ISSN 2303-4521
Methodology/Method

Periodicals of Engineering and Natural Sciences
Vol. 9, No. 4, September 2021, pp.23-30

A cubic regression model to measure the impact of accidents on the number of incidences in the State of Kuwait

Mohammad Zainal[1]
[1]Kuwait University, College of Business Administration, Kuwait

ABSTRACT

The regression analysis theories play a crucial part in safety traffic applications. A cubic regression model was specified in this article to fit the annual traffic incidents in Kuwait during the period 2002-2017 based on annual traffic in Kuwait. Estimation results using the least square estimation and the goodness of fit using the sample autocorrelation function are used to demonstrate the appropriateness of the estimated cubic regression model. As a result, the cubic regression model is supported using the residual analysis through the sample ACF, the sample partial ACF, and the normal probability plot NPP figures of the residuals $W(n)$.

Keywords: Cubic regression model, Traffic accidents, Number of incidences, Maximum likelihood estimation

Corresponding Author:

Mohammad Zainal
Kuwait University, College of Business Administration
P. O. Box 5486, Safat 13055, Kuwait
E-mail: zainal@cba.edu.kw

1. Introduction

Policymakers often have an essential question to ask whether the implementation of speed limits, seatbelt laws, or a sudden increase in population due to a random event(s) impacts the extent and security of traffic-related accidents and the number of incidences.

Traffic accidents form one of the most common modern-day problems that cause deaths and injuries. Today, Traffic Incidences are placed at the top of the list in several countries due to causing high mortality rates and severe injuries. Gulf countries, furthermore, report very high numbers as they suffer from substantial losses, including economic and social impacts and other economic problems related to traffic and the environment in general. Furthermore, most traffic accidents sufferers are young people, which causes a deterioration in a vital part of the population.

Several countries in the Gulf have recognized the importance of traffic safety and the typical association with social and economic growth, public policies, and developed plans to increase the overall traffic safety level. A set of individual budgets have been allocated to overlook projects and plans, including studies in traffic safety, enactment of the advancements in infrastructure, treating harmful accidents, and increasing general traffic safety awareness.

In the past, Gulf states faced an apparent movement towards more democracy, openness, and freedom, resulting in progressing economies and increasing incomes. It subsequently led to an increase in the demand to acquire vehicles of different types, and Kuwait's case was not any different. With increasing numbers of vehicle acquisitions, Kuwait suffered from traffic problems such as traffic congestion, environmental pollution, and traffic accidents which is the primary cause of the increase in the youth mortality rate impacting the Kuwaiti Society and economy negatively.

Consequently, it becomes necessary to study the traffic problem analytically and statistically to find the root of the problem and reduce traffic accidents and state the positive economic and social effects associated with them.

This paper aims to specify a mathematical model to study the annual total number of traffic incidences in relation to the total annual number of traffic accidents in Kuwait. More specifically, a cubic regression model was considered, and estimation was done using the least square estimation (LS) method. The goodness of fit using sample autocorrelation function (ACF) and partial ACF was also used to prove the potentiality and suitability of the proposed cubic regression model in fitting the annual traffic accident data's impact on the total annual number of incidents in Kuwait during the period 2002-2017.

2. Literature review

Traffic accidents in all societies cause one of the most significant social and economic problems. However, different societies have differences in the quality and quantity of the accidents subject to differences in the drivers themselves despite various intersecting factors leading to these accidents. The road type in terms of appearance, extensiveness, trees, rocks, obstacles, and various other factors. Also, the different categories of drivers coming from different educational, demographical characteristics, and psychological characteristics play a major part in traffic accidents. A few previous studies having an indirect or direct relation to the paper's problem will be reviewed.

[5] has pointed out a direct connection between the drivers' cultural and educational level and traffic accidents. He believed that individuals do not feel obliged to obey the public systems and traffic rules in the absence of police officers as drivers may lack driving ethics.

The intersections on the roads which do not have clear signages for drivers are one of the leading causes of traffic accidents [1]. Also, the vehicle's maintenance level affected the accident percentage in the area, necessitating a regular patrol and maintenance of the vehicle to minimize traffic accidents resulting from it.

[12] delivered an inclusive analysis of the modeling techniques and used it to analyze traffic accidents, the severity of the injuries they cause, driver and pedestrian behavior, and the roadway accidents’ operational considerations related to.

According to [4], it is indicated that traffic accidents would be fewer among the higher educational level category due to improved skills and road ethics. Additionally, the negative psychological trends and emotional factors concerning social situations in daily life led to an unstable and turbulent psychological state causing it to be a significant factor in accidents.

The selection of the statistical method is expected to change the impact of an event that has been reported in the literature. Two statistical techniques were used in [13], which addressed risk compensation, the effect of the 1985 seat belt-use law in Illinois. One is related to a before and after method, and the other is related to the analysis of intervention using Autoregressive Integrated Moving Average ARIMA techniques. The developed time series intervention models revealed no statistically significant increase in traffic accidents. Unlike the before-and-after method, the developed time series intervention models revealed no significant change in traffic accidents. Nevertheless, the reported conclusions in that study opposed those reported by [6], [7], [8], [9], who used the traditional before-and-after analysis; in [13], it was claimed that the model misspecification and correlated error terms played against regression techniques.

Other studies have aimed to find the effect of the 55-mph speed limit on traffic accidents in the literature. A domestic speed limit of 55-mph was forced during the 1973 energy crisis in the USA. Meanwhile, numerous studies have assessed the influence of such a speed limit on traffic accidents (TRB, 1984 The impact of a reduced speed limit on the inclination of traffic accident rate has been addressed by [11]. The study specifically dealt with the time intervention when the speed limit was reduced, which significantly impacted traffic accidents.

The study on the traffic in Kuwait determined the leading causes analytically and statistically [2]. He provided probability and statistical models associated with the study's data from the year 2002 up to the year 2013. Also, he aimed to detect the economic and social effects on Kuwait’s society, besides the institutions and government role, to decrease repetitive violations and accidents by using policies, rehabilitation programs, and laws.

A method capable and convenient to fit the traffic accident data which departed from the traditional before-and-after regression techniques and the time series analysis and developed a method capable of fitting the yearly traffic accidents in Kuwait using a convenient lognormal diffusion process was carried out by [3].

[10], specified a rigorous mathematical model to investigate the effect of the Gulf crisis in 1990, imposing numerous families that went back to Jordan on the monthly total number of traffic accidents there. A stochastic diffusion model was specified and estimated in that study. They found that the Gulf crisis did not significantly influence the total number of traffic accidents.

[14], proposed a different method for estimating the frequency of accidents at various severity levels, specifically the two-stage mixed multivariate model, which mixes both accident severity and frequency models. The accident, traffic, and road characteristics data from the M25 motorway and adjacent main roads in England have been collected to show the two-stage model's use. It was noticed that the two-stage mixed multivariate model is an encouraging technique to predict accident frequency concerning their site ranking and severity levels.

3. Statistical analysis

The statistical analysis to calculate the appropriate statistics was founded using SPSS statistical package, such as the following:

1) We use some descriptive statistical measures to calculate the minimum value, maximum value, sample mean and sample standard deviation, in addition to confidence intervals to calculate the 95% confidence intervals for Kuwait's yearly traffic incidence and accidents.

2) We use the regression analysis to fit the effect of the yearly traffic accident data on the annual total number of incidents in Kuwait during the period 2002-2017 and predict the annual number of incidences for the period from 2018 to 2022.

4. Case study

4.1. Statistical measures and confidence intervals

To study the suggested cubic regression model, we obtained the yearly observations of Kuwait's total incidence and accident data from 2002 to 2017 from the Kuwait Traffic Police Department, as shown in Table 1 below.

Table 1. The total yearly number of accidents and incidences

Year	Number of Accidents	Number of Incidents	Year	Number of Accidents	Number of Incidents
2002	37650	2249	2010	65861	786
2003	45376	1332	2011	75194	971
2004	54878	824	2012	86542	9959
2005	56235	863	2013	89527	8977
2006	60410	853	2014	99047	8783

Year	Number of Accidents	Number of Incidents	Year	Number of Accidents	Number of Incidents
2007	63323	1014	2015	80827	9173
2008	56660	1095	2016	71582	10219
2009	61298	670	2017	71582	10305

We summarized the statistical measures and the confidence intervals using the SPSS analysis for the above data in Table 1, as shown in the following Table 2:

Table 2. Some descriptive statistical measures and the confidence intervals for study variables

Variable	Minimum Value	Maximum Value	Mean	Standard Deviation	95% Confidence Interval	
					Lower Limit	Upper Limit
Number of Accidents	37650	99047	67223.19	16333.457	58519.70	75926.67
Number of Incidents	670	10305	4254.56	4283.875	1971.85	6537.28

Table 2 shows that the annual total number of traffic accidents spread is from 37650 to 99047 accidents with a mean and a standard deviation of 67223.19 and 16333.457 accidents. In comparison, the 95% confidence interval for the mean yearly number of traffic accidents varied from 58519.70 to 75926.67 accidents.

The total annual number of traffic incidents was spread from 670 up to 10305 incidents with a mean number of 4254.56 and a standard deviation of 4283.875 incidents. The 95% confidence interval for the mean annual number of traffic incidents was from 1971.85 to 6537.28 incidents.

4.2. Cubic regression model

Assume the dependent variable Y to be the yearly total number of traffic incidents and the independent variable X to be the annual total number of accidents; then the cubic regression model is given by

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \varepsilon \tag{1}$$

Where $\beta_0, \beta_1, \beta_2, \beta_3$ are regression model parameters. And ε is independent sequence and identically normally distributed with zero mean and unit variance σ^2. Now, using the least square estimation method, we get the fitted cubic regression model as follows:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X + \hat{\beta}_2 X^2 + \hat{\beta}_3 X^3 \tag{2}$$

Where the estimated values of the model parameters $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ can be found using SPSS software.

Now, the following results are obtained using the SPSS analysis for the data in Table 1:

Table 3. Model summary results

Model Summary			
R	R Square	Adjusted R Square	Std. Error of the Estimate
.959	.920	.925	2823.788
The independent variable is Accidents.			

Table 3. Model coeficient results

Coefficients					
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Accidents	-3.773	1.814	-14.387	-2.080	.060
Accidents ** 2	5.900E-5	.000	31.238	2.130	.055
Accidents ** 3	-2.806E-10	.000	-16.381	.	.
(Constant)	76159.007	38195.954		1.994	.069

The above results show that the correlation coefficient is 0.959, which indicates that the two variables under investigation are highly correlated. Also, the coefficient of determination is 0.920, which suggests that the cubic regression model is explained by 92% of the variation, resulting in the cubic regression model accurately fitting the data.

We find the estimated parameters using the least square estimation method, as shown in the above output. Thus, the estimated cubic regression model is given by

$$\hat{Y} = 76159.007 - 3.773X + 5.900 \times 10^{-5}X^2 - 2.806 \times 10^{-10}X^3 \tag{3}$$

Figure 1 below demonstrates the actual and predicted annual number of Kuwait incidents from 2002 to 2017, and it shows that the suggested cubic regression model resembles the incident annual traffic data.

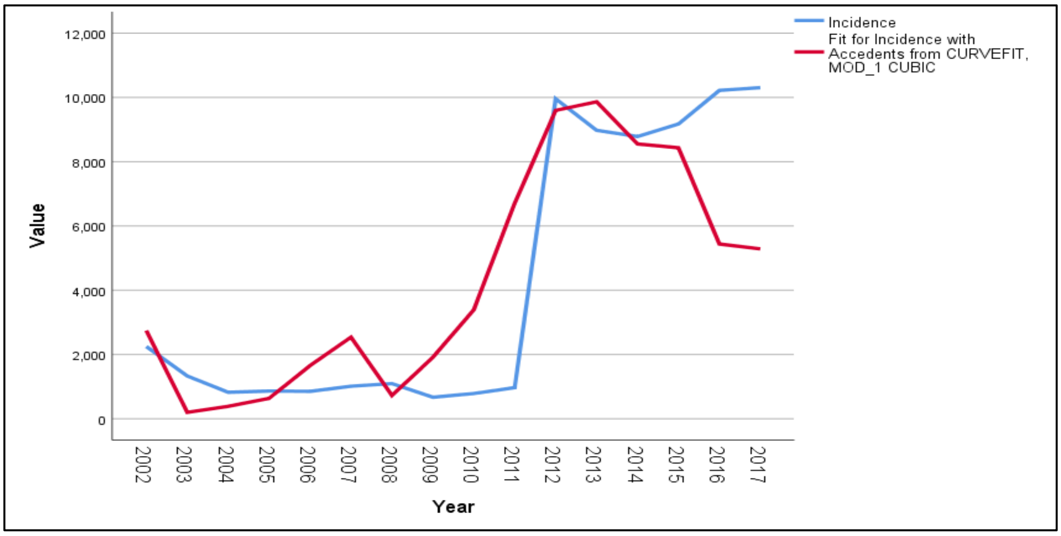

Figure 1. The real and predicted yearly traffic incidents in Kuwait for the period from 2002-2017

5. Residual analysis

In this section, we checked the goodness-of-fit of the suggested cubic regression model using a test based on the sample autocorrelation function ACF and the sample partial ACF as well as the normal probability plot of the rescaled residuals $(W(n))$. Assuming the residual $R(n)$ be as follows (c.f. [3]):

$$R(n) = X(n) - E[X(n)] \tag{4}$$

Let μ and σ be the mean and the standard deviation of the residuals $R(n)$ respectively in this case $\hat{\mu} =$ -0.083and $\hat{\sigma} = 932.46$). Thus, the rescaled residuals are then defined by (c.f. [3]):

$$W(n) = \frac{R(n) - \hat{\mu}}{\hat{\sigma}} \tag{5}$$

Rescaled residuals $(W(n))$ were used to assess the goodness-of-fit of the estimated cubic regression model. Precisely, the residuals should be independently distributed with mean zero and constant variance. The sample autocorrelation function of $(W(n))$ and the values with the bounds $(\pm 1.96/\sqrt{n})$ checks the compatibility with independence. If more than 5% of the values sit out of the limits, then the residual's independence is rejected. Figure 2 shows the sample autocorrelation function of $W(n)$. It is also evident that the model's residuals pass the independence test, which supports the overall goodness-of-fit of the projected cubic regression model.

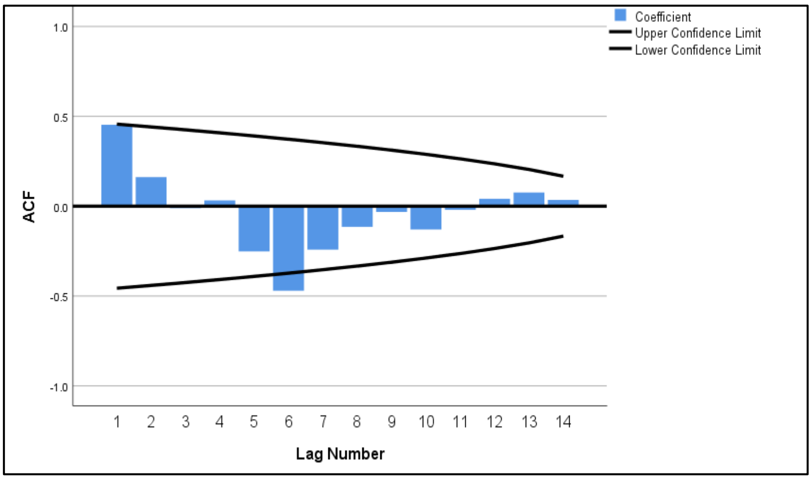

Figure 2. The sample ACF of the scaled residuals

Likewise, the sample partial autocorrelation PACF and the normal probability plot NPP in Figures 3 and 4, respectively of the rescaled residuals $W(n)$ below, help the adequacy of the fitted model.

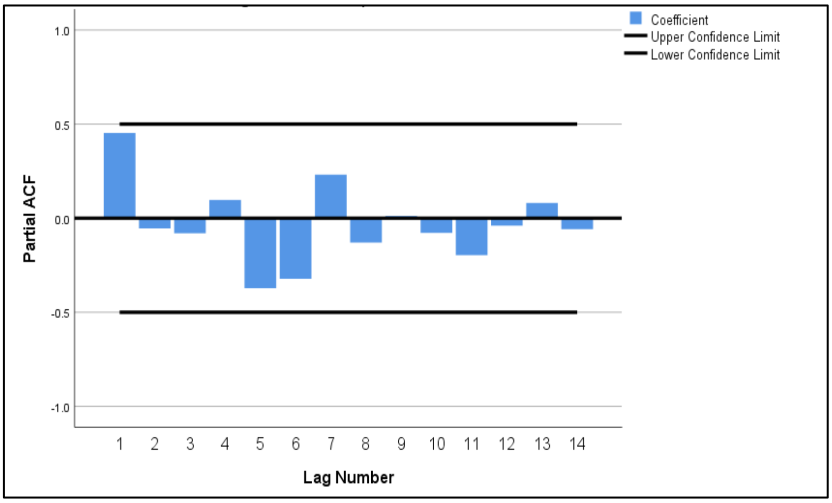

Figure 3. The Sample PACF of the Scaled Residuals

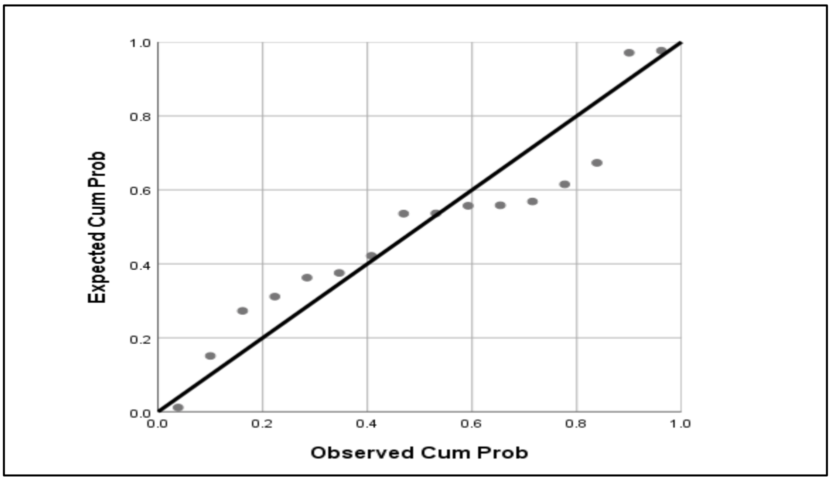

Figure 4. The normal P-P plot of the scaled residuals

6. Prediction of future incidents

For prediction purposes, we first should predict the future annual number of Kuwait traffic accidents. In this case, the average annual rate of increase is computed and found to be equal to 1.085 and then by multiplying it with the number of traffic accidents for the year 2017 to get the number of traffic accidents in the year 2018, and by multiplying it with the number of accidents for the year 2018 to get the number of accidents in the year 2019, until we forecast the number of accidents in the year 2022. Using the fitted cubic regression model obtained in equation (3) above, we get the predicted annual total number of incidences from 2018 to 2022.

Therefore, Table 3 below shows the actual and predicted annual traffic incidents data from 2018 to 2022.

Table 3. The real and the predicted yearly traffic incidents in Kuwait

Year	Real Traffic Accidents	Real Traffic Incidents	Predicted Traffic Incidents
2002	37650	2249	2763.84
2003	45376	1332	219.27
2004	54878	824	413.56
2005	56235	863	663.66
2006	60410	853	1684.19
2007	63323	1014	2571.76
2008	56660	1095	751.02
2009	61298	670	1942.02
2010	65861	786	3425.31
2011	75194	971	6747.05
2012	86542	9959	9644.5
2013	89527	8977	9914.44
2014	99047	8783	8609
2015	80827	9173	8477.07
2016	71582	10219	5475.25
2017	71582	10305	5475.25
2018	76231		7094.94
2019	81299		8600.16
2020	87972		9807.37
2021	95194		9587.7
2022	103076		6809.88

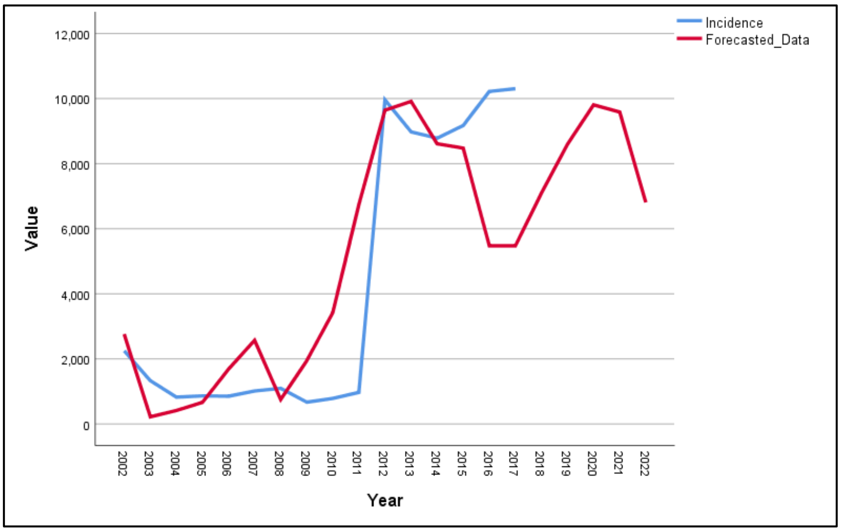

Figure 5. The real and the predicted yearly traffic incidents in Kuwait

Also, Figure 5 shows the predicted annual number of incidents for the same period. Note that the total number of traffic accidents from 2018 to 2022 is predicted recursively using the yearly increase rate, as stated earlier.

7. Concluding remarks

This paper aimed to specify a mathematically sound model to study the annual total number of traffic incidences in relation to the total yearly number of traffic accidents in Kuwait using a cubic regression model, and estimation was done using the least square estimation (LS) method. The goodness of fit using sample autocorrelation function (ACF) and partial ACF was also used to prove the capability and appropriateness of the suggested cubic regression model and fit the yearly traffic accident data's impact on the total annual number of incidents in Kuwait during the period 2002-2017. It was found that it is reasonable to fit the annual total number of traffic incidents data based on the total number of traffic accidents in Kuwait by using a cubic regression model, which is supported by the residual analysis through the figures of the sample ACF and the sample partial ACF as well as the normal probability plot of the residuals $W(n)$. Therefore, this study provides a capable and convenient methodology to fit the effect of traffic accident data on the traffic incident data in Kuwait. In terms of future research, this method could be helpful to study the impact of using traffic control devices on certain transportation services and introducing mandatory laws for traffic in Kuwait.

8. References

[1] J. Abdul Al-Aal, “Traffic accidents and its ruling elements”, a scientific paper presented in "Forty scientific symposium on the methods and means of reducing traffic accidents", *Naif Arab Academy for Security Sciences*, Riyadh: pp. 26-39, 1992.

[2] B. Al-Eideh, “Statistical Analytical Study of Traffic Accidents and Violations in the State of Kuwait and Its Social and Economic Impact on the Kuwaiti Society", *American Journal of Applied Mathematics and Statistics*, vol. 4, no. 2, pp. 24-36, 2016.

[3] B. Al-Eideh, “Modeling the Traffic Accident Data Using a Convenient Lognormal Diffusion Process,” *Asian Research Journal of Mathematics*, vol. 2, no. 3, pp. 1–13, 2017.

[4] A. Al-Saif, “The evolution of the organization of traffic management techniques: *Theoretical and practical aspects*,” Riyadh: p. 25, 1992.

[5] M. Al-Shawan, “The causes of traffic accidents”, *Naif Arab Academy for Security Sciences*, Riyadh: p. 35, 1992.

[6] P. Asch, D. Levy, D. Shea, and H. Bodenhorn, “Risk compensation and the effectiveness of safety belt use laws: a case study of New Jersey,” *Policy Sciences*, vol. 24, pp. 181-197, 1991.

[7] W. Evans and J. Graham, “Risk reduction or risk compensation? The case of mandatory safety-belt use laws”. *Journal of Risk and Uncertainty*, vol. 4, no. 1, pp. 61-73, 1991.

[8] C. Garbacz, "Impact of the New Zealand seat belt laws", *Economic Inquiry*, vol. 29, no. 2, pp. 310-316, 1991.

[9] C. Garbacz, "More evidence on the effectiveness of seat belt laws", *Applied Economics*, vol. 24, no. 3, pp. 313-315, 1992.

[10] M. Hamed, B. Al-Eideh and M. Al-Sharif, "Traffic accidents under the effect of the Gulf crisis", *Safety Science*, vol. 33, no. 1-2, pp. 59-68, 1999.

[11] U. Helfenstein, "When did a reduced speed limit show an effect? Exploratory identification of an intervention time", *Accident Analysis & Prevention*, vol. 22, no. 1, pp. 79-87, 1990.

[12] F. Mannering and C. Bhat, "Analytic methods in accident research: Methodological frontier and future directions", *Analytic Methods in Accident Research*, vol. 1, pp. 1-22, 2014.

[13] S. Rock, "Risk compensation and the Illinois seat belt use law", *Accident Analysis & Prevention*, vol. 25, no. 5, pp. 537-544, 1993.

[14] C. Wang, M. Quddus and S. Ison, "Predicting accident frequency at their severity levels and its application in site ranking using a two-stage mixed multivariate model", *Accident Analysis & Prevention*, vol. 43, no. 6, pp. 1979-1990, 2011.