Analytic Methods in Accident Research 8 (2015) 33–44

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

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

Multilevel models to analyze before and after speed data

Md Tazul Islam [a,*], Karim El-Basyouny [b,1]

[a] Department of Civil and Environmental Engineering, University of Alberta, 4-110 NREF, Edmonton, AB, Canada
[b] City of Edmonton Assistant Professor of Urban Traffic Safety, Department of Civil and Environmental Engineering, University of Alberta, 3-004 NREF, Edmonton, AB, Canada

ARTICLE INFO

Article history:
Received 7 July 2015
Received in revised form
29 October 2015
Accepted 29 October 2015
Available online 19 November 2015

Keywords:
Multilevel model
Heterogeneous variance
Speed modeling
Posted speed limit reduction

ABSTRACT

Analyzing before–after speed data is often limited to a standard comparison of various speed parameters. Although a few studies have used a model-based approach, various limitations exist in terms of both data and methodology. The aim of this paper was to examine the applicability of using multilevel models to analyze before–after speed data and to explore the effect of various temporal, geometrical, and traffic characteristics on traffic speed in an urban residential context. Two multilevel models, one with homogeneous and one with heterogeneous within-site variance, were used for analyzing the hourly free-flow speed data. The study used a dataset collected before and after a posted speed limit (PSL) reduction from 50 km/h to 40 km/h; the reduction was a pilot program in the city of Edmonton, Alberta, Canada. The results demonstrated the appropriateness of using the multilevel model for analyzing speed data. Moreover, the heterogeneous within-site variance model outperformed the homogeneous counterpart in terms of goodness-of-fit and the precision of parameter estimates. The parameter estimations demonstrated intuitive findings with respect to the effect of various factors on mean free-flow speed. In general, the evaluation results showed that the mean free-flow speed was reduced by 4.6 km/h in the after period, when the multilevel model with heterogeneous within-site variances was used. The use of the multilevel model with homogeneous within-site variances slightly underestimated the mean free-flow speed reduction. A separate investigation revealed that the mean free-flow speed was reduced by 5.3 km/h and 4.0 km/h for local and collector roads, respectively.

1. Introduction

It is well established that the traffic speed is related to the severity of crashes (Aarts and van Schagen, 2006; Elvik, 2013; Hauer, 2009; Nilsson, 2004); consequently, speed management is considered to be the key strategy to reduce traffic fatalities and injuries (OECD, 2006). For example, the vision zero or safe system approach adopted by different countries around the world has identified effective speed management as the cornerstone to achieve the vision zero goal. A few examples of speed management initiatives include, but are not limited to, public education or training, intensifying speed enforcement and penalties, improving road infrastructure, lowering the speed limit, and adopting new technologies, such as intelligent speed adaptation (OECD, 2006).

* Corresponding author. Tel.: +1 780 695 5877; fax: +1 780 492 0249.
E-mail addresses: mdtazul@ualberta.ca (M.T. Islam), basyouny@ualberta.ca (K. El-Basyouny).
[1] Tel.: +1 780 492 9564; fax: +1 780 492 0249.

http://dx.doi.org/10.1016/j.amar.2015.10.001
2213-6657/

Analyzing before–after speed data to evaluate the effectiveness of the adopted speed management strategy in reducing traffic speed is an important component of the overall strategy. While many earlier studies used a standard before–after approach (Hoareau et al. 2006; Kamya-Lukoda, 2010; Kloeden et al., 2004), few recent studies explored the value of using a model-based approach in before–after speed data analysis (Heydari et al., 2014; Islam and Basyouny, 2015; Kuo and Lord, 2013). The model-based approaches are considered more reliable, as they can account for the effect of various confounding factors and time trend.

In general, conventional ordinary least square (OLS) regression is the most commonly used method for modeling speed data (TRC, 2011). This single-level regression modeling method assumes that each observation of speed is independent. In reality, the speed data are often multilevel (at least two-level) in nature, as the data are collected for multiple sites with multiple observations from each site. The data collected from different sites can exhibit different speed characteristics because of the dissimilarity in site characteristics, such as geometric design, surrounding environment, etc. Similarly, within-site speed data can show variability because of the difference in driver characteristics, traffic flow, vehicle type, temporal pattern, etc. Therefore, the heterogeneous variances in speed data can be divided into two categories: between-site and within-site (Poe and Mason, 2000). The conventional OLS regression method cannot address these two variances, and hence can result in biases in speed prediction (Park et al., 2010).

Few studies have proposed alternative methods to overcome the flawed assumption of data independence considered in the OLS method. These alternative methods include the mixed model (Eluru et al., 2013; Poe and Mason, 2000; Wang, 2006), panel data analysis (Figueroa-Medina and Tarko, 2004; Tarris et al., 1996), the simultaneous equation model (Himes and Donnell, 2010; Porter, 2007; Shankar and Mannering, 1998), and multilevel model (Cruzado and Donnell, 2010; Park et al. 2010). All of these methods included within-site and between-site variances to address the multilevel nature of speed data.

The methodologies mentioned above provided significant improvement over the conventional ordinary least square (OLS) regression method in speed modeling. Nevertheless, several avenues for improvement in speed modeling warrant further investigation. The single level model for analyzing speed data ignores the fact that observations within a site are correlated. This limitation can lead to a loss of information and consequently results in inefficient and inconsistent econometric estimation (Mannering and Bhat, 2014). The multilevel models, which can be defined as an error term correlation issue in traditional econometrics, can be used to address the within-site and between-site correlation of the speed data. However, the existing applications of a multilevel model in analyzing speed data often constrained the model coefficients to be constant, although, according to Gelman and Hill (2007), one of the key components of a multilevel model is varying coefficients (i.e., varying intercept, varying slope, or both). The literature suggests that restricting the coefficients to be constant/fixed when they actually vary across observations/sites can lead to inconsistent and biased coefficient estimates (Washington et al., 2003). A varying coefficient can also account for the unobserved heterogeneity that is likely to be present in the absence of an exhaustive list of explanatory variables (Anastasopoulos and Mannering, 2009; Mannering and Bhat, 2014). Similar to the concept of a varying coefficient, it is possible that the within-site variances in speed data can vary across sites due to the presence of unobserved heterogeneity. Restricting the within-site variances to be constant/fixed across sites can also lead to biased coefficient estimates and consequent speed prediction. Existing speed modeling studies have hardly investigated the effect of varying within-site variance on model coefficient estimates and speed prediction.

2. Objective of the study

The objective of this study was to explore the application of a multilevel model to analyze before–after speed data. Within the multilevel modeling framework, the model intercept was allowed to vary across speed survey sites. Moreover, the traditional homogeneous/fixed within-site variance multilevel model was extended by considering heterogeneous/varying within-site variance. The models were compared to investigate whether the use of heterogeneous variance contributed to improving the model's goodness-of-fit and the precision of coefficient estimates. Speed data collected before and after an urban residential posted speed limit (PSL) reduction pilot program was used. Therefore, the current study quantified the effect of urban residential PSL reduction on vehicle speed. Furthermore, use of a number of explanatory variables in the models allowed for investigating the effect of various temporal, traffic, and geometrical factors on speed behavior.

3. Literature review

3.1. Speed modeling

Numerous studies have modeled speed data (such as mean free-flow speed, 85th percentile speed) as a function of various geometrical, temporal, traffic flow, and weather characteristics. The motivation was to better understand how these characteristics influence the speed of traffic, which in turn can help in making informed policy decisions towards improving traffic safety. Results from these studies reveal that various geometrical features (such as lane and shoulder width, degree of curve, etc.), traffic flow characteristics (such as daily traffic), and traffic laws (such as the PSL) significantly influence speed behavior (TRC, 2011).