistically different from the mean value of residuals for Tarko. For the Std. Dev. value of our model equal to 0.55 injurious crashes per year and for the Std. Dev. value of Tarkoôs model equal to 0.58 injurious crashes per year the two models produce similar results at the significant level of 3% (13).

RESULTS

The two injurious crash prediction models were calibrated by using the sample data observed on the roadway network of the Province of Salerno in Southern Italy.

The database used to calibrate the prediction model on the multilane roadway covered 223 kilometers where 1,205 injurious crashes occurred from 2003 to 2005.

The Gauss-Newton method based on the ordinary-least-square model (OLS) was developed to perform a final regression equation to predict the injurious accidents per year per kilometer.

All the parameters included in the model are significant to a 95% level of confidence, and the adjusted coefficient of determination (ρ^2) of the model is 67.3%. The variables used in this prediction model are the ADT value (average daily traffic in vehicles/day observed over three years) and the L_u value (length of the analyzed roadway segment).

The database used to calibrate the prediction model on the roads with an undivided roadway (two-lane rural roads and urban roads) covered 2,651 roadway kilometers, where 1,131 injurious crashes took place from 2003 to 2005.

All the parameters included in the model are significant to a 95% level of confidence, and the adjusted coefficient of determination (ρ^2)is equal to 68.0%. The model predicts the number of fatal crashes per year per kilometer and depends on the ADT value (average daily traffic in vehicles per day observed over three years), the L_u value (length of the analyzed roadway segment), the L_a value (roadway width), the V value (mean speed value in free flow conditions on a selected roadway segment), and the slope and tortuosity coefficient.

Two accident prediction models were then validated using an accident database which had not been used to calibrate the prediction models. These two accident prediction models are statistically significant because the residual values are in a limited range around the mean.

Figure 3 shows a diagram of the accident prediction models calibrated for the Roads with undivided roadways once some values have been established. In the case of Figure 3, the L_a value and V value have been established. Respecting all the general features of the roadway shown in Table 5 and the maximum value of injurious crashes per year per Km observed on the roadways analyzed as shown in Table 2, the calibrated accident prediction model on roads with undivided roadways can be used. Other ranges of values for some used variables are to be respected; in particular the V value between 40 Km/h and 65 Km/h, and L_a value must fluctuate between 8.5m and 9.5m; for a V value between 65 Km/h and 80 Km/h, the L_a value must fluctuate between 9.5m and 10.5m.

Assigning maximum value to speed V, in order to avoid an overestimated predicted value thus conferring a pragmatic sense to the results, the L_a value must belong to a specific range. There is an empirical rule following this formula:

$$La = 0.686 \times V - 40.07 \qquad V = \frac{La + 40.07}{0.686} \qquad (8)$$

According to this empirical procedure, once the minimum value of the highway width is established, the maximum speed value to insert in the model is obtained.

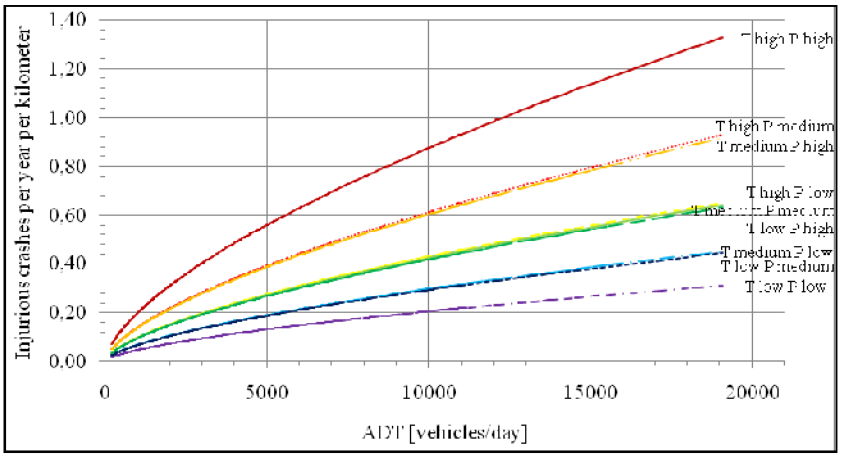

FIGURE 3 Abacus of accident prediction models on roads with undivided roadways where La=10.50 m and V=70 Km/h.

CONCLUSIONS

The proposed objective was to identify the relationship between the existing causality events among the geometric and functional characteristics of the examined network (in the Province of Salerno) and the number of recorded injurious crashes. Once the data was gathered, a database was created in order to process the data.

This project referred only to injurious crashes; a similar procedure could be used for total crash rates or only for injuries or deaths. The proposed models can be used for accident analysis on the road network and they can be arranged next to the detailed models (e.g. crashes at intersections) also through *ad hoc* investigations of specific sites. The suitability of the models to the data can be measured using many coefficients, but the adjusted coefficient of determination ρ^2 has been used in this manuscript- The functional form of the models was designed to reproduce the data; other functional forms could reproduce the data and an automated *stepwise* procedure could be implemented in this application.

The accident prediction models here presented should be improved: more information will be needed and other variables will be analyzed for this purpose while others will require further perfection.

The results obtained could be said to be satisfactory. The two prediction accident models here presented could contribute to the planning phase preceding the design of intervention. In particular, this pertains to the programming of road improvement operations (also laid out in regulations), which make it possible to target the spending of public funds by governing administrations or directing infrastructures, depending on the number of crashes predicted by the model. Especially as it concerns the evaluation and programming of road safety improvement operations, which are to be adopted in the provincial road networks, concentrating on those areas of the infrastructural network that are deemed ñcriticalò from a safety point of view.

In particular it is therefore possible to identify whether:

1. varying the traffic makes it possible to identify priority interventions.
2. to improve a road section through significant interventions.
3. to evaluate the investment related to a project that includes the insertion of a new branch in the network.

ACKNOWLEDGEMENTS
The authors would like to thank the Province of Salerno

REFERENCES

1. Persaud, B. and L. Dzbik. Accident prediction models for freeways. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1401, Transportation Research Board of the National Academies, Washington, D.C., 1993, pp. 55-60.
2. Knuiman, M.W., F. M. Council and D. W. Reinfurt. Association of median width and highway accident rates. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1401, Transportation Research Board of the National Academies, Washington, D.C., 1993, pp. 70-82
3. Fridstrøm, L., et al. Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. *Accident analysis and prevention*, Vol. 27, 1995, pp. 1-20.
4. Hadi, M. A., et al. Estimating safety effects of cross-section design for various highway types using Negative Binomial regression. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1500, Transportation Research Board of the National Academies, Washington, D.C., 1995, pp. 169-177.
5. Persaud, B., R. A. Retting and C. Lyon. Guidelines for the identification of Hazardous Highway Curves. In *Transportation Research Record: Journal of the Transportation Research Board*, No.1717, Transportation Research Board of the National Academies, Washington, D.C., 2000, pp. 14-18.
6. Lord, D. and B. N. Persaud. Accident Prediction Models With and Without Trend - Application of the Generalized Estimating Equations Procedure. . In *Transportation Research Record: Journal of the Transportation Research Board*, No.1717, Transportation Research Board of the National Academies, Washington, D.C., 2000, pp. 102-108.
7. Saccomanno, F. F., L. Fu and R. K. Roy. Geographic Information System-Based Integrated Model for Analysis and Prediction of Road Accidents. In *Transportation Research Record: Journal of the Transportation Research Board*, No.1768, Transportation Research Board of the National Academies, Washington, D.C., 2001, pp. 193-202
8. Nassar, S.A., F. F. Saccomanno and J. Shortreed. Disaggregate Analysis of Road Accident Severities. *International Journal for Impact Engineering*, Vol 15, No. 6, 1994.
9. Nassar, S. A. Integrated Road Accident Risk Model (ARM) Ph.D. thesis. University of Waterloo, Waterloo, Ontario, Canada, 1996.
10. Martin, J.-L. Relationship crash rate and hourly traffic flow on interurban roads with divided roadway. *Accident analysis and prevention*, Vol. 34, No. 5, 2002, pp. 619-629.
11. Golob, T. F. and W. W. Recker. Relationship among urban freeway accidents, traffic flow, weather, and lighting conditions. *Journal of Transportation Engineering*, 129, 2003, pp. 342-353.
12. Golob, T. F., W. W. Recker and V. M. Alvarez. Toll to evaluate safety effect of changes in freeway traffic flow. *Journal of Transportation Engineering*, 130, 2004, pp. 222-230.
13. Hauer, E. Statistical Safety Modeling. Presented at 83rd Annual Meeting of the Transportation Research Board, Washington, D.C., 2004.
14. Hauer, E., F. M. Council and Y. Mohammedshah. Safety Models for Urban Four-lane Undivided Road Segments. In *Transportation Research Record: Journal of the Transportation Research*

Board, No.1897, Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 96-105.

15. Hauer, E. The Harm Done by Tests of Significance. *Accident analysis and prevention - Elsevier*, Vol. 36, 2004, pp. 495-500.
16. Hauer, E. Statistical test of the difference between expected accident frequencies. In *Transportation Research Record: Journal of the Transportation Research Board*, No.1542, Transportation Research Board of the National Academies, Washington, D.C., 1996, pp. 24-29.
17. Hauer, E. Fishing for Safety Information in the Murky Waters of Research Reports. Presented at 81st Annual Meeting of the Transportation Research Board, Washington D.C., 2002.
18. Caliendo C., M. Guida and A. Parisi. A crash-prediction model for multilane roads. *Accident analysis and prevention*, Vol. 39, 2007, pp. 657-670.
19. Tarko, A. P. *Guidelines for roadway safety improvements*. West Lafayette, Indiana : School of Civil Engineering Purdue University, 2006.
20. Tarko, A. P. Calibrating of Safety Prediction Models for Planning Transportation Networks. In *Transportation Research Record: Journal of the Transportation Research Board*, No.1950, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 83-91.
21. Hauer, E. and J. Bamfo. Two tools for finding what function links the dependent variable to the explanatory variables. Presented at the ICTCT Conference, Lund, Sweden, 1997.