STATISTICAL MODEL ANALYSES OF TRAFFIC ACCIDENTS INVOLVING ELEMENTARY STUDENTS FOCUSING ON EXPOSURE VARIABLE

Toyohashi University of Technology, Student Member, Halimatusaadiah Sahdan
Toyohashi University of Technology, JSCE Member, Kojiro Matsuo
Toyohashi University of Technology, JSCE Member, Nao Sugiki

1. Introduction

Active travel to school such as walking and cycling, which contributes to the physical activity and a healthy environment, highly relies upon pedestrian safety. In Japan, most elementary students walk to school via designated school routes. Although the rate of elementary students' accidents (ESA) has declined steadily, the urgency to upgrade modification of built environment as well as to raise awareness among road users to curtail child traffic accidents still remains. To do so, it is important to appropriately evaluate the ESA risk of each location. Although the exposure variables (i.e. the quantity of road users who pass through and exposed to the accident risk of each location) have been thought as a key factor of total accident risk, they have been not easy to be measured so far, particularly on smaller roads or residential roads due to the great number of locations.

Through statistical model analyses, this study tries examining surrogate variables for vehicle exposure and elementary student exposure based on car probe data and digitalized school commuting route data, respectively. This study also aims to reveal the potential ESA risk of each intersection.

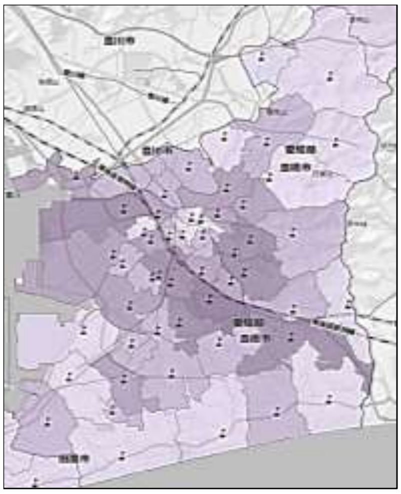

Fig.1 ESA frequency by school area. Darker area indicates higher frequency. (Max:27)

2. Methodology

1) Study area and target accidents

The study area was Toyohashi city, Aichi (**Fig.1**). The target was accidents around intersection on community streets (excluding arterial roads) involving elementary students. There are 52 elementary schools in Toyohashi, and total number of enrolled students (6-12 years old) were around 60700 in 2016. The dominant mode of travel to school was walking (99.95%).

2) Accident data

Police-reported injury accident data from 2008 to 2015 (excluding 2012 due to locating error) was used. Total ESA were 481 (2.15% of all accidents) during the study period. ESA mostly occurred after school time (**Fig.2**), and most of the students were involved in the accidents when they had travelled with no school commuting purpose (accidents with school commuting purpose was only 7.4%). This study defined an intersection area as a 15 m radius area from the center of the intersection (**Fig.3**). The number of ESA occurred inside the intersection area was aggregated as the unit of analysis.

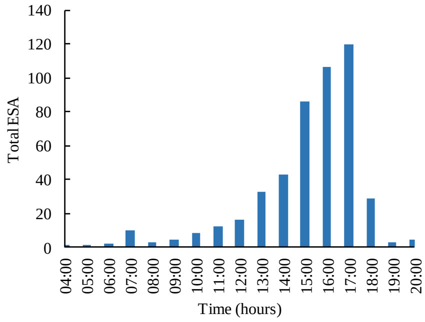

Fig.2 Total ESA by time in hours

3) Vehicle exposure

General vehicle probe data (collected by Pioneer Corp.) for a year in 2013 was used to calculate a surrogate measures for vehicle exposure (traffic volume) of each intersection. The probe data was at first map-matched to digital link data (provided by ZENRIN Corp.), then the number of the probe vehicles passing through each intersection during a year was aggregated.

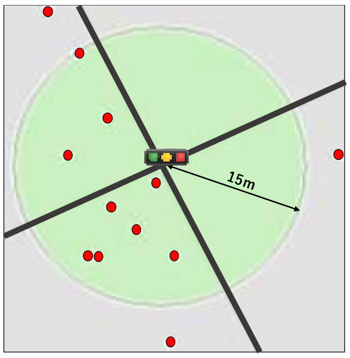

Fig.3 ESA in a 15-m radius area from intersection's center

4) Elementary student exposure

The board of education of Toyohashi have digitalized the designated school routes and the gathering points (students once gather at the neighboring points and then start commuting along the school routes) for all elementary schools with the information of the number of students using each school route. Based on this data, three types of surrogate measures for elementary student exposure (student's activity) around each intersection were calculated: i) the number of students who use school routes within 15 m, 50 m and 100 m from each intersection; ii) the number of students who gather at gathering points within 15 m, 50 m and 100 m from each intersection; iii) distance from each intersection to nearest school route. **Fig.4** shows the ESA frequency at each intersection overlaid with school routes, indicating that many ESA occurred on the school routes. Although the ESA occurred with no school commuting purpose as mentioned above, our hypothesis is that elementary students' activities are concentrated around the school commuting routes even without

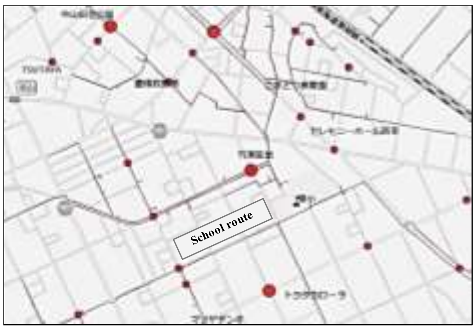

Fig.4 ESA frequency at intersection. Bigger circle indicates higher frequency (Max:3).

school commuting purpose and thus the school routes are spatially related to the ESA risk at intersections.

5) Statistical analysis

In this study, two count data regression models, which are Negative Binomial (NB) model and Zero-inflation Negative Binomial (ZINB) model were applied. The better model was then selected based on smallest value of Akaike Information Criteria (AIC).

a) Negative Binomial (NB) regression model:

$$P_{NB}(Y_i = y|\lambda_i, \phi) = \binom{y+\phi+1}{\phi-1}\left(\frac{\phi}{\phi+\lambda_i}\right)^{\phi}\left(\frac{\lambda_i}{\phi+\lambda_i}\right)^{y},$$

$$\lambda_i = \exp(\boldsymbol{x}_{\boldsymbol{i}}^{t}\beta) = exp\left(\beta_0 + \textstyle\sum_j \beta_j \cdot x_{ij}\right),$$

$$E_{NB}(Y_i|\lambda_i, \phi) = \lambda_i, \quad V_{NB}(Y_i|\lambda_i, \phi) = \lambda_i + \frac{\lambda_i^2}{\phi}$$

where, Y_i is the number of vehicle-pedestrian accidents at an intersection i during the study period; $\boldsymbol{x}_{\boldsymbol{i}}^{t}$ is the predictor variables vector for the intersection i; β is the coefficient parameters vector; and ϕ is the dispersion parameter.

b) Zero-inflation Negative Binomial (ZINB) regression model:

$$P_{ZINB}\left(Y_i = y|p_{i,}\lambda_{i,}\phi\right) = \begin{cases} p_i + (1-p_i)\cdot P_{NB}(Y_i = 0|\lambda_i) & for\ \ y = 0 \\ (1-p_i)\cdot P_{NB}(Y_i = y|\lambda_i) & for\ \ y > 0' \end{cases}$$

$$p_i = P(S_i = Zero\,risk\ \ state \mid \boldsymbol{z_i},\ \boldsymbol{\alpha}) = \frac{1}{1+exp(-\boldsymbol{z_i\alpha})}$$

$$\mathrm{E_{ZINB}}(Y_i|p_i, \lambda_i, \phi) = p_i \cdot 0 + (1-p_i)\cdot\lambda_i$$

where, $\boldsymbol{z}_{\boldsymbol{i}}^{t}$ is the zero-inflation related predictor variables vector for the intersection i; $\boldsymbol{\alpha}$ is the coefficient parameters vector.

3. Results and discussions

Table 1 presents the both final model estimated results. The AIC value is smaller in the ZINB model, which indicates the better model in this study. Hence, results for the ZINB model was discussed in-depth. The exposure variables of 1) probe traffic volume, 2) distance from intersection to nearest school route as well as the road or area environment variables of 3) the number of leg and 4) land use were found to be significant factors in the ZINB model and that associated with ESA at intersection.

Table 1 Results of the NB and ZINB models

Explanatory variable		Negative Binomial		Zero-inflation Negative Binomial			
				Zero-inflation part		Positive count part	
		Coefficient	P value	Coefficient	P value	Coefficient	P value
Logarithm of traffic volume		0.330	<<0.01	-0.589	<<0.01		
Distance from intersection to nearest school route		-0.00792	<<0.01			-0.00775	<<0.01
Number of leg		0.869	<<0.01			0.867	<<0.01
Land use	High-rise residential and commercial land	1.07	0.0339			1.11	0.0235
	Low-density low-rise residential land	0.665	<<0.01			0.669	<<0.01
	Public land	1.03	0.0520			0.977	0.0724
Dispersion parameter		0.645		2.681			
Sample size		15060		15060			
McFadden's ρ^2		0.19					
AIC		2326.042		2323.754			

The greater the logarithm of probe traffic volume, the smaller the probability to be zero risk is. The mixture of high-speed motorized traffic with vulnerable road users influence the accident risk. Higher traffic volumes affect road safety as the number of interactions between road user increases. Thus, students that are assumed to have high risk to accidents are escalating with the increment in traffic volume.

The distance from each intersection to the nearest school route decreases the probability of high frequency of ESA. The expectation would decrease by $\exp(-0.00792 * 10) - 1.0 = 0.07$ (or 7%) for every additional distance of 10 m. This is considering elementary students' activity that mostly happened near school route assuming that they already familiar to the route. On the other hand, the quantities of students assigned to each school route were not significant. The number of students near school route and gathering point for school commuting purposes did not affect directly as other significant factors because mainly most traffic accidents were recorded with no school commuting purpose.

Intersections with a high number of leg have a substantially higher risk since higher level of awareness is required to walk across an intersection with higher number of leg to avoid accident.

Type of land use that were included in this study were agricultural, forest, industrial, residential, commercial, public, roads, and seaside. Only 1) high-rise residential and commercial land, 2) low-density low-rise residential land and 3) public land predict higher ESA risk. Regarding residential land, road traffic is evolving to meet the needs of various occupants' activities. Hence, the consequent risk to road traffic injuries may be high towards elementary students. As for public land which includes airport, stadium, gymnasium and other areas at which students most presumably present, resulting in more complex interactions between elementary students and vehicles, similar to commercial land that can comprise of shopping centers and offices buildings.

4. Conclusion

This study found that in analyzing ESA at intersection, probe data-based traffic volume and the distance from intersection to nearest school route can be reasonable surrogate variables for vehicle exposure and elementary student exposure, respectively. While some road and area environment variables that affect to ESA risk were revealed, it is needed to consider more detail road environment variables in further study.

REFERENCES

1. Gain Lee, Yuna Park, Jeongseob Kim, Gi-Hyoug Cho, 2016. Association between intersection characteristics and perceived crash risk among school-aged children. Accident Analysis and Prevention. 97,111-121