ANALYZE THE SPATIAL DISTRIBUTION OF DELIVERY MOTORCYCLE CRASHES AND IDENTIFY THE RELATED FACTORS

I Gede Brawiswa Putra[1*], Pei-Fen Kuo[1], Chui-Sheng Chiu[1], Umroh Dian Sulistyah[1]

[1] Department of Geomatics, National Cheng Kung University, Taiwan

KEY WORDS: Delivery motorcycle crash, Geographically weighted negative binomial regression, Point of interest.

ABSTRACT:

During COVID-19, the suspension of the dine-in option at restaurants had significantly increased online food delivery crashes in Taiwan. Nevertheless, the majority of current studies remain focused on the common motorcycle, which has distinct driving habits and routes than a delivery motorcycle. Even though some recent studies identified the variables contributing to delivery motorcycle crashes, they still restricted in defining crash severity model and did not account for spatial dependences. In this study, two different models were used in this study: the generalized linear model (GLM), and the geographically weighted negative binomial model (GWNBR) to estimate crash frequency in a non-stationary pattern. In 2020, there were 2314 delivery motorcycle crashes in Taipei, according to the study area. Besides that, the point of interests data from 456 villages in Taipei city was considered as related crash factors for further analysis. According to the results, GWNBR showed the best performance in terms of log-likelihood, Akaike Information Criterion (AIC), and Root Mean Square Error (RMSE). Furthermore, this research reveals that commercial areas and bus stations had a significant impact on delivery motorcycle crashes. As per the coefficient distribution, the effect is exacerbated in rural areas where the traffic policy is still a major concern. As the popularity of delivery food services grows, this topic will become even more important in the future.

1. INTRODUCTION

Online food delivery services, such as Uber Eats and Food Panda, offer a service where users can order food from various restaurants that do not provide their own delivery services. The order is picked up from the restaurant and delivered to nearby customers in collaboration with courier partners (such as Uber drivers). Furthermore, the online food delivery industry has grown rapidly in recent years, which has been aided by the COVID-19 quarantine and the suspension of dining options in restaurants. With the rise of food delivery activities, the number of motorcycle crash events has recently increased. According to a Taipei Police Department report in 2020, from October 2018 to November 2020, there were 2,325 delivery motorcycle crashes, or 5.4 crashes per day. As a result, food delivery platforms were forced to accept some of the blame and devise strategies to improve driver safety.

According to the Taipei Police Department, the most common crash-related factors are distracted driving, speeding, and failure to yield. Such reckless behavior is common since many food delivery services prioritize speed over safety. Several employees have complained that food delivery platforms will reassign an order to a different driver if it is not delivered within a stipulated period of time. This philosophy, in fact, encourages speeding behavior rather than safe driving. In order to minimize driver fatigue, New Taipei City has narrowed the number of working hours per day for delivery services to 12 hours. However, if a food delivery platform is registered in multiple cities and customer orders in City A, the order will be transferred to nearby City B to be fulfilled. As a result, even if the platform's drivers cause more accidents, the platform is not required to submit a safety plan to City A. Managers may also be unaware of a driver's total work hours if they register with multiple platforms. Despite claims that Food Panda and Uber Eats have implemented a policy to hire safer drivers who are 20 or older, have one to two years of driving experience, and have no serious traffic violations, these platforms are unable to manage contract workers rather than formal employees.

On the other hand, environmental factors such as land use may be associated with the frequency of delivery motorcycle crashes. Land use can have an impact on the socioeconomic and demographic characteristics of a given area, which can influence the likelihood of a crash. It also has the potential to alter traffic patterns and volume, thereby influencing the occurrence and number of traffic accidents (Pulugurtha et al., 2013; Xu et al., 2020). Previous studies that land-use data to investigate its impact on traffic crashes yielded varying results. For example, Levine et al. (1995) explored motorcycle crash patterns in Honolulu, Hawaii, and found that the majority of crashes occurred nearby business districts rather than residential zones. Similarly, Ukkusuri et al. (2012) also identified that regions with a larger ratio of industrial and commercial areas seem to be more prone to have vehicle accidents. Meanwhile, areas with a larger ratio of residential zones had lower accident risks, particularly for pedestrians. In contrast to previous studies' findings, Kim et al. (2010) found a significant relationship between vehicle accidents and commercial land use, residential land use, and urban construction areas with mixed residential and commercial land use (Pulugurtha et al., 2013; Xie et al., 2019). In addition, they also indicated that higher population density could increase the likelihood of crashes. In summary, while they discovered a positive association between vehicle accidents and residential land use, there was no direct relationship exists between vehicle accidents and population density. Those incidents arise because this type of land use attracted more pedestrians and increased the volume of traffic, with population density and activity being the main factor.

[*] Corresponding author

https://doi.org/10.5194/isprs-archives-XLIII-B4-2022-163-2022

Many studies have extensively investigated the macroscopic and microscopic analysis of vehicle accidents in various spatial entities (Abdel-Aty et al., 2013; Lee et al., 2015; Osama et al., 2016; Cai et al., 2017; Pljakić et al., 2019; Wang et al., 2019). In the field of traffic safety, the macroscopic spatial analysis is used to identify whether the frequency of vehicle crashes within a given space of space is correlated with the characteristics (such as demographic or land-use) of the space (Levine et al., 1995). This macroscopic spatial analysis is more concerned with the aggregate effects of accidents in units of space rather than individual accidents (Lee, 2014). On the other hand, Wang et al. (2016) argue that micro-level analysis is more effective at pinpointing specific places where road safety issues occur, which is preferred for detailed traffic safety analyses. However, in practical uses, macro-level safety analysis allows the identification of problems in a greater geographical area, which assists in generating a long-term road safety plan. Hence there are more studies that include macro-level analysis as well as spatial units such as countries (Huang et al., 2010), regions (Huang and Abdel-Aty, 2010), traffic analysis districts (Abdel-Aty et al., 2013; Cai et al., 2017), census blocks (Wang and Kockelman, 2013) and traffic analysis zones (Wang et al., 2012; Pulugurtha et al., 2013; Lee et al., 2018). When it comes to macro-level analysis, the primary criterion is to include various important characteristics of spatial data in the modeling (Lee et al., 2014). So far, the aforementioned studies have relied on the Geographical Information System (GIS) to collect and visualize spatial data that represent particular characteristics of the spatial unit (Wang et al., 2012). As a result, we chose the village level as our spatial unit in this study because it can represent land use and demographic characteristics. These characteristics can vary from place to place, creating an opportunity to conduct a detailed local analysis.

The majority of macro-level research on crash influence mechanisms has focused mainly on global regression (Jia et al., 2018). In that type of regression, both the predictor and response variables are assumed to be spatially stationary. However, disregarding the spatial characteristics of traffic accidents can result in biased results. As a consequence, several local modeling methods for capturing spatial heterogeneity have been recommended. One of the most widely used local regression methods is geographically weighted regression (GWR) (Chen et al., 2020). It has been discovered that geographically weighted Poisson regression (GWPR) and geographically weighted negative binomial regression (GWNBR) provide more spatial randomness than the generalized linear model (GLM) because they allow coefficients to vary spatially across observations (Hadayeghi et al., 2010). Among the GWR models, GWPR and GWNBR are the best suitors in crash analysis and prediction (Mathew et al., 2022). The GWNBR model could account for overdispersion, which is commonly observed in crash data. Failure to properly address overdispersion will result in the underestimation of standard errors and misleading inference for the coefficients (Chen et al., 2020). Although the negative binomial model (NB) has been used to address overdispersion in the crash analysis as an alternative to the Poisson regression model, it still fails to account for spatial heterogeneity.

Therefore, there is an urgent need to investigate the potential key crash-related factors of these online delivery services. If there is a higher crash risk for online food delivery service workers, these companies must be held accountable because they receive a large percentage of the workers' pay (up to 30 percent of the order). Surprisingly, very few studies on the safety issues associated with delivery motorcycle crashes have been conducted. Chung et al. (2014), for example, used data from Korea to identify possible related factors to motorcycle crashes. Furthermore, their dataset did not include a comparison of delivery motorcycles to other modes of transportation during the same time period, let alone platforms. Since the Taipei City DOT has begun to separate delivery motorcycle crashes from non-delivery motorcycle crashes, we will use this valuable information to see if there are any differences between delivery motorcycle-related crashes and others

2. METHODOLOGY

2.1 Study Area

This study used Taipei city as the study area. Taipei city has an estimated population and area of 2,668,572 and 271.7997 km^2, respectively. Figure 1 illustrates the population density and the land use of Taipei city. According to Figure 1a, the majority of people live in the central area of Taipei city which is represented by the yellow, orange, and red areas. These areas are made up of various land uses, such as a residential area, a school, and a commercial area, as shown in Figure 1b by the yellow, purple, and red areas, respectively. In the western area of Taipei, the population density is lower than in the city's central area (light green and yellow area in Figure 1a) since that area is mostly the commercial area and shopping center of Taipei city (red area in Figure 1b). Meanwhile, the eastern, southern, and north-eastern areas of Taipei City are classified as sub-urban areas because of their lower population density (mostly have the light green area in Figure 1a). However, those areas have different types of land use, with the southern and northern areas of Taipei dominated by residential areas (yellow area in Figure 1b), but still have a few schools and commercial areas (purple and red area respectively in Figure 1b). Meanwhile, the western area is dominated by an industrial area (brown color in Figure 1b) and contains a few residential areas. Finally, the northern part of Taipei City is a national park, with almost no other land use type other than conversational open spaces.

The city of Taipei was chosen for this study because it is Taiwan's capital city with the busiest transportation activity and the highest traffic crashes. According to the statistics of the Ministry of Transportation and Communications, the total number of delivery motorcycle crashes in Taipei in 2020 was 2,314. The data available consists of point data from each motorcycle crash that occurred during the one-year period from January 2020 to December 2020. In terms of explanatory variables, this study collected land use data related to points of interest in Taipei City. Previous studies (Jia et al., 2018; Chen et al., 2019) used that variable, which may have a significant impact on delivery motorcycle crashes. This study included POI data from restaurants, supermarkets, shopping malls, schools, hotels, and bus stops (please table 1 for more detail information).

https://doi.org/10.5194/isprs-archives-XLIII-B4-2022-163-2022 |

Variables	Mean	Min	Med	Max	Std dev
Delivery motorcycle crashes Weekday	3.662	0	3	26	3.680
Delivery motorcycle crashes Weekend	1.412	0	1	13	1.706
Restaurant	18.206	0	14	212	18.090
Supermarket	0.511	0	0	5	0.725
Shopping mall	0.066	0	0	2	0.273
School	0.605	0	0	6	0.812
Hotel	0.044	0	0	5	0.321
Bus Stop	0.375	0	0	12	0.977

Table 1. Descriptive Statistics.

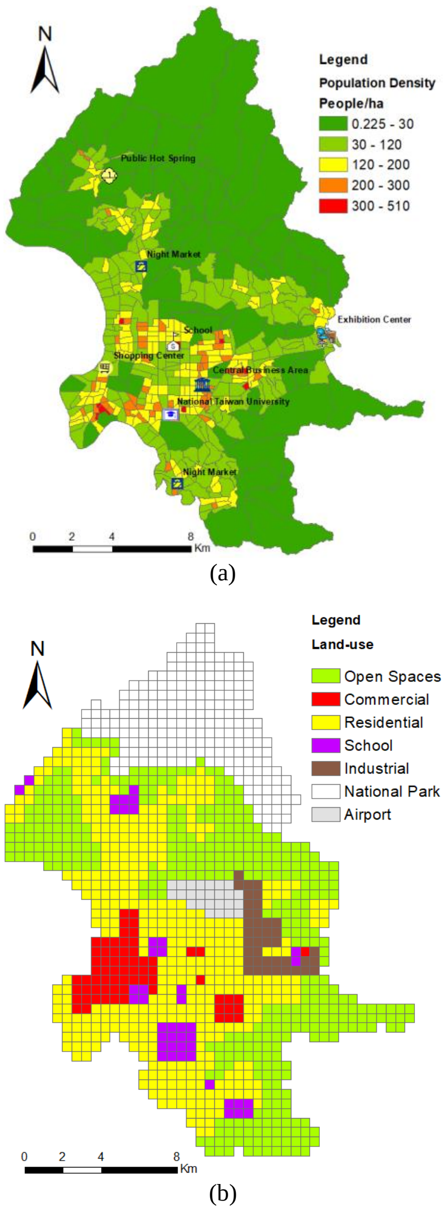


Figure 1. Illustration of the study area; (a) population density distribution; and (b) land use of Taipei City.

2.2 Generalize Linear Model (GLM)

One of the basic assumptions of traditional linear regression models is that the response variable needs to have a normal distribution; however, this hypothesis is rarely fit in crash frequency data. For example, if the frequency of motorcycle accidents was used as the response variable, the distribution would no longer be a normal distribution but rather a Poisson or negative binomial distribution. Then, alternatively, generalized linear models such as the Poisson and negative binomial regression models were used in this study. When the response variable is over-dispersed counts data, the negative binomial regression model is commonly used. Nonetheless, the negative binomial distribution assumes that the data has a high variance, which is common in crash data.

$$y_i \sim NB\left[t_i \exp\left(\sum_k \beta_k\, x_{ik}\right), \alpha\right] \tag{1}$$

where NB stands for negative binominal distribution, y_i is the number of delivery motorcycle crashes in the i^{th} ($i = 1,\ldots,n$) spatial unit, x_{ik} is the kth explanatory variable for spatial unit i, β_k ($k = 0,1,\ldots,p$) are the coefficients of each variable, t_i is the number of spatial unit i, which is the offset variable, and α is the parameter of overdispersion.

2.3 GWNBR

The Geographically Weighted Negative Binomial Model (GWNBR) was commonly used to investigate the relationship between crashes frequency and its related factors. According to Xu and Huang (2015), GWNBR was used to model overdispersed count data, which was limited by the available software Statistical Analysis System (SAS). This software was developed by SAS Institute to model spatial heterogeneity and overdispersion simultaneously. The GWNBR model can be described as follows:

where t_i is an offset variable, which is the number of spatial units, β_k is the coefficient for the independent variable x_k, for $k = 1,\ldots n$, y_i is the number of delivery motorcycle crashes in the i_{th} spatial unit, and α is the parameter of overdispersion.

$$y_i \sim NB\left[t_i \exp\left(\sum_k \beta_k\,(u_i, v_i) x_{ik}\right), \alpha\,(u_i, v_i)\right] \tag{2}$$

The basic concept of GWNBR is deduced from the first law of geography, which states that observations nearby to the location I have a greater influence on the estimation of βk(ui,vi) than the observations with the further location. The magnitude of the influence can be effectively represented by a kernel function, and bi-square is one of the most commonly used kernel functions, which is used in this study:

$$w_{ij} = \left(1 - \left(\frac{d_{ij}}{b_{i(k)}}\right)\right)^2 \; if \; d_{ij} < b_{i(k)} \tag{3}$$

where d_{ij} is the Euclidian distance between spatial unit i and spatial unit j, and $b_{i(k)}$ is the bandwidth value.

Bandwidth has a significant impact on parameter estimation. The corrected Akaike information criterion (AICc) is a commonly used method for determining optimal bandwidth, which is described as follows:

https://doi.org/10.5194/isprs-archives-XLIII-B4-2022-163-2022 |

$$AICc = -2L\,(\beta, \alpha) + 2k + \frac{2k\,(k+1)}{n-k-1} \qquad (4)$$

Where L(β,α) is the log-likelihood of GWNBR model and k is the effective number of parameters. The GWNBR's k could be calculated based on k=k1+k2, where k1 and k2 are the effective numbers of parameters of α and β. The GWNBR model can evolve into two models depending on whether the overdispersion parameter varies over space or constant over space. The one with spatially varied data is known as local GWNBR, and the one with the same data across the entire research area is known as global GWNBR. The global GWNBR has a k2 of one, and the local GWNBR has been difficult to estimate until now.

On the other hand, the root mean squared error (RMSE) is used as other criteria to evaluate the model performance, which can be presented as:

$$RMSE = \sqrt{\frac{1}{n}\sum\left(y_j - \hat{y}_j\right)^2} \qquad (5)$$

where y_j is the observed number of delivery motorcycle crashes, $\hat{y}_j$ is the predicted number of delivery motorcycle crashes, and n is the number of spatial units.

3. RESULTS AND DISCUSSIONS

3.1 Model Performances

The count data models were chosen for this study because the response variable was the frequency of crashes, which has a skew distribution. Furthermore, this study attempts to incorporate over-dispersion distribution into a GWR model in order to observe the impact of overdispersion on non-stationary crash frequency modeling. Based on the aforementioned methodology, two models, the GLM and GWNBR, were developed to examine the effect of overdispersion on crash frequency analyses. The models described above were calculated using R and SAS software macros created by Da Silva and Rodrigues (2014). By minimizing the AICc value, the optimal bandwidth for GWNBR was obtained.

To compare the performance of the three models mentioned above, three criteria were used: the correct Akaike information criterion (AICc), log-likelihood, and RMSE. The lower value of RMSE and AICc of the models indicates the better the model's performance. While the models with higher log-likelihood values show a better fitting and have an advantage over others. The results are shown in Tables 2 and 3 where the GLM model had a lower log-likelihood than the GWNBR model. The spatial model clearly outperforms the non-spatial model because the local model provides more spatial variation than the GLM, which allows coefficients to vary spatially across observations. In terms of AICc, the GLM model performed worse than the GWNBR model. Similarly, the RMSE value for the GWNBR model was lower, indicating that this model produces the least amount of errors.

Table 4 presents the Moran's I statistics and the corresponding p-value for the three models' residuals. First of all, the Moran's I value decreased considerably after incorporating spatial effects and overdispersion in the data. Second, it should be noted that the spatial dependency becomes insignificant in the GWNBR models, which indicates that the spatial autocorrelation between the models' residuals can be effectively explained by the overdispersion and spatial heterogeneity.

Variable	GLM	GWNBR			
		Mean	**Min**	**Median**	**Max**
Intercept	0.722*	0.800	0.593	0.827	0.870
Commercial area POI					
Restaurant	0.018*	0.017	0.014	0.017	0.022
Supermarket	0.079	0.081	0.048	0.078	0.122
Mall	0.281	0.273	0.223	0.274	0.305
Other POI					
School	0.121	0.142	0.120	0.141	0.166
Hotel	0.192	0.179	0.112	0.179	0.239
Bus stop	0.053	0.057	0.025	0.046	0.127
Model Performance					
AIC	2123.3	2113.3782			
2 x Log-likelihood	-2107.328	-2090.046			
RMSE	7.049	6.455			

Table 2. Model comparison of delivery motorcycle crashes during the weekday.

Variable	GLM	GWNBR			
		Mean	**Min**	**Median**	**Max**
Intercept	-0.244*	-0.111	-0.472	-0.095	0.094
Commercial area POI					
Restaurant	0.016*	0.015	0.009	0.014	0.024
Supermarket	0.147*	0.113	0.063	0.113	0.187
Mall	0.367*	0.359	0.038	0.413	0.462
Other POI					
School	0.075	0.093	-0.026	0.111	0.152
Hotel	0.186	0.184	-0.028	0.204	0.341
Bus stop	0.104*	0.101	0.073	0.091	0.170
Model Performance					
AIC	1426.5	1420.026			
2 x Log-likelihood	-1410.54	-1385.828			
RMSE	2.544	1.821			

Table 3. Model comparison of delivery motorcycle crashes during the weekend.

https://doi.org/10.5194/isprs-archives-XLIII-B4-2022-163-2022

Model	Moran's I	P-value
NB-Delivery Motorcycle Crashes on Weekdays	0.177	0.0001
NB-Delivery Motorcycle Crashes on Weekends	0.149	0.0001
GWNBR-Delivery Motorcycle Crashes on Weekdays	0.036	0.463
GWNBR-Delivery Motorcycle Crashes on Weekends	0.018	0.822

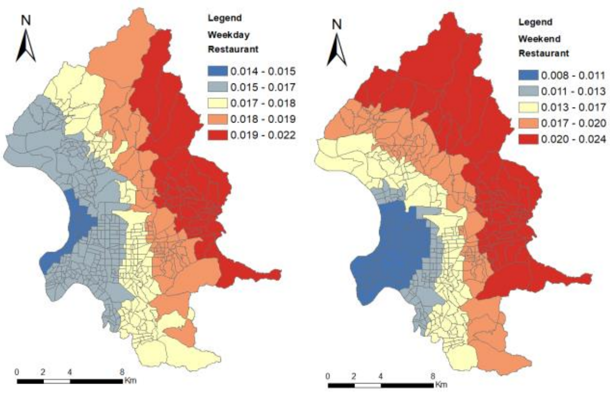
Table 4. Global Morans' I from each model.

3.2 Discussions

Tables 2 and 3 also show the results of the coefficient estimation. The coefficients of the global model (GLM) are provided in the second column, as are descriptive statistics for coefficients estimated by local models (GWNBR), such as the average, minimum, median, and maximum values are illustrated in the third until sixth column. In terms of magnitude, the parameters estimated by GWNBR models were comparable to GLM. The coefficients of GWNBR models, on the other hand, vary spatially, whereas the parameters of the GLM model are constant in the study area. In terms of the coefficients' significance, the number of restaurants had a significant relationship with delivery motorcycle crashes on weekdays and weekends. On the other hand, particular POI such as supermarket, shopping mall, and bus stop only had a significant association with delivery motorcycle crashes on weekend.

3.2.1 Spatial Analysis of Restaurant Coefficient: Figure 2 illustrates a positive and significant association between restaurant with delivery motorcycle crashes, meaning that areas with more restaurants tend to have more crashes. This result is consistent with previous studies, which concluded that restaurants may attract high pedestrian activity and more temporary parking, which may interfere with the traffic situation around this POI, then increasing the probability of traffic crashes if the driver is not aware of the situation nearby (Chen et al., 2019).

According to the local model results, the trend of coefficient values increased from urban to rural (from west to the upper east) during weekdays and weekends. This result is expected since there are few restaurants in the rural area, especially in northern Taipei where it belongs to the National Park. Therefore, increasing one restaurant in a rural area may attract neighboring citizens to visit and increase the probability of motorcycle crashes. However, there is a slight difference in the trend from weekdays to weekends. During the weekend, the cold spot gets bigger which indicates there is a lower probability of delivery motorcycle crashes in the urban area. This hypothesis might be true since, on a working day, the delivery service might be tended to go to the commercial and office area (located in the center of Taipei city) which could increase the traffic value and crash probability in the center area of Taipei during the weekday. Meanwhile, during the days off or weekend, online delivery service may tend to go to the residential area which is majority located in the suburban area of Taipei.

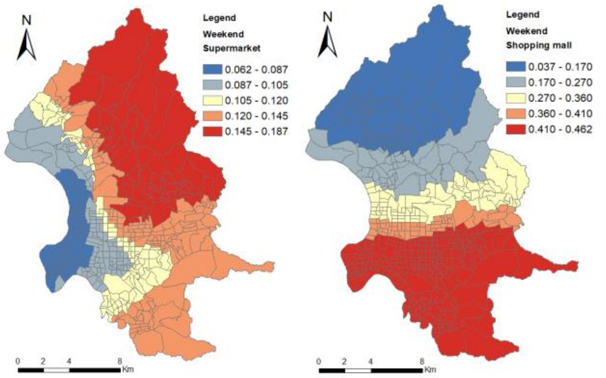


(a) Weekday (b) Weekend

Figure 2. The spatial distribution of restaurant parameter estimation delivery motorcycle crashes during weekdays and weekends.

3.2.2 Spatial Analysis of Commercial Areas with Delivery Motorcycle Crashes: Figure 3 shows commercial areas such as supermarkets have a significant positive relationship with delivery motorcycle crashes during weekends. This is consistent with the previous studies where that have noted that commercial land attracted more pedestrians and vehicles, leading to frequent traffic crashes, especially near large commercial centers (Huang et al., 2018). One of the reasons is because Taiwanese supermarkets recently collaborated with the online delivery service, making this POI become one of the most visited destinations for delivery services, and potentially increasing the delivery motorcycle crash probability significantly. For the local coefficient distribution, first, the supermarket has a similar spatial distribution to the restaurant where the coefficient value increases from the middle area (urban area) to urban to suburban and rural areas. In the other words, increasing the number of supermarkets in rural area may increase the probability of delivery motorcycle crashes even higher since it could attract nearby citizens.

On the other hand, shopping mall also has a significant positive association with delivery motorcycle crash on the weekend where the local coefficient value is increasing from urban to the suburban area (southern part of Taipei city). This result is expected since developing a new shopping area in the sub-urban area (which tends to have more residential land use) will greatly attract pedestrians and increase the chance of having traffic crashes. Furthermore, according to Pulugurtha et al (2013) and Xie et al (2019), the areas with mixed land-use of residential and commercial areas were found to be a hotspot of vehicle crashes. This condition may become worse during the weekend, where shopping malls may have higher traffic density during the weekend, then it also may raise the probability of delivery crash frequency from reckless drivers who prioritize their delivery speed.



(a) Supermarket (b) Shopping mall

Figure 3. The spatial distribution of commercial area (a) Supermarket; (b) shopping mall parameter estimation delivery motorcycle crashes on weekends

3.2.3 Spatial Analysis of Bus Stops with Delivery Motorcycle Crashes: According to figure 4, bus stop has a significant positive association with delivery motorcycle crashes on weekend. The positive signs suggest that a greater volume of bus stops is likely to increase the likelihood of public transport such as bus and delivery motorcycles intersecting. This result is consistent with a previous study from Wei and Lovegrove (2013) that indicated there was a potential conflict between motorcycles and buses under the similar road line. Public transportation, such as buses, frequently makes an immediate right turn to drop off passengers, which is extremely dangerous for delivery motorcycles riding in the right lane at the high speed. However, since both bus and delivery motorcycles are important transportation modes in promoting sustainable transportation, neither of them should be discouraged. To optimally isolate road traffic at bus station areas requires speed control, eyesight checking, and regulations on driving and riding behaviors. Furthermore, developing individual bus stop lines which is separating the bus lane and motorcycle lane could reduce collisions between motorcycles and buses, such as those already built in Taipei's urban area (please see Figure 4). This becomes one possible reason why the local coefficient value of bus stop increases from urban to the rural area. Since the bus stop in rural areas still does not have an individual line, increasing the number of bus stop in the rural area will produce a higher probability of delivery motorcycle crashes which often violates the speed limit on a rural road.

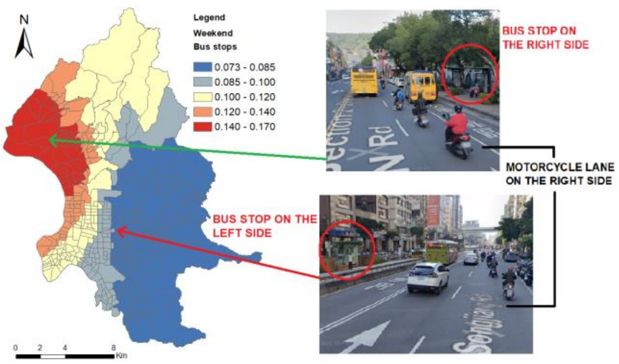


Figure 4. The spatial distribution of bus stops parameter estimation delivery motorcycle crashes on weekends

4. CONCLUSIONS

This study develops several statistical models to analyze the association between environmental factors with delivery motorcycle crashes. Among the tested models, the GWNBR model has been proven to be a powerful methodology for crash modeling compared to the GLM, which could capture spatial heterogeneity in crash data. According to the AICc, log-likelihood, and RMSE values, the GWNBR model outperformed the GLM. On the other hand, the Morans' I result showed that the spatial dependency of GWNBR models residual was insignificant, indicating that the spatial autocorrelation between the residuals can be effectively explained by overdispersion and spatial heterogeneity. In conclusion, the findings of this study demonstrated that incorporating overdispersion into spatial heterogeneity could improve crash modeling performance.

The GWNBR model's estimated coefficients show that number of restaurants have a significant association with delivery motorcycle crashes on weekday and weekend. As for other POI, commercial area, and bus stop only have a significant impact on delivery motorcycle crashes during weekend. The majority of the coefficient distribution is shifting from urban to suburban and rural areas, which indicate that rural area may be at a higher risk of vehicle crashes due to a lack of traffic policy. In the future, this result can be used to evaluate the daily safety situation and predict the number of crashes. In addition, these models can also be used to evaluate the efficacy of current policing policies or countermeasures in a specific area

REFERENCES

Abdel-Aty, M., Lee, J., Siddiqui, C. and Choi, K., 2013. Geographical unit based analysis in the context of transportation safety planning. *Transportation Research Part A: Policy and Practice, 49, pp.62-75.*

Cai, Q., Abdel-Aty, M. and Lee, J., 2017. Macro-level vulnerable road users crash analysis: A Bayesian joint modeling approach of frequency and proportion. *Accident Analysis & Prevention, 107, pp.11-19.*

Chen, Y., Ma, J. and Wang, S., 2019. Spatial Regression Analysis of Pedestrian Crashes Based on Point-of-Interest Data. *Journal of Data Analysis and Information Processing, 8(01), p.1.*

Chen, J., Liu, L., Xiao, L., Xu, C. and Long, D., 2020. Integrative analysis of spatial heterogeneity and overdispersion of crime with a geographically weighted negative binomial model. *ISPRS International Journal of Geo-Information, 9(1), p.60.*

Chung, Y., Song, T.J. and Yoon, B.J., 2014. Injury severity in delivery-motorcycle to vehicle crashes in the Seoul metropolitan area. *Accident Analysis & Prevention, 62, pp.79-86.*

Da Silva, A.R. and Rodrigues, T.C.V., 2014. Geographically weighted negative binomial regression—incorporating overdispersion. *Statistics and Computing, 24(5), pp.769-783.*

Hadayeghi, A., Shalaby, A.S. and Persaud, B.N., 2010. Development of planning level transportation safety tools using Geographically Weighted Poisson Regression. *Accident Analysis & Prevention, 42(2), pp.676-688.*

Huang, H. and Abdel-Aty, M., 2010. Multilevel data and Bayesian analysis in traffic safety. *Accident Analysis & Prevention, 42(6), pp.1556-1565.*

Huang, H., Abdel-Aty, M.A. and Darwiche, A.L., 2010. County-level crash risk analysis in Florida: Bayesian spatial modeling. *Transportation Research Record, 2148(1), pp.27-37.*

Huang, Y., Wang, X. and Patton, D., 2018. Examining spatial relationships between crashes and the built environment: A geographically weighted regression approach. *Journal of transport geography, 69, pp.221-233.*

Jia, R., Khadka, A. and Kim, I., 2018. Traffic crash analysis with point-of-interest spatial clustering. *Accident Analysis & Prevention, 121, pp.223-230.*

Kim, K., Pant, P. and Yamashita, E., 2010. Accidents and accessibility: Measuring influences of demographic and land use variables in Honolulu, Hawaii. *Transportation research record, 2147(1), pp.9-17.*

Lee, J., 2014. Development of Traffic safety zones and integrating macroscopic and microscopic safety data analytics for novel hot zone identification.

Lee, J., Abdel-Aty, M. and Jiang, X., 2014. Development of zone system for macro-level traffic safety analysis. *Journal of transport geography, 38, pp.13-21.*

Lee, J., Abdel-Aty, M. and Jiang, X., 2015. Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. *Accident Analysis & Prevention, 78, pp.146-154.*

Lee, J., Yasmin, S., Eluru, N., Abdel-Aty, M. and Cai, Q., 2018. Analysis of crash proportion by vehicle type at traffic analysis zone level: A mixed fractional split multinomial logit modeling approach with spatial effects. *Accident Analysis & Prevention, 111, pp.12-22.*

Levine, N., Kim, K.E. and Nitz, L.H., 1995. Spatial analysis of Honolulu motor vehicle crashes: I. Spatial patterns. *Accident Analysis & Prevention, 27(5), pp.663-674.*

Mathew, S., Pulugurtha, S.S. and Duvvuri, S., 2022. Exploring the effect of road network, demographic, and land use characteristics on teen crash frequency using geographically weighted negative binomial regression. *Accident Analysis & Prevention, 168, p.106615.*

Osama, A. and Sayed, T., 2016. Evaluating the impact of bike network indicators on cyclist safety using macro-level collision prediction models. *Accident Analysis & Prevention, 97, pp.28-37.*

Pljakić, M., Jovanović, D., Matović, B. and Mićić, S., 2019. Macro-level accident modeling in Novi Sad: A spatial regression approach. *Accident Analysis & Prevention, 132, p.105259.*

Pulugurtha, S.S., Duddu, V.R. and Kotagiri, Y., 2013. Traffic analysis zone level crash estimation models based on land use characteristics. *Accident Analysis & Prevention, 50, pp.678-687.*

Ukkusuri, S., Miranda-Moreno, L.F., Ramadurai, G. and Isa-Tavarez, J., 2012. The role of built environment on pedestrian crash frequency. *Safety science, 50(4), pp.1141-1151.*

Wang, X., Jin, Y., Abdel-Aty, M., Tremont, P.J. and Chen, X., 2012. Macrolevel model development for safety assessment of road network structures. *Transportation research record, 2280(1), pp.100-109.*

Wang, X., Yang, J., Lee, C., Ji, Z. and You, S., 2016. Macro-level safety analysis of pedestrian crashes in Shanghai, China. *Accident Analysis & Prevention, 96, pp.12-21.*

Wang, X., Zhou, Q., Yang, J., You, S., Song, Y. and Xue, M., 2019. Macro-level traffic safety analysis in Shanghai, China. *Accident Analysis & Prevention, 125, pp.249-256.*

Wang, Y. and Kockelman, K.M., 2013. A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. *Accident Analysis & Prevention, 60, pp.71-84.*

Wei, F. and Lovegrove, G., 2013. An empirical tool to evaluate the safety of cyclists: Community based, macro-level collision prediction models using negative binomial regression. *Accident Analysis & Prevention, 61, pp.129-137.*

Xie, B., An, Z., Zheng, Y. and Li, Z., 2019. Incorporating transportation safety into land use planning: Pre-assessment of land use conversion effects on severe crashes in urban China. *Applied geography, 103, pp.1-11.*

Xu, P. and Huang, H., 2015. Modeling crash spatial heterogeneity: Random parameter versus geographically weighting. *Accident Analysis & Prevention, 75, pp.16-25.*

Xu, C., Wang, Y., Ding, W. and Liu, P., 2020. Modeling the Spatial Effects of Land-Use Patterns on Traffic Safety Using Geographically Weighted Poisson Regression. *Networks and Spatial Economics, 20(4), pp.1015-1028.*