REVISITING THE MODIFIABLE AREAL UNIT PROBLEM IN
MACRO-LEVEL CRASH MODELING

Qing Cai
University of Central Florida
4000 Central Florida Blvd, Orlando, Florida 32816
Phone: +01 4074064132 E-mail: qingcai@knights.ucf.edu

Mohamed Abdel-Aty, University of Central Florida; Jaeyoung Lee, University of Central Florida

1. Introduction

Macro-level safety analysis has been an important component in transportation safety planning. Although various geographic units have been employed for previous macro-level crash studies, no studies provide a specific guidance to select the best geographic units. In this study, a comparative analysis was conducted to determine the optimal geographic unit for macroscopic crash modeling including census tracts (CTs), traffic analysis zones (TAZs), and traffic analysis districts (TADs). According to the U.S. Census Bureau, the CTs are small, relatively permanent subdivisions of a county or equivalent entity to present statistical data such as poverty rates, income levels, etc. The TAZs are geographic entities delineated by state or local transportation officials to tabulate traffic-related data such as journey-to-work and place-of-work statistics (Abdel-Aty et al., 2013). The TADs are new, higher-level geographic entities for traffic analysis (FHWA, 2011). The TADs are developed by aggregating current TAZs, block groups or CTs. In Florida, the average area of CTs, TAZs, and TADs are 155, 65, and 103 square miles, respectively. Across the three geographic units, a TAD is considerably larger than a CT and TAZ while a TAZ is the smallest in most of cases.

2. Method

Six models including the Poisson lognormal (PLN) models without and Poisson lognormal conditional autoregressive (PLN-CAR) model with consideration of spatial autocorrelation for total, severe, and non-motorized mode crashes were developed based on CTs, TAZs, and TADs. Based on the estimated models, predicted crash counts for the three geographic units were computed. To compare the model performance between different geographic units, this study proposed a new method by adopting grid structures of different scales. The grid structure is developed with uniform length and shape across the whole state without any artifact impacts. Also, the numbers of grids remain the same for all candidate geographic units thereby providing a common comparison platform. Considering the average area of each geographic unit, ten grid structures with dimensions ranging from 1 to 10 miles were created for the comparison analysis. Each grid intersected with several geographic units (i.e., CTs, TAZs, and TADs) and can be divided into several regions accordingly. The observed crash counts for each grid were directly obtained with GIS while the different predicted crash counts were transformed into the grids that each geographic unit intersects with. Specifically, the transformed predicted crash counts in a grid were calculated by summing up the predicted crash counts of all intersected regions. The mean absolute error (MAE) and root mean squared error (RMSE) were calculated for the observed and different transformed predicted crash counts of different grid structures. By comparing the MAE and RMSE values, the best geographic unit as well as model for macroscopic crash modeling can be identified with the same sample size.
3. Results

The comparison results indicated that the models based on TADs offered the best fit for all crash types. Also, the comparison results highlighted that models with the consideration of spatial effects performed consistently better than those not considering spatial effects. The modeling results based on different geographic units had different significant variables, which demonstrated the zonal variation. Besides, the results highlighted the importance of several explanatory variables such as traffic (i.e., VMT and heavy vehicle mileage), roadway (e.g., proportion of local roads in length, signalized intersection density, length of sidewalks), and socio-demographic characteristics (e.g., population density, commuters by public transportation, walking, and cycling, median household income).

4. Conclusions

The results indicated that the proposed comparison method could work properly and reasonably in comparing different geographic units for macro-level crash analysis. The modeling results demonstrated the importance of considering spatial autocorrelation for the crash modeling at macro-level. Based on the comparison results and motivation for developing the different geographic units, it is recommended adopting TADs for transportation safety analysis and planning. Also, according to the modeling results, the hotspot could be identified and appropriate countermeasures could be suggested at the planning level.

REFERENCES
