Macro Safety Analysis for Non-motorized Vehicles Based on Roadway and Safety Education Improvement

Zhicheng Dai, Xuesong Wang, Wenbin Luo Tongji University, China

PAPER NO:21-02415

ABSTRACT

Due to differences between crash locations and the locations of the crash-involved road users' residences, traditional macro safety analysis strategy could be inefficient when comprehensive improvements such as traffic safety education programs are considered. This paper:

- Proposing a new analysis strategy of separately aggregating crashes for roadway engineering improvement and road users for education improvement.
- Collecting non-motorized crashes, crash-involved road users, roadway, socioeconomic and land use characteristics from 213 sub-districts in Shanghai.
- Developing a multivariate Poisson lognormal conditional autoregressive (CAR) model as prediction model due to correlation between the location of the crash and the residence of the crash-involved road users.
- Applying potential for safety improvement (PSI) to identify hot zones for engineering and education improvement respectively.
- Developing False Positive (FP) and False Negative (FN) indexes to identify the differences of the hot-zone distributions quantitatively.

The purpose of this study is to find the most appropriate strategy to apply effective roadway and safety education countermeasures for safety improvement.

ANALYSIS STRATEGY

TRADITIONAL ANALYSIS STRATEGY:

- Using crash location to identify hot zones.

However, the analysis strategy may be inefficient when other countermeasures such as traffic education improvement are considered, due to the difference between the location of crashes and the location of the residence of the involved road users.

PROPOSED ANALYSIS STRATEGY:

- Aggregating crashes and involved road users for engineering and education improvement, respectively.

A higher number of crashes represents a urgent need for engineering improvement, and a higher number of crash-involved road users in a given area usually represents a more urgent need for education program implementation.

- All non-motorized crash-involved road users were included in the analysis strategy

As traffic education programs provide information not only on traffic regulations, but also on self-protection skills, all non-motorized road users can benefit.

DATA PREPARATION

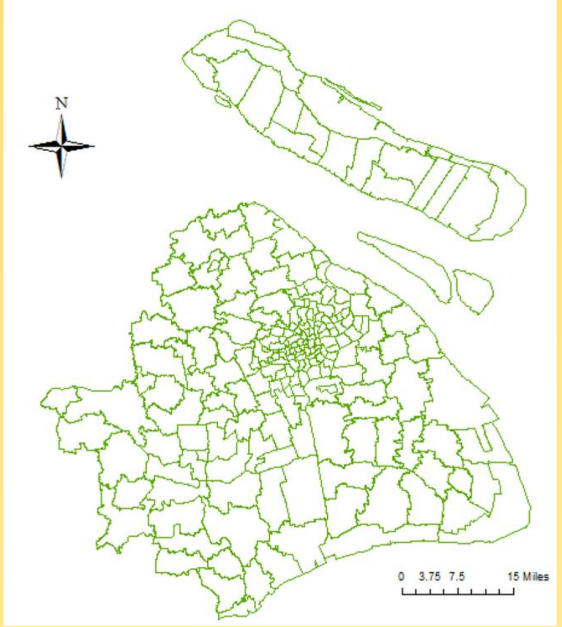
In a coordinated effort to assist local administrators ready to apply new data to traffic safety improvement, detailed data from the 213 sub-district administrative units in Shanghai were collected for this study.

INDEPENDENT VARIABLES:

- Roadway network data, including information on road type (e.g., highway, street) and intersections, was obtained from the Shanghai traffic police.
- socioeconomic influence was represented by registered population.
- Land use characteristics were represented by area, area type and number of metro stations.

DEPENDENT VARIABLES:

- The location and road user residence for each crash were precisely located on the map by longitude and latitude, and the numbers of each were gathered and counted in each sub-district using ArcGIS software.
- These two dependent variables were divided into two severity levels, fatal and injury crashes (FI) and property damage only crashes (PDO), for analysis of correlation between severity levels.



STATISTICAL MODELING

Macro-Level Safety Model:

- A Bayesian multivariate Poisson-lognormal conditional autoregressive model was developed. The framework for Bayesian inference was as follows:

$$\pi(\theta y) = \frac{L(y\theta)\pi(\theta)}{\int L(y\theta)\pi(\theta)d\theta}$$

- The Poisson-lognormal model assumes that the dependent variable y_{ij} follows the Poisson distribution as follows:

$$y_{ik} \sim Poisson(\lambda_{ik})$$

- The logarithm was used as a function to link the expected value of y_{ij} with independent variables as follows:

$$\log(\lambda_{ik}) = \beta_{0k} + X_{ik}\beta_k + \varepsilon_{ik} + \varphi_i$$

ε_{ik} denote the unobserved heterogeneity for sub-district i and subject variable k, and follow the multivariate normal distribution.

- The proximity matrix W with entry w_{ij} indicates the spatial relationship between unit i and unit j:

$$w_{ij} = \begin{bmatrix} \frac{1}{d_{ij}} & i \neq j \\ 0 & i = j \end{bmatrix}$$

- The conditional distribution of φ_{ij} and φ_{-ij} turned out to be a normal distribution (φ_{-ij} is the set of $\varphi_{i'j}$):

$$\varphi_{ij} \mid \varphi_{-ij} \sim N\left(\sum_{i' \neq i} \frac{w_{i'j,ij}}{w_{j+}} \varphi_{i'j}, \frac{\sigma_\varphi^2}{w_{j+}}\right)$$

The estimation process used the Bayesian estimation approach implementing the Markov Chain Monte Carlo (MCMC) algorithm in the open source software WinBUGS® to conduct estimation. In the process of model calibration, two MCMC chains of 40,000 iterations were obtained, with the first 5,000 iterations discarded as burn-in.

Subject	*FI Crash*		*PDO Crash*		*FI Road User*		*PDO Road User*	
Variable	Mean	90% BCI	Mean	90% BCI	Mean	90% BCI	Mean	90% BCI
Intercept	2.4810	(2.2070, 2.7610)	4.2230	(4.0340, 4.4310)	2.5360	(2.3600, 2.7150)	4.3220	(4.1850, 4.4760)
Highway Length (Levels 1-2)	0.0718	(0.0345, 0.1082)	0.0604	(0.0355, 0.0841)	0.0300	(0.0058, 0.0547)	0.0297	(0.0097, 0.0492)
Street Length (Levels 1-2)	-0.0736	(-0.0982, -0.0493)	-0.0042*	(-0.0239, 0.0149)	-0.0323	(-0.0462, -0.0188)	-0.0057*	(-0.0182, 0.0060)
Street Length (Levels 3-4)	0.0197	(0.0080, 0.0309)	0.0037*	(-0.0038, 0.0127)	0.0089	(0.0021, 0.0154)	9.585E-04*	(-0.0040, 0.0064)
Intersection Density	-0.0090	(-0.0164, -0.0016)	-0.0086	(-0.0134, -0.0038)	-0.0107	(-0.0150, -0.0063)	-0.0069	(-0.0099, -0.0037)
Number of 4-legged Intersections	0.0090	(0.0058, 0.0124)	0.0143	(0.0116, 0.0168)	0.0053	(0.0031, 0.0077)	0.0104	(0.0084, 0.0124)
Registered Population	0.0034	(0.0003, 0.0066)	0.0060	(0.0037, 0.0085)	0.0086	(0.0065, 0.0107)	0.0100	(0.0083, 0.0116)
Area	-0.0101	(-0.0144, -0.0055)	-0.0067	(-0.0095, -0.0036)	-0.0040	(-0.0071, -0.0011)	-0.0060	(-0.0081, -0.0035)
Number of Metro Stations	0.0100*	(-0.0579, 0.0836)	0.0771	(0.0357, 0.1239)	-0.0321*	(-0.0753, 0.0116)	0.0273	(0.0003, 0.0569)
Area Type	0.6813	(0.4445, 0.9213)	-0.5550	(-0.7351, -0.3696)	0.0046	(-0.1467, 0.1583)	-0.3327	(-0.4528, -0.2094)
CAR Effect	0.4770	(0.1789, 0.6105)	-	-	-	-	-	-
DIC Value	6617.85	-	-	-	-	-	-	-

Mean values noted by * are nonsignificant at the 0.1 level; "-" means there is no corresponding value for this parameter..

Analysis of Results:

- Roadway characteristics had similar effects on both crashes and crash-involved road users, but the land use characteristics had different influence towards the studied subjects.
- The street length of Levels 1-2 was negatively associated with the studied subjects, while the length of Levels 3-4 streets had an opposite safety affect, probably because a higher level of street tends to have a better riding environment.
- Urban area type was found to be positively associated with FI crashes and involved road users but negatively associated with PDO crashes and involved road users., probably because a highly mixed traffic composition in urban areas that can pose a higher risk for serious crashes.

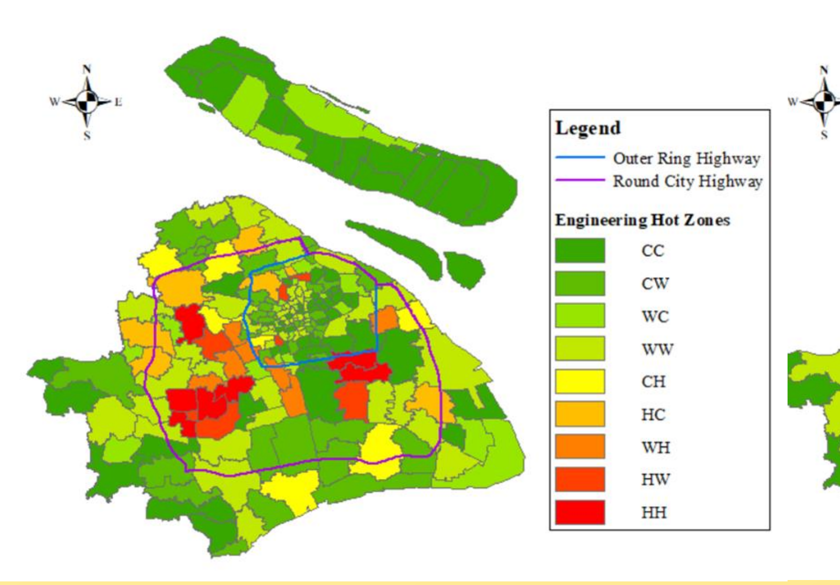
HOTZONE IDENTIFICATION

POTENTIAL SAFETY IMPROVEMENT:

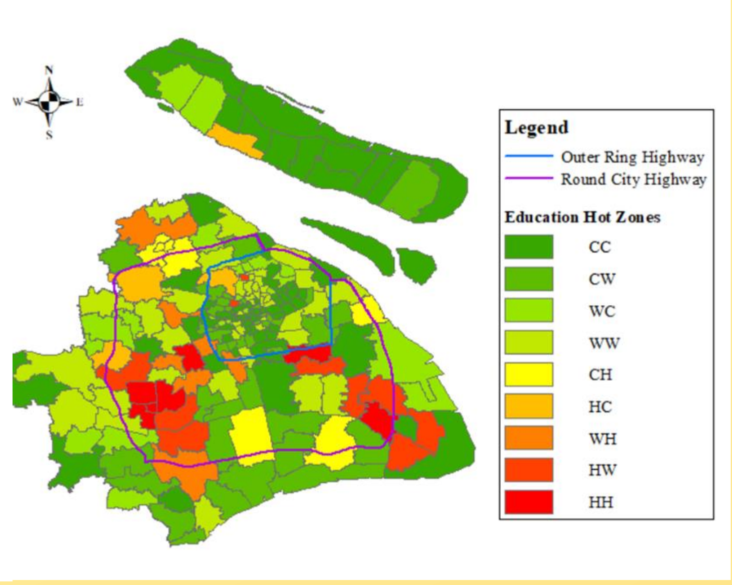
Potential Safety Improvement (PSI) refers to the difference between the expected and predicted number of crashes (or crash-involved road users). PSIs for the four subjects (FI crashes, PDO crashes, FI road users, and PDO road users) were calculated using the multivariate model.

- Hot zones (H) were identified as sub-districts with top 10% of PSI values.
- Warm zones (W) were identified as having PSIs greater than zero, but not in the top 10%.
- Cold zones (C) were the sub-districts with PSIs lower than zero.

Engineering-potential hot zones and education-potential hot zones were identified and integrated according to crash and road user categories, respectively, of severity. The first letter denotes the FI category, and the second letter denotes the PDO category; for example, HW is used to represent sub-districts classified as hot for FI crashes and warm for PDO crashes.



Engineering Hot Zones



Education Hot Zones

COMPARISON OF HOT-ZONE IDENTIFICATION RESULTS:

- False Positive (FP) index refers to the ratio of sub-districts that were identified as engineering hot zones but education cold or warm zones.
- False Negative (FN) index refers to the ratio of sub-districts identified as education hot zones but engineering cold or warm zones.

Index	FI	PDO	Integrated Results
False Positive	10/21	10/21	16/35
False Negative	9/21	9/21	16/35

These results indicated that improvement countermeasures would have been implemented incorrectly in nearly half of the identified hot zones using the traditional analysis strategy.

SUMMARY AND CONCLUSIONS

This study proposed a new strategy for macro safety analysis based on comprehensive improvement countermeasures, which was accomplished in this study by separately aggregating crashes for roadway engineering improvement, and crash-involved road users for traffic safety education improvement.

- A multivariate Poisson lognormal CAR model was developed to examine the relationships between regional characteristics and traffic safety.
- Integrated identification results showed significant differences between engineering hot zones and education hot zones.
- A false positive index and a false negative index were further developed to more accurately identify the differences, which indicated that nearly half the identified hot zones were inconsistent in unnecessarily prioritizing either engineering or education improvement.

Due to the separate analyses of crashes and crash-involved road users, the new analysis strategy proposed in this study is thus far more practical for safety improvement. Conclusions drawn from this study suggest a different approach to non-motorized vehicle traffic safety for local administrators and transportation planners and provide a comprehensive understanding of hot zones for a more effective use of limited resources.