Accident Analysis and Prevention 83 (2015) 203–213

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Assessing the safety effects of multiple roadside treatments using parametric and nonparametric approaches

Juneyoung Park*, Mohamed Abdel-Aty[1]

Department of Civil, Environmental and Construction Engineering, University of Central Florida, Orlando, Florida 32816-2450, United States

ARTICLE INFO

Article history:
Received 10 March 2015
Received in revised form 3 July 2015
Accepted 6 July 2015
Available online 25 August 2015

Keywords:
Safety effectiveness
Crash modification factors and functions
Multiple treatments
Cross-sectional method
Generalized nonlinear models
Multivariate adaptive regression splines

ABSTRACT

This study evaluates the safety effectiveness of multiple roadside elements on roadway segments by estimating crash modification factors (CMFs) using the cross-sectional method. To consider the nonlinearity in crash predictors, the study develops generalized nonlinear models (GNMs) and multivariate adaptive regression splines (MARS) models. The MARS is one of the promising data mining techniques due to its ability to consider the interaction impact of more than one variables and nonlinearity of predictors simultaneously. The CMFs were developed for four roadside elements (driveway density, poles density, distance to poles, and distance to trees) and combined safety effects of multiple treatments were interpreted by the interaction terms from the MARS models. Five years of crash data from 2008 to 2012 were collected for rural undivided four-lane roadways in Florida for different crash types and severity levels. The results show that the safety effects decrease as density of driveways and roadside poles increase. The estimated CMFs also indicate that increasing distance to roadside poles and trees reduces crashes. The study demonstrates that the GNMs show slightly better model fitness than negative binomial (NB) models. Moreover, the MARS models outperformed NB and GNM models due to its strength to reflect the nonlinearity of crash predictors and interaction impacts among variables under different ranges. Therefore, it can be recommended that the CMFs are estimated using MARS when there are nonlinear relationships between crash rate and roadway characteristics, and interaction impacts among multiple treatments.

1. Introduction

Crash modification factors (CMFs) are multiplicative factors that express the expected changes of crash frequency as a result of a specific treatment (or countermeasure) implemented on roadways. Among four main parts of the Highway Safety Manual (HSM) (AASHTO, 2010), part D provides a variety of CMFs for different roadway facilities such as rural two-lane, rural multilane roadways, and urban arterials. CMFs in part D have been developed using high-quality observational before-after studies that account for the regression to the mean threat. Observational before-after studies are well known methods for evaluating safety effectiveness and calculating CMFs of specific roadway treatments (Gross et al., 2010). Moreover, the cross-sectional method has been commonly applied to derive CMFs due to the ease with which data can be obtained compared to the before-after approaches. According to the HSM, the cross-sectional method is used when (1) the date of the treatment installation is unknown, (2) the data for the period before treatment installation are not available, and (3) the effects of other factors on crash frequency must be controlled for creating a crash modification function (CMFunction) (Abdel-Aty et al., 2014; Lee et al., 2015a).

Although the current HSM provides various CMFs for single treatments, there are no CMFs for multiple treatments to roadway segments and intersections. Due to the lack of sufficient CMFs for multiple treatments, the HSM provides combining method (i.e. multiplication of single treatments) to assess the combined safety effect. However, it is cautioned in the HSM that the combined safety effect of multiple CMFs may be over or under estimated. In particular, since the roadside elements are usually simultaneously applied to roadways and implemented at the same location, interaction effects among multiple roadside features need to be considered to overcome the issue of over- or under- estimation. In general, most previous studies have estimated single treatment effect with no attention for multiple treatments since it is hard to consider the safety effect of single treatment from other multiple treatments implemented at the same time using the observational

* Corresponding author. Tel.: +1 407 823 0300; fax: +1 407 823 3315.
E-mail addresses: jypark@knights.ucf.edu (J. Park), m.aty@ucf.edu (M. Abdel-Aty).
[1] Tel.: +1 407 823 5657; fax: +1 407 823 3315.

http://dx.doi.org/10.1016/j.aap.2015.07.008
0001-4575/

before-after studies (Harkey et al., 2008; Stamatiadis et al., 2011). According to Bonneson et al. (2007),Gross et al. (2009), Li et al. (2011), Park et al. (2014), and Park et al. (2015b), the CMFs need to be developed with consideration of simultaneous impact of more than one roadway characteristic to account for the combined safety effects of multiple treatments.

In order to assess safety effects of multiple roadway characteristics, CMFs have been evaluated using the cross-sectional method (Lord and Bonneson, 2007; Stamatiadis et al., 2009; Li et al., 2011; Carter et al., 2012; Park et al., 2014; Abdel-Aty et al., 2014; Park et al., 2015a; Lee et al., 2015a). To estimate the CMF using the cross-sectional method, development of safety performance functions (SPFs) or crash prediction models (CPMs) is required. Due to its strength of accounting for over-dispersion, generalized linear model (GLM) with negative binomial (NB) distribution has been widely used to develop SPFs. The CMFs can be calculated from the coefficient of the variable associated with specific treatment. However, the estimated CMFs from GLM cannot account for the nonlinear effect of the treatment since the coefficients in the GLM are assumed to be fixed. As one of the efforts to account for the nonlinear effects of crash predictors, many previous researchers have used the logarithm of Annual Average Daily Traffic (AADT) instead of AADT in the analysis (Abdel-Aty and Radwan, 2000; Harwood et al., 2000; Wong et al., 2007; Abdel-Aty and Haleem, 2011; Park et al., 2014; Wang and Abdel-Aty, 2014). Moreover, some previous studies found a nonlinear relationship between crash frequency and roadway characteristics (e.g. lane width and shoulder width) (Xie et al., 2007; Li et al., 2008; Li et al., 2011; Lee et al., 2015a).

Therefore, researchers have tried to apply different techniques to account for the nonlinearity of variables on crash frequency. For instance, an application of using generalized nonlinear model (GNM) was proposed by Lao et al. (2013). In GNMs, the nonlinear effects of independent variables to crash analysis can be captured by the development of nonlinearizing link function. The study found that GNM performs better than GLM since it can reflect nonlinear effects of AADT, shoulder width, grade, and truck percentage on rear-end crashes. Similar to this study, Lee et al. (2015a) estimated CMFs for changes of lane width using GNMs. The study developed nonlinearizing link functions to reflect the nonlinear effects of lane width and speed limit on crash frequency. The CMFs estimated using the GNMs reflect that narrower lanes reduce crashes for the lane width less than 12 ft whereas wider lanes reduce crashes for lane widths greater than 12 ft. It was concluded that the CMFs estimated using GNMs clearly reflect variations in crashes with lane width, which cannot be captured by the CMFs estimated using GLMs. Park et al. (2015b) found that the nonlinear relationship between safety effects of widening urban roadways and time changes. The study developed crash modification functions (CMFunctions) using a Bayesian regression model including the estimated nonlinearizing link function to incorporate the changes in safety effects of the treatment over time. It was found that including the nonlinearizing link functions in developing CMFunctions shows more reliable estimates, if the variation of CMFs with specific parameters has a nonlinear relationship. Moreover, data mining techniques have been applied in the evaluation of safety impacts of roadway features to consider nonlinear effects. Li et al. (2011) utilized the generalized additive model (GAM) to estimate the safety effects of combinations of lane and shoulder width on rural frontage roads in Texas. Similarly, Zhang et al. (2012) applied the GAM to determine the nonlinear relationships between crash frequency and exposure for different segment types. However, most studies investigated only the main effect of each variable, but not the effects of interaction between variables. In addition, the applicability of random parameters modelling approaches has been discussed and tested in order to account for the variations of the effects of variables (or heterogeneity) across observations (Eluru et al., 2008; Anastasopoulos and Mannering, 2009; Venkataraman et al., 2013; Xu and Huang, 2015). However, although the variation of the effects of variables is not fixed and the approach can account for heterogeneity among different sites, interaction impacts between variables were not considered in most studies.

In order to account for both nonlinear effects and interaction impacts between variables, another data mining technique, the multivariate adaptive regression splines (MARS), have been used in safety evaluation studies. According to Briand et al. (2004), the MARS accommodate nonlinearity of independent variables and interaction effects for complex data structure. Unlike other data mining and machine learning techniques, the MARS is a non-black-box model and making it advantageous in the analysis of traffic safety (Haleem et al., 2013). Harb et al. (2010) applied MARS to assess safety effects of toll-lane processing time. Haleem et al. (2010) used MARS to analyze rear-end crashes at un-signalized intersections in Florida. Both studies found that the MARS can be superior to the traditional models and have high model performance. Haleem et al. (2013) also applied MARS to develop CMFs for changes of median width and inside and outside shoulder widths on urban freeway interchange influence areas for total and injury crashes. The study shows that MARS models outperformed the NB models based on their prediction performance and goodness-of-fit statistics. However, the uniform truncated basis functions were used for both total and injury crashes although the rate of changes can vary within the range for different crash types or severity levels. A number of studies addressed the safety effects of roadside features on roadway crashes. The roadside countermeasures have been known as one of the most important treatments for roadway safety to reduce injury crashes (Elvik et al., 2009). The study summarized the aggregate effects of roadside features on injury crash reduction. Other studies have assessed the safety effects of particular roadside elements such as rumble strips, shoulder widths, guardrails, barriers, poles, bridges, signs, ditches and side slopes (Turner, 1984; Good et al., 1987; Gattis et al., 1993; Hadi et al., 1995; Zegeer and Council, 1995; Viner, 1995; Kennedy, 1997; Reid et al., 1997; Bateman et al., 1998; Ray, 1999; Griffith, 1999; Lee and Mannering, 2002; Carrasco et al., 2004; Patel et al., 2007; Jovanis and Gross, 2008; Harkey et al., 2008; Torbic et al., 2009; Wu et al., 2014; Park et al., 2014; Park and Abdel-Aty, 2015 Wu et al., 2014; Park et al., 2014; Park and Abdel-Aty, 2015). As stated by Park et al. (2014), although it is important to examine the interaction impact of multiple treatments implemented on the same location such as roadside, there is a lack of studies that have dealt with this issue.

Thus, the objectives of this study are (1) to analyse the safety effects of multiple roadways and roadside elements using NB, GNM, and MARS, and (2) to develop the CMFs using cross-sectional method for single and multiple treatments for different crash types and severity levels. The remainder of this study is organized as follows. The second section describes data collection and preparation. The third section describes methodologies. The fourth section presents and discusses the results. The final section draws conclusions. In this paper, we refer to ‘All crash types (KABCO severities)’ as Total crashes, ‘All crash types (KABC severities)’ as Injury crashes, ‘All crash types (KAB severities)’ as Severe crashes, and ‘Run-off roadways crashes (KABCO severities)’ as ROR crashes for different crash severity levels. Crash severities were categorized according to the KABCO scale as follows: fatal (K), incapacitating injury (A), non-incapacitating injury (B), possible injury (C) and property damage only (O).

2. Data preparation

In this study, the road geometry data for roadway segments were identified for 5 years (2008–2012) and crash records were collected for 5 years (2008–2012) from multiple sources