Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia - Social and Behavioral Sciences 48 (2012) 2454 – 2463

Procedia
Social and Behavioral Sciences

Transport Research Arena– Europe 2012

Integrating real-time traffic data in road safety analysis

Zoi Christoforou[a,*], Simon Cohen[b], Matthew G. Karlaftis[c]

[a]*Department 'City, Environment, Transportation', Ecole Nationale des Ponts et Chaussees, 6&8 av.Blaise Pascal, Cite Descartes, Champs-sur-Marne, F-77455 Marne-la-Vallee cedex 2, France*
[b]*Laboratoire Génie des Réseaux de Transport et Informatique Avancée (GRETIA), Institut National de Recherche sur les Transports et leur Sécurité (INRETS)- «Le Descartes 2», 2 rue de la Butte Verte, 93 166 Noisy-Le-Grand, France*
[c]*Department of Transportation Planning and Engineering, School of Civil Engineering, National Technical University of Athens, 5, Iroon Polytechniou, 15773 Zografou Campus, Greece*

Abstract

Traffic data aggregation has been a serious factor of inaccuracy in most road safety studies. The Average Annual Daily Traffic (AADT) has been the most commonly used measure to reflect traffic conditions. In this paper, we establish a framework for the integration of real-time traffic data in road safety analysis. To this end, we explore the effects of traffic parameters on type of road crash and on the injury level sustained by vehicle occupants. Univariate and ordered Probit models are specified on 4-years of data from the A4-A86 highway section in the Ile-de-France region, France. Empirical results indicate that multi-vehicle crashes tend to occur under low or very high traffic speeds, while single-vehicle crashes appeared to be largely geometry-dependent. Increasing traffic volume was found to have a consistently positive (i.e. decreasing) effect on injury severity, while speed appears to have a differential effect on severity depending on flow conditions. Also, while in higher traffic volumes higher traffic speeds aggravate severity outcomes, in lower traffic volumes speed does not significantly influence severity in a consistent pattern.

Keywords: road safety; real-time traffic data; Probit; crash type; severity

* Corresponding author.
E-mail address: zoic@civil.ntua.gr.

doi:10.1016/j.sbspro.2012.06.1216

1. Main text

1.1. Introduction

Road accidents are an issue of major concern for all countries independently to their level of development; highly developed countries have 60% of the total motor vehicle fleet but they contribute only to 14% of the global road accident deaths (Jacobs et al, 2000). An estimated 1.2 million of people are killed worldwide each year, while another 50 million are injured (WHO, 2004). WHO estimates that road accidents will become the third leading cause of death by the year 2020 (after heart disease and deaths linked to mental illness) if no effective actions and efficient measures are taken. Apart from the loss of lives, accidents have multiple collateral effects: material damage, environmental damage, pain to society, loss of productivity, impact on freight transport, health costs, delays, congestion, and so on. All these effects correspond to a certain cost, which cumulatively results to be extremely high.

Research on road safety has attracted considerable research interest in the past three decades. Major factors known to affect safety are driver characteristics, vehicle features, exposure to risk (e.g., traffic volumes), traffic control, weather conditions, and roadway design characteristics. To predict the safety of transportation systems, traffic engineers model crash rate or frequency as a function of the above mentioned factors. These measurable factors do not completely explain accident occurrence and, so, models typically used are stochastic models including a disturbance error term. Despite their frequent application, the ability of such models to reliably identify important accident predictors is open to question (Davis, 2004). Regardless of modeling techniques, a serious factor of inaccuracy – in most past studies – has been data aggregation (Lord and Mannering, 2010) and sample size insufficiency (Pande and Abdel-Aty, 2006).

Traditionally, models were macroscopic in nature, where researchers mainly used summary statistics rather than microscopic measures to develop the models. The Average Annual Daily Traffic (AADT) has been most commonly used to reflect prevailing traffic conditions (Kim et al., 2006; Mouskos et al., 1999; Qin et al., 2004). AADT is an aggregate measure of exposure; the use of AADT to approximate vehicle kilometers traveled at a site might reduce the natural variance that exists in exposure data and may result in heavy underdispersion (Pasupathy et al., 2000). Later, many authors used aggregated data over shorter periods of time (month or day) for developing the same models; others used deduced hourly traffic characteristics by combining AADT and a 1-day hourly traffic profile for the site analyzed (Ivan et al., 2000). Nevertheless, even hourly measures cannot consider the short-term variation of traffic flow and are rather not well suited for application to real-time operations. As most freeways are equipped with continuous surveillance systems, disaggregate traffic data collection is possible as well as readily available. Disaggregate traffic data have been used in only a limited number of studies (Abdel-Aty et al., 2007; Kockelman and Ma, 2007; Lee et al., 2003; Madanat and Liu, 1995). While detailed vehicle movement data in a section would be the best data source, traffic data from several consecutive detectors in a section can be a good surrogate to identifying traffic dynamics that may lead to accidents (Oh et al., 2001).

Real-time traffic data have little been utilized in road safety analyses. The exploration of the influence of real time traffic variables on accident patterning (in terms of crash type) could provide significant insight in the accident mechanism of occurrence, while proving highly beneficial to traffic and incident managers. Pande and Abdel-Aty (2006) underlined the importance of by-crash-type analysis, particularly when it comes to real-time risk assessment. They suggested that the conditions preceding crashes are expected to differ by type of crash and, therefore, any approach towards proactive traffic management should be type-specific in nature. Real-time data integration to accident severity analyses offers the possibility to associate accident attributes to the actual traffic flow characteristics at the time of the accident. Based on the analysis of historical data, typical traffic patterns recorded prior to accidents may then act as real-time identifiers (Abdel-Aty and Pande, 2007). Such explorations are useful for both researchers and practitioners in estimating accident and congestion external costs and in transportation planning. Further, such analyses may enable practitioners and authorities to locate hazardous – on severity grounds – spots on the road networks. Finally, they may provide additional insight regarding the factors that contribute to higher probabilities context for severe injuries (given that an accident occurs).

The objective of this paper is to establish a framework for the integration of real-time traffic data in road safety analysis. To this end, we explore the effects of traffic parameters on type of road crash and on the injury level sustained by vehicle occupants. Multivariate and ordered Probit models are specified on 4-years of data from the A4-A86 highway section in the Ile-de-France region, France. We use a disaggregate approach in which the units of analysis are the crashes themselves (rather than aggregations of crashes over time), and traffic data are measurements of volume, speed, and density over 6-minute intervals.

1.2. Empirical Setting

To explore the factors that determine accident occurrence by crash type, the A4-A86 highway section from a dense urban area a few miles to the east of Paris was selected (Figure 1). The A4-A86 junction has a length of 2.3 kilometers and includes four lanes per direction (to and from Paris). In particular, we used measurements from 3 stations per direction situated at kilometers 5.50, 6.00, 7.05 (direction to Paris) and 5.50, 6.14, 7.03 (direction from Paris). The A4-A86 junction is a particular site as it is the point where the Ile-de-France Ring Road (Périphérique-A86) coincides with the Autoroute de l'Est (A4) and merging is prevalent; five lanes are reduced to four on each direction. All stations are situated on the common part of the two highways; i.e. after the merging and before their separation.

Accident data were extracted from B.A.A.C. (Bulletins d'Analyse des Accidents Corporels) along with the Verbal Proceedings of the Accidents from a previous INRETS study (Aron and Seidowsky, 2004). The BAAC files provide a wealth of useful information such as crash type for all accidents, location and time, lighting conditions, and infrastructure characteristics such as road curvature and alignment. Detailed weather data are available on a 30-minute basis. We extracted such data directly from the closest meteorological station and for the 30-minute interval into which the reported time of the accident occurred. In total, 381 accidents were recorded during the period 2000-2002 and 2006. We statistically checked and found no significant difference between the 2000-2002 and 2006 registrations.

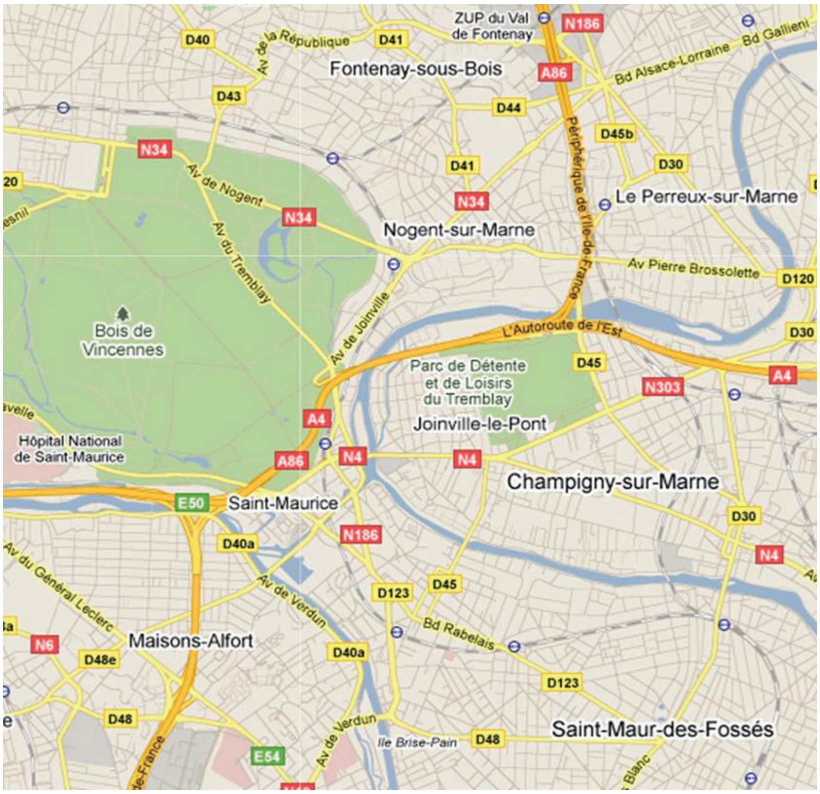

Fig. 1. The A4-A86 junction

Traffic data (flow, speed, occupancy) were provided as part of the same INRETS study (Aron and Seidowsky, 2004), and cover the period 2000-2002 and 2006; data are recorded on 6-minute intervals. Such intervals may be too large to capture short-term variations; however, data averaged on shorter intervals are not available. For each 6-minute period, the traffic database provides a series of speed, volume and occupancy measurements for each lane. The recorded traffic volume and speed – used in the thesis – were for the six-minute period ending 6 minutes before the accident (from the closest downstream detector). This time lag was used to avoid the impact of the crash itself on the traffic variables and as a buffer to compensate for any ‘inaccuracies’ in the exact time of the accident. For example, if an accident occurred at 9:00h, the traffic data considered were obtained from the 8:48-8:54 period. Similar techniques have been applied in other real-time data analyses (e.g., Abdel-Aty et al., 2007).

Loop detectors often suffer from problems that may result in unreasonable values for speed, volume, and occupancy. We reviewed all data sequences based on time series deviations, deviations across lanes, and logical rules derived from reasonable volume, occupancy and speed relationships. Aberrant values (e.g. speed>200 km/h or speed>0 along with flow=0) were discarded from the database. Accidents with traffic data unavailability were also discarded. Each observation in the dataset is a record of the crash type of each accident, the corresponding traffic conditions, and various external factors.

1.3. Methodology

The general specification for a univariate Probit model (for an event n resulting in an outcome i) can be expressed as (Washington et al., 2003):

$$Y_{in} = \beta_i X_{in} + \varepsilon_{in} \qquad (1)$$

Where:

Y_{in} defines an unobserved variable representing the latent utility (or propensity) for alternative i,
X_{in} is a vector of observed characteristics determining the outcome of the event *n*,
β_i represents a vector of unknown coefficients to be estimated for the alternative *i*, and
ε_{in} represents a vector of error terms.

If further assumed that ε_{in} follows the normal distribution and that *i=1,2*, we obtain the specification of the binomial Probit model (Washington et al., 2003):

$$P_n(1) = P(\beta_1 X_{1n} - \beta_2 X_{2n} \geq -\varepsilon_{1n} + \varepsilon_{2n}) \qquad (2)$$

Equation 2 estimates the probability of occurrence of alternative *1* for event *n*. The terms $\varepsilon_{1n}, \varepsilon_{2n}$ are normally distributed with zero mean and variances σ_{12} , σ_{22}.

Any road accident can be regarded as an event whose outcome is the type of crash that finally occurred (rear-end, side-swipe etc.). The binomial Probit model of Equation 2 can be used for estimating factors contributing or preventing a specific crash type versus all other types. Under this assumption, Equation 2 provides the probability of occurrence of a crash type (alternative 1) for each of the n accidents.

The severity function determining the severity level for each individual (*in*) can be defined as follows (Greene, 2003):

$$y^*_{in}=\beta X_{in}+\varepsilon_{in} \qquad (3)$$

Where:

y^*_{in} denotes the latent injury risk propensity for each individual involved in a road accident,
X_{in} is a vector of the independent variables considered, β is the vector of estimable coefficients, and
ε_{in} is a random error term assumed to follow the standard normal distribution across individuals.

It is reasonable to assume that unobserved values of injury y^*_{in} correspond to observed values of injury y_{in} as follows (Greene, 2003):

$$y_{in}= \begin{cases} 0 \ \ if -\infty<y_{in}*<\mu_1 & \text{(no injury)} \\ 1 \ \ if\ \mu_1<y_{in}*<\mu_2 & \text{(slight injury)} \\ 2 \ \ if\ \mu_2<y_{in}*<+\infty & \text{(severe or fatal injury)} \end{cases}$$

Thresholds μ_1, μ_2 ($\mu_1< \mu_2$) are constant and to be estimated along with the β. Then, the predicted probability of the injury level *l* (*l = 0,1,2*) for given X_{in} is given by (Greene, 2003):

$$\hat{P}(y_{in} = 0|X_{in}) = F(\widehat{\mu_1}\_X_{in}\hat{\beta})$$
$$\hat{P}(y_{in} = 1|X_{in}) = F(\widehat{\mu_2}\_X_{in}\hat{\beta}) - F(\widehat{\mu_1}X_{in}\hat{\beta}) \quad (4)$$
$$\hat{P}(y_{in} = 2|X_{in}) = 1 - F(\widehat{\mu_2}\_X_{in}\hat{\beta})$$

In the present analysis, the road user is the unit of analysis (and not the specific accident). Since each individual has specific characteristics that may influence the severity outcome differentially, there is a possibility of (additional) heterogeneity in the model. However, in the standard ordered probit model the distributional assumption does not allow for additional heterogeneity between individual realizations; the random parameter models allow the influences of variables affecting accident injury-severity proportions to vary across observations. This is achieved by adding an error term that is correlated with the unobserved factors in ε_{in} and translating individual heterogeneity into parameter heterogeneity as follows (Greene, 2003):

$$\beta_{in}=\beta + \varphi_{in} \quad (5)$$

where φ_i is a randomly distributed term.

The severity function now becomes (Greene, 2003):

$$y^*_{in}=\beta X_{in}+\varepsilon_{in}' \quad (6)$$

where ε_{in}' is the new error term (Greene, 2003):

$$\varepsilon_{in}'=X_{in}\,\varphi_{in}+ \varepsilon_{in} \quad (7)$$

1.4. Empirical Results

Separate binomial Probit models were applied for each crash type considered; parameters β of Equation 1were estimated using maximum likelihood. Univariate estimation results for crash type are presented in Table 1.

Table 1. Model estimation for crash type models

	Dependent Variables									
	rear-end with 2 vehicles		sideswipe with 2 vehicles		rear-end with > 2 vehicles		other multiple collisions		single-vehicle crash	
Independent Variables	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
constant	1.07	1.46	-1.48	-5.70	-1.79	-5.00	-1.28	-4.67	-0.78	-4.03
speed (km/h)	-0.17	-2.21					0.01	1.68		
density (veh/km)	-0.22	-2.47			0.08	1.79				
volume (veh)			0.01	1.92						
nighttime	-0.32	-1.51					-0.41	-1.75		
holiday					0.44	1.76				
gradient			0.03	1.54			-0.47	-1.75	-0.62	-2.01
curve									-0.51	-2.13
observations	235									

p-value (x^2 test)	0.040	0.068	0.007	0.039	0.039

Table 1 provides model estimation results for rear-end crashes involving two vehicles; significant factors were the average speed for all lanes and over 6 minutes (in km/h), the average traffic density per lane over 6 minutes (in veh/km), and lighting conditions. In particular, 2-vehicle rear-end crashes seem to be more probable (compared to all other crash types) during daytime compared to nighttime. Regarding traffic flow parameters, rear-end crashes were found to be more probable for lower values of both density and average speed. Sideswipe crashes involving two vehicles were found to be positively associated with both per lane and over 6 minutes (in vehicles) and road gradient. This suggests that the probability of occurrence of such a sideswipe crash (versus all other crash types) increases on 'non-flat' road segments and for high volumes of traffic. The occurrence of rear-end crashes involving more than two vehicles was found to be positively associated with average traffic density per lane over 6 minutes (in veh/km), and type of day (whether holiday/Sunday or working day). In particular, rear-ends involving more than two vehicles are more probable to occur (compared to all other crash types) on Sundays and on holidays and for high levels of traffic density. Multi-vehicle collisions other than rear-ends seem more probable to occur at high speeds, during daylight conditions, and on flat road segments. Single-vehicle crashes seem to be more probable (compared to all other crash types) on straight and flat road segments.

The probit specification of Equation (*7*) was estimated using simulation-based maximum likelihood, as maximum likelihood estimation of the random parameters ordered probit model is computationally cumbersome (Anastasopoulos and Mannering, 2009). Halton draws were used to estimate the parameters that maximized the simulated log-likelihood function and normal, triangular and uniform distributions were considered for the functional form of the parameter density function. The standard deviation of the parameter distribution was significantly different from 0 for all but two of the variables included in the final model; average speed in low traffic conditions and traffic volume was found to have fixed parameters across the population of road users. The normal distribution was found to provide the best statistical fit for the density function of the random parameters. Model estimation results are shown in Table 2. Results indicate that the random parameters model significantly outperforms the fixed parameters model based on both the log-likelihood at convergence ($LL(\beta)$) and the overall which improve noticeably when moving from the fixed to the random parameters specification (Washington et al., 2003).

Table 2. Model estimation for injury severity models

Independent variables	Fixed Parameters		Random Parameters	
	coefficient	t-statistics	coefficient	t-statistics
constant	0.426	0.33	2.429	6.48
holiday	-0.737	-7.71	-1.680	-9.62
nighttime	0.930	10.17	1.544	10.18
wet road surface	-0.508	-5.32	-1.064	-6.39
curve	0.629	7.09	0.764	5.59
'*Ancienw1*'	-0.009	-8.69	-0.067	-3.46
2-wheels	0.627	-6.74	1.805	-10.79
heavy vehicle	-0.626	6.74	-0.481	1.69
'*VQ1*'	0.017	1.26	0.015	8.59
volume (veh)	-0.001	-0.83	-0.012	-6.96
Thresholds				
μ_1	0.988	22.36	1.93	13.52
μ_2	2.011	33.13	4.33	16.65
Number of observations		893		

Log-likelihood with constant only LLI	-1207.487	
Log-likelihood at convergence LL(β)	-1045.860	-590.441
ρ_c^2=(LLI-LL(b))/LL(c)	0.134	0.511

The type of the day (working day or weekend/holiday) on which the accident occurred was found to have an effect on the severity outcome. In particular, accidents occurrence on weekdays seems to increase the severity outcome sustained by most vehicle occupants, but this effect varies across the population of vehicle occupants. The lighting conditions were also found to be significant in the injury outcome. Specifically, daylight conditions were found to exercise a positive (i.e. decreasing) and variable influence on the severity outcome of road users involved in accidents. Even though weather conditions were not found to determine severity, road surface conditions seem to significantly affect it. Normal road surface conditions (vs. wet) significantly increase accident severity, but their effect is not the same across different individuals. Further, as anticipated, vertical road curvature was found to increase the injury severity of individuals involved in crashes. The variable capturing the combined effect of rainy weather and drivers' experience ('ancienw1') was defined to explore the possible association between severity and the behavior of inexperienced drivers (holding recent driving licenses) under adverse weather conditions. Indeed, it was found that under severe weather conditions such as snow, severity levels increase for 'recent driver licenses' (inexperienced drivers). Indeed, it was found that under severe weather conditions such as snow, severity levels increase for 'recent driver licenses' (inexperienced drivers). This finding implies that under fine weather conditions, there is no significant effect of the driver's experience on severity outcome. On the contrary, increased experience causes a significant reduction in the probability of severe accidents in rainy weather. Two of the explanatory variables examined refer to the type of vehicle in which the road user was travelling at the time of the accident. Results indicate that practically all 2-wheels riders have significantly higher probabilities of getting severely injured if involved in accidents. In contrast, practically all heavy vehicle drivers and passengers (all distribution above zero) have significantly lower probability of suffering a severe injury when involved in crashes.

Turning to traffic characteristics, model estimation results indicate that they have fixed parameters across observations as the standard deviation was not found to be significantly different from zero. The negative sign of the traffic volume coefficient implies that for lower traffic volumes, the probability of more severe accidents is significantly higher. The latter comes to verify the common assumption that under free flow, drivers tend to travel at higher speeds and, thus, the severity level of potential accidents increases (Golob et al., 2008; Quddus et al., 2009; The Scottish Office Central Research Unit, 1997). Indeed, the average speed developed under dense traffic conditions (>1,120 vehicles/lane/hour) was found to have a significant and positive association with severity as the corresponding variable ('VQ1') has a positive parameter coefficient (0.015). This implies that beyond a given traffic volume level (1,120/lane/hour), higher speeds imply higher probability for more severe accidents. However, there appears to be no significant difference beneath this traffic volume; this can be attributed to the dispersion of speeds under free flow at a rather random manner and does not affect severity in a consistent pattern.

1.5. Conclusions

We examined the effects of various traffic parameters on crash type and injury severity. Probit models were specified on 4-years of data from the A4-A86 highway section in the Ile-de-France region, France. The empirical results indicated a strong and critical impact of prevailing traffic conditions upon accident occurrences. Traffic speed and volume were found to almost exclusively define crash type and to significantly affect the injury severity level sustained by vehicle occupants involved in accidents. This overall conclusion suggests that similar accident investigations should consider the actual traffic conditions at the moment of the accident occurrences.

The paper contribution to research state of the art and practice is manifold as real-time traffic data are readily available in most freeways and show significant potential for research and applications. However, related work has been limited in the past years. In freeway incident research, most studies use aggregation of exposure data neglecting their natural variance which may result in heavy underdispersion. We used traffic data averaged over a 6-minute interval and collected real-time at the moment of the incident's occurrence. From a methodological standpoint, such disaggregation minimizes possible bias and provides better estimates. Moreover, the exploration of the influence of real-time traffic variables on incident outcomes (in terms of both type and severity) provided significant insight in the incident's mechanism of occurrence.

1.6. Discussion

Based on the analysis of historical data performed, typical traffic patterns recorded prior to accidents may then act as real-time identifiers (Abdel-Aty and Pande, 2007). Such research is useful for researchers and practitioners in estimating accident and congestion external costs and in transportation planning. Further, it may enable practitioners and authorities to locate hazardous spots on the road networks by utilizing real-time data widely available. Once a location is identified as being susceptible to a given crash type occurrence, it may be flagged with warnings through variable message signs (VMS). Furthermore, the concept of variable speed limits could be used to intervene on driver behavior and to reduce speed variation. The presence of traffic police on the designated locations could also serve as a crash prevention measure.

In addition to real-time monitoring of safety levels, a safety performance tool could be developed and used in project evaluation and planning. Safety aspects of costs and benefits can be assessed by comparing the levels of safety before and after implementation of a treatment (Golob and Recker, 2004). Finally, a procedure that uses real-time data on traffic flow, speed, and occupancy and the relationship between these variables and crash-type occurrence could be used to develop congestion mitigation strategies that incorporate safety (Garber and Subramanyan, 2001).

In conclusion, the attempt to further study and develop accident models, and in particular the integration of real-time data, can significantly contribute to the elaboration of a better-structured incident response system with predictive power. Thus, accident counts would be decreased and their consequences would be further limited. Apart from human lives saved, an economic burden would be taken off from societies; non-recurrent congestion would be decreased, while environmental gains would occur.

References

Abdel-Aty, M., & Pande, A. (2007). Crash data analysis: collective vs. individual crash level approach. *Journal of Safety Research*, 38, 581-587.

Abdel-Aty, M., Pande, A., Lee, C., Gayah, V., Cunningham, R., Dhindsa, A., & Dilmore, J. (2007). Linking Crash Patterns to ITS-Related Archived Data: Phase II, Volume I: Real-time Crash Risk Assessment Models. Florida Department of Transportation, BD-550-5.

Anastasopoulos, P.Ch., &Mannering, F.L. (2009). A note on modeling vehicle accident frequencies with random-parameters count models. *Accident Analysis & Prevention*, 41, 153-159.

Aron M., & Seidowsky R. (2004). Diagnostic d'exploitation et de sécurité du tronc commun autoroutier A4-A86 ; Analyse statistique des accidents survenus entre 1996 et 2002. Rapport de convention SISER-INRETS.

Davis, G. (2004). Possible aggregation biases in road safety research and a mechanism approach to accident modeling. *Accident Analysis & Prevention*, 36, 1119-1127.

Garber, N, & Subramanyan, S. (2001). Incorporating crash risk in selecting congestion-mitigation strategies: Hampton Roads Area (Virginia) case study. *Transportation Research Record*, 1746, 1-5.

Golob, T., &Recker, W. (2004). A method for relating type of crash to traffic flow characteristics on urban freeways. *Transportation Research Part A*, 38, 53-80.Greene, W. (2003). Econometric Analysis, 5th Edition, Prentice-Hall Press, New York.

Ivan, J., Wang, C., & Bernardo, N. (2000). Explaining two-lane highway crash rates using land use and hourly exposure. *Accident Analysis & Prevention*, 32, 787-795.

Jacobs, G. D., Aeron-Thomas, A. & Astrop, A. (2000). Estimating global road fatalities. TRL Report 445. Transport Research Laboratory, Crowthorne.

Kim, D-G, Washington, S., & Oh, J. (2006). Modeling crash types: new insights into the effects of covariates on crashes at rural intersections. *Journal of Transportation Engineering*, 132, 282-292.

Kockelman, K., & Ma, J. (2007). Freeway speeds and speed variations proceeding crashes, within and across lanes. *Journal of the Transportation Research Forum*, 46, 43-61.

Lee, C., Hellinga, B., & Saccomanno, F. (2003). Real-time crash prediction model for the application to crash prevention in freeway traffic. In: *Proceedings of the 82nd Annual Meeting of the Transportation Research Board*, Washington, DC.

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. *Transportation Research Part A*, 44, 291-305.

Madanat, S., & Liu, P-C. (1995). A prototype system for real-time incident likelihood prediction. IDEA program final report, Transportation Research Board, US Department of Transportation.

Mouskos, K., Sun, W., & Qu, T. (1999). Impact of Access Driveways on Accident Rates at Multilane Highways. New Jersey Department of Transportation NJ-99-008-NCTIP5, US Department of Transportation.

Oh, C., Oh, J., Ritchie, S., & Chang, M. (2001). Real time estimation of freeway accident likelihood. In: *Proceedings of the 80th Annual Meeting of Transportation Research Board,* Washington, DC.

Pande, A., & Abdel-Aty, M. (2006). Assessment of freeway traffic parameters leading to lane-change related collisions. *Accident Analysis & Prevention*, 38, 936-948.

Pasupathy, R., Ivan, J., & Ossenmbruggen, P. (2000). Single and multi-vehicle crash prediction models for two-lane roadways. USDoT, Region I University Transportation Center, Project UCNR9-8, Final Report.

Qin, X., Ivan, J., & Ravishanker, N. (2004). Selecting exposure measures in crash rate prediction for two-lane highway segments. *Accident Analysis & Prevention*, 36, 183-191.

Washington, S.P., Karlaftis, M.G., & Mannering, F.L. (2003). Statistical end Econometric Methods for Transportation Data Analysis. Chapman & Hall/CRC.

WHO (2004). *World report on road traffic injury prevention*. World Health Organization, Geneva.