Available online at www.sciencedirect.com
SCIENCE DIRECT
Safety Science 43 (2005) 541–557
SAFETY SCIENCE
www.elsevier.com/locate/ssci

Analysis of freeway accident frequencies: Negative binomial regression versus artificial neural network

Li-Yen Chang *

Graduate Institute of Transportation and Logistics, National Chia-Yi University, 300 University Road, Chia-Yi 600, Taiwan

Received 27 October 2004; received in revised form 13 April 2005; accepted 17 April 2005

Abstract

The Poisson or negative binomial regression model has been employed to analyze vehicle accident frequency for many years. However, these models have the pre-defined underlying relationship between dependent and independent variables. If this assumption is violated, the model could lead to erroneous estimation of accident likelihood. In contrast, the artificial neural network (ANN), which does not require any pre-defined underlying relationship between dependent and independent variables, has been shown to be a powerful tool in dealing with prediction and classification problems. Thus, this study employs a negative binomial regression model and an ANN model to analyze 1997–1998 accident data for the National Freeway 1 in Taiwan. By comparing the prediction performance between the negative binomial regression model and ANN model, this study demonstrates that ANN is a consistent alternative method for analyzing freeway accident frequency.

Keywords: Accident frequency; Artificial neural network; Negative binomial regression; Freeway

1. Introduction

The impact that traffic accidents have on society is significant. For example, there are approximately 3000 people are killed and thousands more injured by traffic accidents in

* Tel.: +886 5 271 7982.
E-mail address: liyen@mail.ncyu.edu.tw

0925-7535/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ssci.2005.04.004

Taiwan each year. Individuals injured (or killed) in traffic accidents must deal with pain and suffering, medical costs, wage loss, higher insurance premium rates, and vehicle repair costs. For society as a whole, traffic accidents result in enormous costs in terms of lost productivity and property damage. Therefore, public agencies have put much effort into preventive measures, such as illumination and enforcement. However, the annual number of traffic accidents has not yet significantly decreased. Therefore, there should be further research studies on the risk factors for traffic accidents. This study focuses on the non-behavioral factors of freeway accident risk, specifically highway geometric characteristics and environmental conditions. Although past statistics indicated that most traffic accidents resulted from drivers' errors (behavioral factors), non-behavioral factors also play an important role in traffic safety. Not only can they contribute to certain types of driver errors (e.g., speeding often occurs at downgrades), but accidents will be likely to occur at the same location repeatedly if the problem is not mitigated. In addition, with a better understanding of non-behavioral factors of freeway accidents, transportation engineers will be able to design freeways to higher safety standards.

Past research analyzing accident frequencies mainly relied on statistical models such as linear regression models, Poisson regression or/and negative binomial regression models because the occurrence of accidents on a highway section can be regarded as a random event. Another major advantage of applying these statistical models is the ability to identify a broad range of risk factors that can contribute significantly to accidents. However, most of the statistical models have their own model assumptions and pre-defined underlying relationships between dependent and independent variables. If these assumptions are violated, the model could lead to erroneous estimation of accident likelihood. Artificial neural networks (ANN) which do not require any pre-defined underlying relationship between dependent and independent variables have been widely employed in financial analysis, decision problems, and pattern recognition. The ANN has been shown to be a powerful tool, particularly in dealing with prediction and classification problems. There has also been an increased interest in applying ANN in the field of transportation since the 1990s, such as driver behavior analysis, pavement maintenance, vehicle detections, and so on (Dougherty, 1995). However, the applications of ANN to analyze traffic safety problems have been relatively few. Therefore, this study examines whether ANN can be used to analyze the relationship between risk factors and accidents. This is done by evaluating the prediction performance between the negative binomial regression model and ANN model. The paper begins with a review of previous literature on modeling accident frequencies and then presents the methodological approach. A description of the available data and an assessment of the model estimation results follow this. The paper concludes with a summary and directions for future research.

2. Literature review

Past research on modeling accident frequencies has been diverse, both empirically and methodologically. From an empirical standpoint, most research studies (Shankar et al., 1995; Milton and Mannering, 1998; Carson and Mannering, 2001) have focused on non-behavioral risk factors of accidents on the freeway or arterial roadways. These non-behavioral factors included highway geometry (e.g., horizontal and vertical alignments, and shoulder width), traffic characteristics (e.g., average annual daily traffic (AADT) and percentage of trucks) and weather conditions (e.g., rain or snow). The