Accident Analysis and Prevention 86 (2016) 59–67

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Integration of geometric consistency contributory factors in three-leg junctions collision prediction models of Portuguese two-lane national highways

Jocilene Otilia da Costa[a,*], Maria Alice Prudêncio Jacques[b], Francisco Emanuel Cunha Soares[c], Elisabete Fraga Freitas[d]

[a] Academic Centre of Agreste Region (CAA), University Federal of Pernambuco, Brazil
[b] Postgraduate Program in Transportation (PPGT-UnB), University of Brasilia, Brazil
[c] University of Minho, Portugal
[d] Territory, Environment and Research Centre (C-TAC), University of Minho, Portugal

ARTICLE INFO

Article history:
Received 31 May 2015
Received in revised form 18 August 2015
Accepted 3 October 2015

Keywords:
Three-leg junctions
Portuguese national roads
Collision prediction models
Geometric consistency
Generalized estimating equations
Random-effects models
Random-parameters models

ABSTRACT

This paper aims at developing a collision prediction model for three-leg junctions located in national roads (NR) in Northern Portugal. The focus is to identify factors that contribute for collision type crashes in those locations, mainly factors related to road geometric consistency, since literature is scarce on those, and to research the impact of three modeling methods: generalized estimating equations, random-effects negative binomial models and random-parameters negative binomial models, on the factors of those models. The database used included data published between 2008 and 2010 of 177 three-leg junctions. It was split in three groups of contributing factors which were tested sequentially for each of the adopted models: at first only traffic, then, traffic and the geometric characteristics of the junctions within their area of influence; and, lastly, factors which show the difference between the geometric characteristics of the segments boarding the junctions' area of influence and the segment included in that area were added. The choice of the best modeling technique was supported by the result of a cross validation made to ascertain the best model for the three sets of researched contributing factors. The models fitted with random-parameters negative binomial models had the best performance in the process. In the best models obtained for every modeling technique, the characteristics of the road environment, including proxy measures for the geometric consistency, along with traffic volume, contribute significantly to the number of collisions. Both the variables concerning junctions and the various national highway segments in their area of influence, as well as variations from those characteristics concerning roadway segments which border the already mentioned area of influence have proven their relevance and, therefore, there is a rightful need to incorporate the effect of geometric consistency in the three-leg junctions safety studies.

1. Introduction

The World Health Organization – WHO has revealed that over 1.24 million people die and 50 million are injured every year on the road all over the world (WHO, 2013). In Portugal, the National Highway Safety Authority (Autoridade Nacional de Segurança Rodoviária – ANSR) informed that 518 people were killed and 38,872 were injured as a result of road severity in the year 2013. This means that about 2 people were killed and approximately 106 people were injured daily due to road crashes in Portugal. ANSR also revealed information about the distribution of the fatalities and serious injuries within the national highway network – about 35% and 30% respectively occurred on national roads (NR's) (ANSR, 2013).

Considering the huge costs that road crashes entail for society, knowing the factors that affect the probability of occurrence of a crash has been the subject of research for many decades. Most researchers who have approached this issue (Lord and Persaud, 2000; Wang and Abdel-Aty, 2006; Anastasopoulos and Mannering, 2009, 2011; Lord and Mahlawat, 2009; Cafiso et al., 2010; Gomes, 2013; and others) have focused on the understanding of the factors that affect the rate of crashes occurring in certain

* Corresponding author.
E-mail addresses: jocilene.mt@gmail.com (J.O. da Costa), mapjacques@gmail.com (M.A.P. Jacques), a61864@alumni.uminho.pt (F.E.C. Soares), efreitas@civil.uminho.pt (E.F. Freitas).

http://dx.doi.org/10.1016/j.aap.2015.10.003
0001-4575/

road elements during a certain period of time (week, month, year or set of years).

The development of the prediction models has distinct purposes, such as: to research the connections between the crashes and the main characteristics of the road environment (Fitzpatrick et al., 2008); and to elaborate a quantitative prediction of road crashes (Vogt and Bared, 1998). It requires special attention in order to make sure that the results and their interpretations are accurate, namely regarding: (i) the specification of functional relations; (ii) the choice of explanatory variables and the type of model; (iii) the interpretation of the found relations causes; (iv) the fit assessment; (v) the performance assessment of the estimation model; and (vi) the assessment of potential causes of errors.

The data to be modeled on road crashes will always be made up of nonnegative integers, to which regression models for counting data are proper, or others that can encompass the nature of this kind of information. Lord and Mannering (2010) present an analysis of different methodological alternatives which can be used in the development of these models, while also point out the potentialities and limitations of each approach according with the characteristics of the available database. Among the approaches, Kim and Washington (2006), for instance, have developed models based on generalized linear models' modeling techniques – GLM with a negative binominal distribution, for junctions within rural roads, while Wang and Abdel-Aty (2006) used the generalized estimating equations' modeling technique – GEE, also with a negative binominal distribution for error, in the modulation for the frequency of rear-end collisions in signalized junctions located in inter-urban roads, while taking into consideration the time and space correlation between data. As for Kumara and Chin (2003), they worked with zero-inflated models in order to explain why these crashes happened in signalized junctions. In these past few years, other modeling techniques have been used. Such is the case of Anastasopoulos et al. (2012) who used the random-parameters tobit model to study the crashes rate in roadway segments; Castro et al. (2012) have developed a model to predict the crash frequency at junctions using a special case of generalized ordered-response models; Ferreira and Couto (2013) have modeled the traffic flow-accidents relationship for urban junctions by using the translog function; and Bhat et al. (2014) have proposed a model designated by CEMPS (count model with endogenous multinomial probit selection) models to model road crash frequency at junctions.

In most of the reviewed studies, the authors used variables from two sets of factors: (1) crash risk exposure factors (usually: traffic volume and distance covered) and (2) factors associated with the road and road environment that are directly connected with the road element in question (geometry, type of control, etc.).

As for the introduction of variables which allow to assess the impact of the geometric consistency of the design where crashes occur (which is included in the set of factors associated with the road), there are very few papers in literature. Mayora and Rubio (2003) and Cafiso et al. (2010) papers, for instance, which focus on road segments, introduce the issue of geometric consistency through an estimate of the operational speed of vehicles (V85) in consecutive road segments. In literature review, no papers dealing with the issue of geometric consistency for the modeling of road crashes frequency in junctions were found. It should be noted that the geometric consistency concerns the geometry conditions of the design, specifically, its horizontal and vertical alignment which allow drivers to drive their vehicle without being surprised by a sudden change in geometry that will make them substantially reduce the speed of the vehicle. According to Polus and Mattar-Habib (2004), a similarity between the project characteristics contributes for road safety. Levelled junctions are operationally complex places within the transportation system in the crash rate study due to the level of conflicts, as Mitra and Washington (2012) point out. Within the context of geometric consistency, these junctions in themselves cause, by their nature, a severe change in the design. Nevertheless, if the fact that the junction's area of influence encompasses varied road segments with certain geometric characteristics is considered, it is important to make some research in order to ascertain if the geometric consistency of those segments, with regard to neighboring segments, is an aggravating factor in the occurrence of crashes.

The models obtained in crash studies available in literature can be used as generalized tool, however, it is necessary to adapt them properly to the road context and to the level of approach they are meant to be used in. From a practical point of view, it is important to guarantee that the models are useful for the purposes that justify their development. Different studies point out that using the developed models directly on cases that differ from where the used data was collected leads to unsatisfactory results (Sawalha and Sayed, 2006; Gomes, 2010; Sacchi et al., 2012; Cunto et al., 2013). These studies have found several inconsistencies which show that applying these models in a different context from which they were obtained, is not recommended.

Within this context, several crash studies were conducted in Portugal in the past few years. These studies have dealt with different road elements, segments or junctions (Ferreira, 2010; Gomes, 2010, 2013; Couto and Ferreira, 2011; Gomes and Cardoso, 2012; Gomes et al., 2012; Costa et al., 2015). However, none of these papers deals with collisions which occur in different types of junctions in Portuguese NR's. This type of highways represents about 23% of the entire Portuguese road network system with the exception of streets in urban centers. In 2012, over 59% of crashes registered in roads outside the country's urban centers (ANSR, 2013) took place in NR's. Collisions represented around 51% of those incidents. NR's located in the Northern Region of Portugal, which connect several cities and industrial areas in the country, registered about 53% of all collisions occurred. As to the type of junctions, there are several three-leg junctions in the NR's of that region. For instance, in the case of the stretches of NR's selected for this study, 74% of existing junctions are of a three-leg type.

Therefore, the goal of this paper is to develop a collision prediction model for three-leg junctions located in national roads (NR) in Northern Portugal. The focus is to identify the factors that contribute for collision type crashes in those locations, so that management bodies responsible for road safety in Portugal can prevent crashes by creating new corrective measures based on the identified factors. Specifically, the aim of this research is: (i) to assess the impact of geometric consistency on the occurrence of collisions in the junctions targeted in this study and (ii) to investigate the impact the adoption of different modeling techniques would have in the contributive factors identified using available data and/or in the importance of those factors (coefficients of the regression model). The impact of the geometric consistency is determined through the inclusion in the list of variables researched as contributing factors, of variables which show the difference between the geometry of the road segment in which the junction is located and the geometry of its neighboring segments. For the study of the impact of different modeling techniques, the authors tried to compare the results obtained with three methods that can be applied to panel data (Generalized Estimating Equations – GEE, Random-effects negative binomial models – RENB, and Random-parameters negative binomial models – RPNB).

The starting point of this study was as an example of what was already ascertained for road segments, the hypothesis that the occurrence of collisions in junctions is affected by deep changes between the physical properties of the segment included in the junction's area of influence and corresponding changes observed in the segments connected to that area.