International Journal on Advanced Technology, Engineering, and Information System
https://ojs.transpublika.com/index.php/IJATEIS
Online ISSN 2828-5425

https://doi.org/10.55047/ijateis.v5i3.2240

Original Article

Spatial Analysis of Factors Affecting Traffic Accident Frequency in Jakarta Using a Geographically Weighted Regression Approach

Gamas Gagah Pangestu[1*], Sutanto Soehodo[2], Martha Leni Siregar[3]

[1-3]Department of Civil Engineering, Faculty of Engineering, Universitas Indonesia, Indonesia
Email: 1) gpgamas3@gmail.com, 2) sutantos@ui.ac.id, 3) leni@eng.ui.ac.id

Received : 23 April - 2026
Accepted : 01 June - 2026

Abstract

Traffic accidents are a complex issue influenced by road network characteristics, demographics, land use, and traffic conditions. This study analyzes the effects of these factors on accident frequency in DKI Jakarta and examines their spatial variation at the sub-district level. Using secondary data on accidents, road networks, population, land use, traffic, and regional activity, analysis was conducted using global regression (Ordinary Least Squares/OLS) and spatial regression (Geographically Weighted Regression/GWR), considering variations by severity, time period, and lighting. Results show that all factor groups significantly influence accident frequency, with effect magnitudes varying across areas. The GWR model captures local variations better than OLS, reflected by higher R^2 values in most conditions, while OLS remains effective for explaining global patterns. Findings indicate accident variations are more influenced by differences in the strength of factor effects across regions rather than the types of factors themselves, and are also affected by traffic operational conditions. This confirms that traffic accident characteristics are inherently spatial and contextual. The study highlights the importance of area-specific transportation safety planning, with tailored policy approaches to enhance intervention effectiveness in reducing urban accident rates.

Keywords: Geographically Weighted Regression (GWR), Ordinary Least Squares (OLS), Spatial Variation, Traffic Accidents, Transportation Safety.

1. Introduction

Road traffic accidents are one of the transport issues that have a significant impact on public safety, the economy and quality of life. According to the World Health Organization (2018), road traffic accidents are one of the leading causes of death worldwide, particularly among people of working age; consequently, this issue is not merely a safety concern but also has far-reaching social and economic implications. In major urban areas such as Jakarta, the complexity of this problem is increasing in line with high population mobility, vehicle growth, and the intensity of traffic interactions between road users, making the city a hub for national economic and governmental activities.

Conceptually, road traffic accidents are a multidimensional phenomenon influenced by various interacting factors. Previous research has consistently identified that road network characteristics such as traffic density, road type, and the number of junctions have a significant influence on the potential for traffic conflicts (Abdel-Aty & Radwan, 2000; Xie & Yan, 2013; El-Basyouny & Sayed, 2009). Furthermore, demographic factors, as indicators of road user exposure, such as age structure and gender, have also been shown to reflect differences in driving behaviour that influence accident risk (Islam & Mannering, 2006; Williams & Shabanova, 2003). Land-use activities in commercial and residential areas also contribute to increased traffic interactions that have the potential to cause accidents (Ewing & Dumbaugh, 2009; Pulugurtha et al., 2007), whilst traffic conditions such as congestion levels have a complex relationship with the frequency and severity of accidents (Zheng et al., 2010). Although these studies have provided a strong foundation, the majority of research has yet to integrate these various factors simultaneously within a single comprehensive analytical framework, nor has it considered variations in operational traffic conditions such as differences in time periods and lighting conditions. This gap in factor integration is further compounded by limitations in the analytical methods commonly employed to study them, which often fail to account for the spatial complexity inherent in urban accident patterns.

Given this knowledge gap, this study introduces several innovations that distinguish it from previous research. Firstly, this study simultaneously integrates various explanatory factors within a single analytical framework that encompasses road network characteristics, demographics as an exposure indicator, land use, and traffic conditions. Secondly, this study systematically combines OLS and GWR analyses to provide insights at both the global and local levels. Thirdly, this study considers variations in traffic operational conditions based on time periods (peak and off-peak) and lighting conditions (daylight and nighttime), which is still rarely done in area-based transport safety research. Fourthly, the analysis is conducted at a more detailed spatial level (the sub-district level) thereby enabling the depiction of more specific local variations and providing a more contextual interpretation of the magnitude of the influence between variables.

Based on the above, this study aims to: (1) analyse the influence of road network characteristics, population demographics as an exposure indicator, land-use patterns, and traffic conditions on the frequency of road traffic accidents in Jakarta; (2) analyse the presence of spatial heterogeneity (spatial non-stationarity) in the influence of these factors to identify differences in the magnitude and direction of influence between sub-districts; and (3) to analyse the most dominant spatial factors influencing local variations in accident frequency at the sub-district level and formulate their implications for area-based transport safety planning in Jakarta. With this more deepeer and spatially-based approach, the research is expected to make a tangible contribution to the development of transport safety analysis whilst supporting the formulation of more targeted policies in line with the characteristics of each region.

2. Literature Review

From a methodological perspective, there has been a significant shift from global regression approaches towards those that account for spatial aspects. Global regression approaches such as Ordinary Least Squares (OLS), which were widely used in previous research, assume that the relationship between variables is constant across the entire area (Lord & Mannering, 2010; Hadayeghi et al., 2003). This assumption is often not met in the context of heterogeneous urban areas such as Jakarta, given the differences in characteristics between sub-districts in terms of population density, road network structure, and land-use patterns. To address these limitations, Geographically Weighted Regression (GWR) was developed as an approach that allows for the estimation of different regression parameters for each location, thereby enabling a more representative capture of variations in influence across regions (Brunsdon et al., 2002; Yu & Abdel-Aty, 2014). Several studies have also demonstrated spatial associations in accident distribution, indicating that accident occurrences exhibit specific, non-random patterns across regions (Aguero-Valverde & Jovanis, 2006; Miaou & Lum, 1993). Nevertheless, most studies still examine global and spatial aspects separately without systematic integration (Chiou & Fu, 2015; Erdogan, 2009; Yu & Abdel-Aty, 2014).

Despite the growing body of spatial accident research, studies specifically examining Jakarta remain limited in scope. Existing studies on Indonesian urban road safety have tended to focus on corridor-level or citywide aggregate analyses rather than sub-district-level spatial heterogeneity. Studies simultaneously integrating road network, demographic, land-use, and traffic condition variables within a spatially explicit framework at the sub-district level in Jakarta are also limited. These gaps motivate the present study's approach of combining OLS and GWR within a multi-factor analytical framework.

3. Methods

3.1. Type of Research

The type of research was determined based on three main factors identified by Yin (2003), namely the type of research question, the extent of the researcher's control over events, and the focus on contemporary or historical events. Based on these considerations, this study employs an archival analysis strategy to address all the research questions, given that the research questions are of a "how" and "what" nature, the researcher has no control over the events under investigation, and the data used are secondary data that may be of a contemporary or historical nature. For the third research question, which is more interpretative in nature, a case study approach is also employed to identify dominant factors and their implications for region-based transport safety planning.

3.2. Research Stages

This study was conducted through eight systematic stages that integrated global and spatial analysis. The first stage involved the collection of secondary data, including data on road traffic accidents, road networks, population demographics, land use, traffic conditions, and points of interest (POIs). The second stage involved data processing, including data cleaning, integration across data sources, aggregation at the sub-district level, and the calculation of derived variables such as road density, junction density, and the Travel Time Index (TTI). The third stage comprised descriptive analysis to understand the distribution characteristics of variables through descriptive statistics and spatial visualisation.

The fourth stage is global modelling using OLS accompanied by classical assumption tests. The fifth stage is the evaluation of the OLS model based on AIC values, R^2, and the results of assumption tests as a basis for further analysis. The sixth stage is spatial analysis using GWR, preceded by a Moran's I test to identify spatial patterns, followed by GWR modelling with the determination of optimal bandwidth, local parameter estimation, coefficient analysis, area classification, and coefficient map visualisation. The seventh stage involved the systematic interpretation of model outputs in relation to regional characteristics and established transport theory, providing the analytical basis for drawing substantive conclusions. The eighth stage involved the development of region-based policy recommendations through a structured synthesis of findings, mapping intervention priorities to the specific spatial and demographic characteristics identified for each sub-district.

3.3. Types and Sources of Data

This study utilises secondary data obtained from various official agencies and reliable sources. The data used comprises spatial data in the form of the administrative boundaries of Jakarta's sub-districts as the unit of analysis, as well as non-spatial data covering all research variables. Specifically, traffic accident data is used as the dependent variable; road network data is obtained from OpenStreetMap (OSM); demographic data from the DKI Jakarta Health Department, integrated with Civil Registration and Population Administration (Dukcapil) data; land use data from the Geospatial Information Agency (BIG) and OSM; intersection density and traffic light data from the DKI Jakarta Communication and Information Technology Agency (Diskominfotik); and POI data from OSM.

Traffic condition data is represented by the Travel Time Index (TTI), derived from TomTom data processing using an edge-to-edge travel time approach. Data collection was carried out by identifying 50 sampling points at the boundaries of each sub-district, forming origin–destination pairs, and retrieving travel time data via the TomTom Routing API across five time scenarios: free-flow conditions, daylight peak (07:00), daylight off-peak (11:00), nighttime peak (18:00), and nighttime off-peak (23:00). The TTI value is then calculated using the formula:

$$TTI = \frac{t_{actual}}{t_{freeflow}}$$

and averaged across all pairs of points within a sub-district to produce a single representative value per area.

3.4. Data Processing Methods

Data processing was carried out in six stages. First, data cleaning to address incomplete data, duplicates and format inconsistencies. Second, the integration of data from various sources into a single database based on a consistent sub-district unit. Third, data aggregation from point or segment forms into aggregate values per sub-district, for example, road length into road density (km/km²) and the number of junctions into junction density (units/km²). Fourth, variable transformation through the calculation of densities, proportions, and ratios such as the TTI. Fifth, standardisation and normalisation to reduce scale differences between variables. Sixth, the compilation of a final dataset ready for use in all stages of analysis. The tools used include R for statistical analysis and GWR, QGIS for spatial processing and visualisation, Python for retrieving TTI data via the TomTom API, and Microsoft Excel for initial data processing.

3.5. Data Analysis Methods

3.5.1. Descriptive Analysis

Descriptive analysis was carried out to provide an overview of the frequency distribution of road traffic accidents across regions, temporal patterns, and spatial distribution in the form of thematic maps, in order to identify trends in the data prior to statistical modelling.

3.5.2. Linear Regression (Ordinary Least Squares/OLS)

The OLS model is used as an initial approach to identify the influence of independent variables on the overall accident rate. The OLS model is expressed as:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i$$

where y_i is the dependent variable, x_{ik} is the kth independent variable, β_k is the regression coefficient, and ε_i is the error term. Variable selection was performed through correlation analysis, stepwise regression, and multicollinearity testing using the Variance Inflation Factor (VIF). Although OLS has the advantage of being easy to interpret, this method is unable to capture spatial variation (spatial non-stationarity), so it was further developed using GWR.

3.5.3. Test of Classical Assumptions

Classical assumption tests were conducted to ensure that the OLS estimator is unbiased, efficient and consistent (Gujarati & Porter, 2010).

1) Normality Test

Normality tests were conducted using the Shapiro-Wilk test and the Jarque-Bera test. The Shapiro-Wilk statistic is calculated as:

$$W = \frac{\left(\sum_{i=1}^{n} a_i \, x_{(i)}\right)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Meanwhile, the Jarque-Bera statistic is formulated as:

$$JB = \frac{n}{6}\left(S^2 + \frac{(K-3)^2}{4}\right)$$

where S are ∗skewness∗ and K is *kurtosis*. The residuals are said to be normally distributed if the p-value $>$ 0,05.

2) Multicollinearity Test

Multicollinearity was detected using the Variance Inflation Factor (VIF):

$$VIF_k = \frac{1}{1 - R_k^2}$$

where R_k^2 is the coefficient of determination for the regression of the kth variable against the other independent variables. A VIF value of 10 or more indicates a high degree of multicollinearity.

3) Heteroscedasticity Test

The Breusch-Pagan test is used to detect heteroscedasticity, with the test statistic:

$$BP = n \cdot R^2$$

where R^2 obtained from a regression of the squared residuals against the independent variables. A p-value of ≤0.05 indicates the presence of heteroscedasticity.

4) Autocorrelation Test

The Durbin-Watson test is used to detect autocorrelation in the residuals, with the statistic:

$$DW = \frac{\sum_{i=2}^{n} (e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}$$

A DW value of approximately 2 indicates the absence of autocorrelation. In the context of spatial research, this test is supplemented by Moran's I to detect spatial autocorrelation, which forms the basis for the use of GWR.

3.5.4. Spatial Analysis

1) Spatial Autocorrelation Test (*Moran's I*)

Moran's I test is used to determine whether the distribution of accidents exhibits a specific spatial pattern (Anselin & Arribas-Bel, 2013). Moran's I statistic is defined as:

$$I = \frac{n}{S_0} \cdot \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \, (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

where:

n= number of observations
x_i= the value of the variable at i position
$\bar{x}$= average of the variables
w_{ij}= spatial weight between locations i and j

$$S_0 = \sum_i \sum_j w_{ij}$$

In addition to Moran's I, Local Indicators of Spatial Association (LISA) are also used to identify local clusters:

$$I_i = \frac{x_i - \bar{x}}{m^2} \sum_{j=1}^{n} w_{ij} \, (x_j - \bar{x})$$

The LISA results classify the regions into four cluster types: High–High (HH), Low–Low (LL), High–Low (HL) and Low–High (LH).

2) Geographically Weighted Regression (GWR)

GWR is an extension of OLS that allows regression parameters to vary spatially in order to accommodate spatial non-stationarity (Fotheringham et al., 2002). The GWR model is expressed as:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k \, (u_i, v_i) x_{ik} + \varepsilon_i$$

where (u_i, v_i) are the geographical coordinates of location i, and $\beta k(ui,vi)$ is a local parameter that may vary at each location. The parameters are estimated using the weighted least squares method:

$$\hat{\boldsymbol{\beta}}(u_i, v_i) = (X^T W_i X)^{-1} X^T W_i Y$$

where W_i is the spatial weight matrix for location i. The distance between locations is calculated using the Euclidean distance:

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2}$$

The spatial weights are calculated using the bisquare kernel function:

$$w_{ij} = \begin{cases} \left(1 - \left(\frac{d_{ij}}{b}\right)^2\right)^2, & d_{ij} < b \\ 0, & d_{ij} \geq b \end{cases}$$

where the bandwidth is determined using adaptive bandwidth via the Cross-Validation (CV) method to obtain the optimal model. The GWR output consists of local regression coefficients, local significance values, and local R² values for each sub-district, which are visualised in the form of coefficient maps to identify dominant factors and variations in influence across regions as a basis for formulating region-based transport safety policies.

4. Results and Discussion

4.1. Research Findings

4.1.1. Description of Accident Data

The road traffic accident data used in this study were collected for the period 2020–2024 at the sub-district level within the Jakarta Special Capital Region. The main variables used in this study include the number of accidents and accident density (number of accidents per km²), which serve as indicators to describe the spatial distribution of accidents.

Table 1. Traffic accident data for 2020–2024 at sub-district level in the Jakarta Special Capital Region

Variables	Min	Max	Mean	Std Dev
Number of accidents	7	704	136,75	91,69
Number of accidents per km²	2,82	46,2	11,15	7,44

Based on Table 1, the number of road traffic accidents in DKI Jakarta shows considerable variation between sub-districts, with a minimum of 7 incidents and a maximum of 704 incidents. The mean value of 136.75, with a standard deviation of 91.69, indicates a high degree of heterogeneity in the distribution of accidents across regions. Accident density (number of accidents per km²) also shows significant variation, with an average of 11.15 and a standard deviation of 7.44. This indicates that the intensity of accidents is not spatially uniform, with some areas having higher accident rates than others.

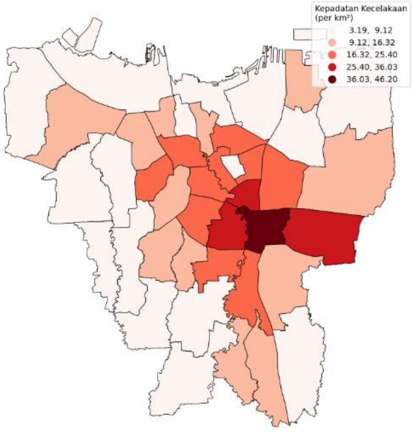


Figure 1. Spatial Distribution of Road Traffic Accidents by Sub-District in Jakarta

Spatially, the distribution of road traffic accidents in Jakarta is uneven. Figure 1 shows that areas with high accident rates tend to be concentrated in zones with heavy transport activity, such as the city centre and major transport corridors. Meanwhile, areas with lower accident rates are generally found in zones with relatively lower levels of activity.

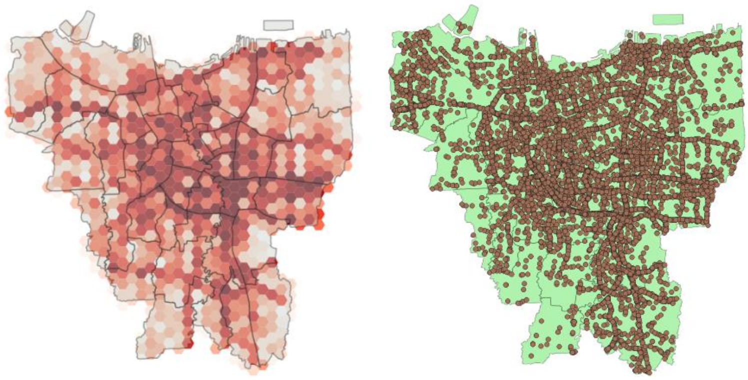


Figure 2. Spatial Distribution and Density of Road Traffic Accident Locations Using a Grid-Based Approach, Jakarta

In addition to representations based on administrative regions, the distribution of accidents can also be examined by location. Figure 3 shows the distribution of accident locations and the density of incidents using a grid-based approach, which highlights concentrations of incidents in specific areas.

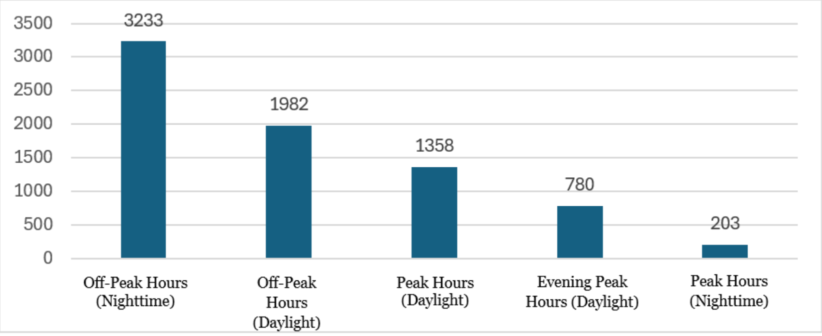


Figure 3. Distribution of Road Traffic Accidents by Time Period and Lighting Conditions

As shown in Figure 3, the distribution of accidents exhibits significant variation across different time and lighting conditions. The highest number of accidents occurred during off-peak hours (nighttime), with 3,233 incidents, followed by off-peak hours (daylight), with 1,982 incidents. Meanwhile, during peak hours, the number of accidents tended to be lower, with 1,358 incidents in well-lit conditions, 780 in the evening, and the lowest figure of 203 during peak hours (nighttime). These findings indicate that accidents are influenced not only by traffic density but also by lighting conditions and vehicle activity patterns, with off-peak conditions—particularly at night—still posing a high risk of accidents.

4.1.2. Global Analysis of Factors Influencing Road Traffic Accidents

A global analysis was conducted using the Ordinary Least Squares (OLS) regression method to identify the influence of independent variables on the frequency of road traffic accidents under various time and lighting conditions. This model was used as an initial approach, assuming that the relationship between variables is constant across the entire study area (Wooldridge, 2016).

Table 2. Results of the Ordinary Least Squares (OLS) Regression Model

No	Types of Accidents	Conditions	Significant Variables	Direction of Influence	R^2	Breusch-Pagan Test (P-value)	Durbin-Watson Test (P-value)
1.	Heavy	Peak (Daylight)	Arterial_road_density, Population_density, Male_proportion	+, +, -	0.36	0.51	0.58
		Off-peak (Daylight)	Motorway_density, Commercial_land_use_ratio	+, +	0.33	0.19	0.99
		Peak (Nighttime)	Collector_road_density, Signalised_junction_density, POI_per_km²	+, -, -	0.35	0.21	0.11
		Off-peak (Nighttime)	Arterial_road_density, Collector_road_density	+, +	0.4	0.22	0.77
2.	Medium	Peak (Daylight)	Motorway_density, Arterial_road_density, Residential_land_use_ratio, Commercial_land_use_ratio	+, +, +, -	0.48	0.23	0.94

No	Types of Accidents	Conditions	Significant Variables	Direction of Influence	R^2	Breusch-Pagan Test (P-value)	Durbin-Watson Test (P-value)
		Off-peak (Daylight)	Motorway_density, Population_density, Proportion_of_young_population, Commercial_land_use_ratio	+, +, -, -	0.43	0.43	0.78
		Peak (Nighttime)	Transport and utility land use ratio, TTI	+, +	0.32	0.22	0.65
		Off-peak (Nighttime)	Motorway density, Arterial road density, Residential land use ratio, POI_KM2	+, -, +, -	0.48	0.72	0.83
3.	Light	Peak (Daylight)	Motorway_density, Collector_road_density	+, +	0.29	0.46	0.45
		Off-peak (Daylight)	Motorway_density, Arterial_road_density	+, +	0.44	0.58	0.11
		Peak (Nighttime)	Junction_density, Proportion_of_adult_population, Commercial_land_use_ratio	+, -, -	0.37	0.05	0.79
		Off-peak (Nighttime)	Arterial_road_density	+	0.35	0.45	0.13

Based on the results of the analysis in Table 2, the values of the coefficient of determination (R^2) indicate the model's ability to explain traffic accidents. The R^2 values range from 0.24 to 0.48, indicating that the model is able to explain approximately 24% to 48% of the variation in traffic accidents. The highest value was found for moderate accidents, particularly under daylight peak and nighttime off-peak conditions ($R^2 = 0.48$), indicating that the variables used are sufficiently representative in explaining accident occurrences under those conditions.

4.1.3. Analysis of the Spatial Patterns of Road Traffic Accidents

To identify spatial correlations between areas, a spatial autocorrelation test was conducted using Moran's I index. The aim of this test is to determine whether accident rates in a given area are correlated with those in neighbouring areas (Anselin & Arribas-Bel, 2013).

Table 3. Spatial Autocorrelation Test (Moran's I)

Indicator	Value
Moran's I	0,0713
Expected I	-0,0244
Z-score	4,6402
p-value	0,00000174

Based on Table 3, a Moran's I value of 0.0713 indicates a tendency towards clustering, as it is positive. However, as the value is relatively small, the degree of spatial clustering observed is considered weak. Meanwhile, the p-value, which is significantly smaller than 0.05, indicates that the results are statistically significant. This means that the pattern formed is not a coincidence, but rather indicates a spatial relationship between regions. Thus, it can be concluded that the distribution of traffic accidents in the study area is not entirely random, but exhibits a tendency towards spatial clustering with a relatively low level of strength.

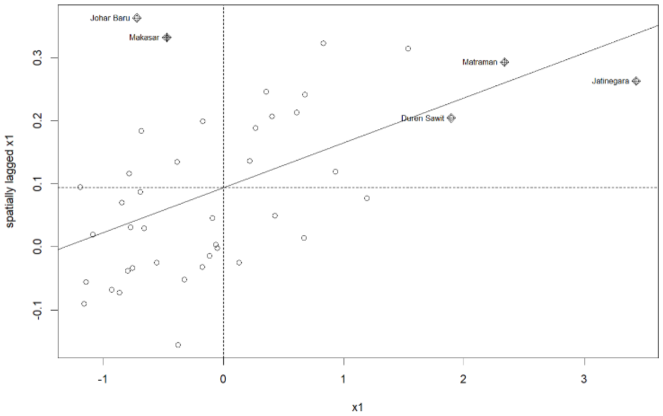


Figure 4. Moran's Scatterplot

A Moran's scatterplot is used to illustrate the relationship between accident rates in a given area and the average accident rate in neighbouring areas (spatial lag). This graph is divided into four main quadrants representing patterns of spatial relationships. The High–High quadrant (Quadrant I) indicates areas with high accident rates surrounded by areas with high rates. This quadrant indicates the presence of high-accident clusters. The Low–Low quadrant (Quadrant III) shows areas with low accident rates surrounded by areas with low rates, indicating low-accident clusters. The High–Low and Low–High quadrants indicate the presence of spatial outliers, i.e. areas with characteristics that differ from those of their surrounding areas. Based on Figure 4, it can be seen that some areas are located in the High–High and Low–Low quadrants, indicating a tendency towards spatial clustering. Areas such as Jatinegara and Matraman fall into the High–High category, indicating a high concentration of accidents.

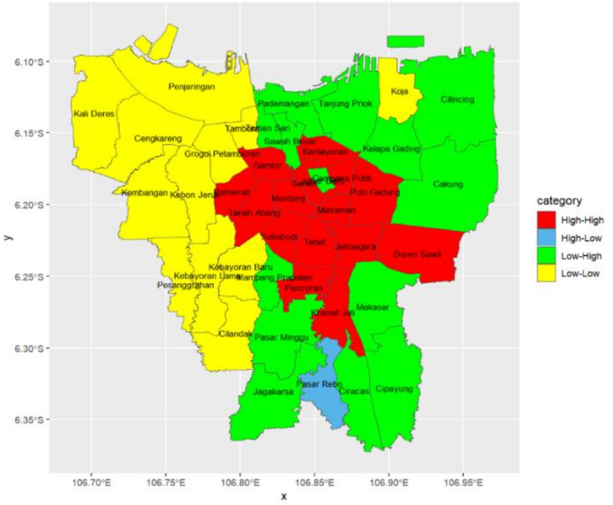


Figure 5. Local Indicators of Spatial Association (LISA) Analysis

The results of the analysis show that the central and parts of the eastern areas of Jakarta are dominated by the High–High category, indicating the presence of high-accident clusters. This suggests that areas with high accident rates tend to be located near other areas that also have high accident rates. Conversely, the western areas of Jakarta are dominated by the Low–Low category, indicating low-accident clusters. In addition, there are several areas falling into the High–Low and Low–High categories, indicating the presence of spatial outliers. As shown in Figure 5, these spatial patterns confirm the presence of localised clustering in accident distribution across Jakarta's sub-districts.

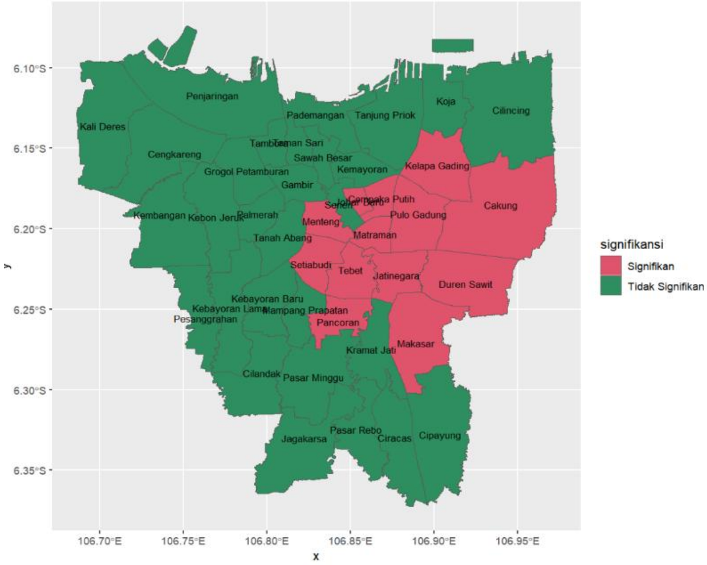


Figure 6. Significance Map

The significance map in the Figure 6 indicates that not all areas exhibit significant spatial association. Significant areas indicate a strong spatial relationship, whilst non-significant areas suggest that the patterns formed tend to be random. This indicates that although there is a tendency for spatial clustering, these patterns are not uniform across the entire region.

Based on the results of the Moran's I analysis, Moran Scatterplot, and LISA, it can be concluded that there is spatial correlation in the distribution of road traffic accidents, although the degree of clustering is relatively weak. Furthermore, the variation in local patterns shown by the LISA analysis indicates spatial heterogeneity, where the relationships between variables are not entirely uniform across the region. Therefore, global models such as OLS are deemed insufficient to capture this variation, necessitating the use of a Geographically Weighted Regression (GWR) approach capable of accommodating differences in influence between regions.

4.1.4. Spatial Analysis of Traffic Accident Factors Using GWR

1) Evaluation and Comparison of the OLS and GWR Models

An evaluation and comparison between the global model (OLS) and the spatial model (GWR) was carried out to assess the performance of each model in explaining variations in road traffic accidents. The GWR model used in this comparison was formulated through a parameter selection process, including bandwidth and kernel function, which will be explained in the following sub-section. This evaluation aims to assess the performance of each model in explaining variations in traffic accidents under various conditions. Model assessment was carried out using the coefficient of determination (R^2) and the Akaike Information Criterion (AIC). The R^2 value indicates the model's ability to explain data variation, whilst the AIC is used to evaluate the balance between the model's fit and its complexity. The results of the comparison between the OLS and GWR models are presented in Table 4.

Table 4. Comparison of OLS and GWR Model Performance by Accident Condition

No.	Accident Model	R^2 OLS	R^2 GWR	AIC OLS	AIC GWR
1.	Heavy – Peak Daylight	0.36	0.5	-52.49	-34.49
2.	Medium – Peak Daylight	0.48	0.51	-13.09	-10.34
3.	Light – Peak Daylight	0.29	0.37	50.76	53.02
4.	Heavy – Off-peak Daylight	0.33	0.46	-33.3	-23.45

5.	Medium – Off-peak Daylight	0.43	0.48	4.83	12.5
6.	Light – Off-peak Daylight	0.44	0.49	55.39	61.04
7.	Global – Peak Nighttime	0.28	0.5	249.39	249.46
8.	Heavy – Peak Nighttime	0.35	0.43	-126.53	-120.47
9.	Medium – Peak Nighttime	0.32	0.37	-127.57	-123.6
10.	Light – Peak Nighttime	0.37	0.49	-26.23	-23.81
11.	Global - Off-peak Nighttime	0.31	0.42	448.58	452.09
12.	Heavy – Off-peak Nighttime	0.4	0.43	-6.7	1.43
13.	Medium – Off-peak Nighttime	0.48	0.52	30.39	34.52
14.	Light – Off-peak Nighttime	0.35	0.42	59.4	62.13

Based on Table 4, the GWR model generally exhibits a higher R² value than the OLS model under most conditions. This indicates that GWR is better able to capture local variations in the data. An increase in the R² value is observed across almost all combinations of accident conditions, whether based on severity, time period, or lighting conditions.

However, when assessed by the AIC value, not all GWR models demonstrate better performance than OLS. Under certain conditions, the OLS model actually has a lower AIC value, such as in the Global–Off-peak Daylight, Light–Peak Daylight, and Light–Off-peak Nighttime conditions. This suggests that the global model remains more optimal for explaining the data as a whole under specific conditions.

The difference between the R² and AIC values indicates that whilst GWR is capable of improving the model's explanatory power locally, higher model complexity does not always result in a better model overall. In other words, an increase in R² with GWR is not always accompanied by an improvement in the AIC value. Consequently, the OLS model remains relevant as a global model, whilst GWR is used as an approach to explore variations in influence across regions. Therefore, further analysis focuses on GWR modelling to gain a deeper understanding of the spatial variations in traffic accident factors.

2) Selection of the GWR Model (Kernel and Bandwidth)

GWR modelling was carried out to capture the local variation in the influence of accident factors at the sub-district level. Three kernel functions were tested—Bisquare, Gaussian and Exponential—with selection criteria based on AIC (lower is better) and R² (higher is better).

Table 5. Results of GWR Kernel Function Selection

No.	Accident Model	Bisquare		Gaussian		Exponential	
		AIC	R2	AIC	R2	AIC	R2
1	Heavy – Peak Daylight	-15.54	0.58	-32.22	0.48	-34.49	0.5
2	Medium – Peak Daylight	-4.05	0.62	-10.34	0.51	-9.37	0.57
3	Light – Peak Daylight	56	0.45	53.02	0.37	54.8	0.44
4	Heavy – Off-peak Daylight	-14.81	0.55	-23.45	0.46	-18	0.577
5	Medium – Off-peak Daylight	25.34	0.66	12.5	0.48	16.19	0.55
6	Light – Off-peak Daylight	68.34	0.59	61.04	0.49	66.79	0.58
7	Heavy – Peak Nighttime	-114.08	0.51	-120.47	0.43	-118.03	0.49
8	Medium – Peak Nighttime	-117.67	0.44	-123.6	0.37	-121.18	0.42
9	Light – Peak Nighttime	-15.35	0.62	-23.81	0.49	-20.7005	0.57
10	Heavy – Off-peak Nighttime	14.62	0.5	1.43	0.43	5.78	0.51
11	Medium – Off-peak Nighttime	40.91	0.63	34.52	0.52	36.82	0.58
12	Light – Off-peak Nighttime	65.1	0.49	62.13	0.42	64.59	0.48

Based on Table 5, no single kernel type consistently yields the lowest AIC value across all models. However, the Gaussian and Exponential kernels tend to be more stable than the

Bisquare kernel. In some models, such as Global–Peak Light, the Exponential kernel produces a lower AIC than the Bisquare kernel (383.71 vs 408.06) and yields a higher R² than the Gaussian kernel (0.53 vs 0.42). A similar pattern is observed in the Global–Nighttime Peak model, where the AIC values for the Gaussian (249.79) and Exponential (249.46) kernels are relatively close, yet both outperform the Bisquare kernel (251.77).

For severity-based models, such as Severe–Daylight Peak, the Exponential kernel yields the lowest AIC (-34.49), whilst for Medium–Daylight Peak, the Gaussian kernel demonstrates the best performance (AIC -10.34). Meanwhile, for Medium–Daylight Peak, the Gaussian kernel also yields the lowest AIC (53.02). Under off-peak conditions, a similar pattern is observed, where no single kernel is absolutely dominant. However, the Gaussian and Exponential kernels generally provide a better balance between AIC and R² values.

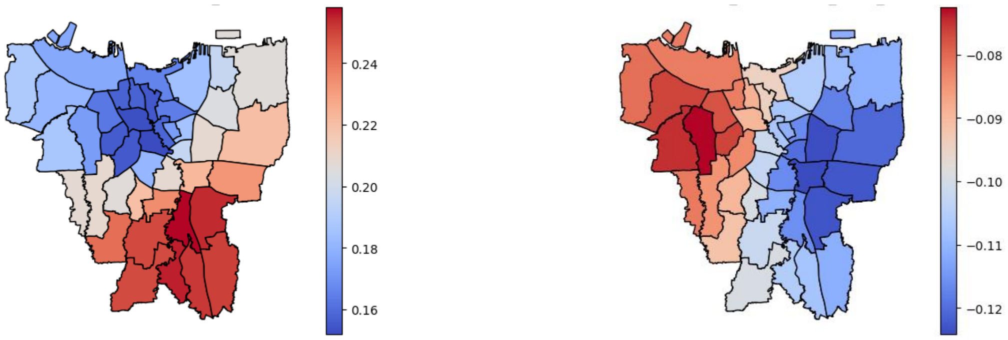
Overall, although the R² values for the Bisquare kernel tend to be higher in some models (for example, 0.83 in Global–Peak Light), the higher AIC values indicate that these models are less efficient than the other kernels. GWR analysis was performed on all combinations of these conditions, but the visualisations focus on representative conditions.

3) Analysis of GWR Local Coefficients

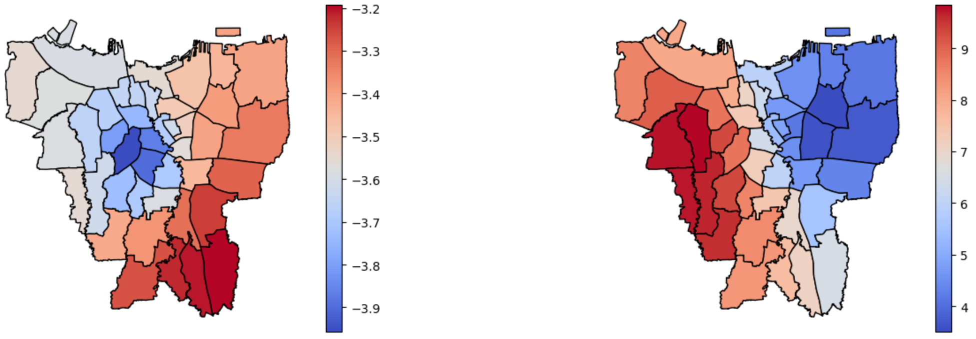
To understand the spatial variation in the influence of traffic accident factors, a map visualisation of the coefficients from the Geographically Weighted Regression (GWR) model was produced. The coefficient map was used to illustrate the magnitude and direction of the influence of each variable in every sub-district.

In this study, the analysis focused on minor accidents during peak periods in the dark (nighttime), representing a combination of high traffic intensity and limited visibility. The variables visualised include intersection density, commercial land ratio, proportion of adult population, and proportion of males, which are significant variables in the model.

Colour variations on the map indicate differences in the magnitude of influence, where red indicates a stronger influence (coefficient values are larger in absolute terms), whilst blue indicates a weaker influence. The visualisation of coefficient maps in this study focuses on the most representative key variables to facilitate interpretation. The complete coefficient maps for all variables and conditions are presented in the Figure 7 as supporting information.



a) b)



c) d)

Figure 7. GWR Local Coefficient Maps for Key Variables Influencing Minor Road Traffic Accidents During Peak Nighttime Conditions by Sub-District, Jakarta

Note: (a) GWR coefficients for intersection density; (b) GWR coefficients for commercial land use ratio; (c) GWR coefficients for adult population proportion; and (d) GWR coefficients for male population proportion

As illustrated in Figure 7, the mapping results show that intersection density has a positive effect across the entire region, with clear variations in magnitude. The southern and eastern areas of Jakarta have higher coefficients, indicating that an increase in intersections is associated with a higher potential for vehicle conflicts, whilst the north and west show a weaker effect. The ratio of commercial land has a negative effect across most of the region: the centre and west have coefficients close to zero, whilst the east is more negative, indicating that commercial activity does not necessarily increase the risk of accidents thanks to better infrastructure and traffic management. The proportion of the adult population shows a consistent negative influence; the south has a less negative coefficient, the centre a more negative one, which reduces the risk of accidents due to more cautious driving behaviour. The proportion of males has a positive influence, with the west and south having higher coefficients, the east lower, linked to a tendency towards aggressive driving. Overall, the magnitude and direction of the variables’ effects are not uniform, reflecting spatial heterogeneity. Nighttime-peak conditions amplify the effects of some variables due to high vehicle volumes and limited visibility. A further analysis of the range of coefficient values was conducted for each variable across all regions.

Table 6. Range of GWR Local Coefficient Values and Direction of Influence for Significant Variables Under Peak Nighttime Conditions, Jakarta

Variables	Min	Max	Direction
Intersection_density	0.15	0.26	+
Proportion_of_men	3.51	9.83	+
Proportion_of_adult_population	-3.96	-3.19	-
Commercial_land_use_ratio	-0.12	-0.07	-

As presented in Table 6, an analysis of the direction and magnitude of the effects was conducted by examining the range of minimum and maximum coefficient values for each variable across the entire region. The results of the analysis show that each variable has a different range of values, reflecting spatial variations in the strength of the effects. The variables of intersection density and male population proportion showed a positive direction of influence across the entire region, indicating that an increase in the values of these variables is associated with an increase in traffic accidents.

Conversely, the variables of adult population proportion and commercial land use ratio consistently showed a negative direction of influence, indicating that an increase in the values of these variables is associated with a decrease in traffic accidents. Although the direction of the variables' effects tends to be consistent, there is variation in the magnitude of the coefficients across regions. This suggests that the strength of each variable's influence is not spatially uniform. Thus, these results indicate the presence of spatial heterogeneity in the magnitude of the effects, although no variation in the direction of the effects was found across regions.

4) Identification of Dominant Factors

Dominance analysis was conducted based on the largest coefficient value in each region. The results show that the proportion of males is the most dominant factor across all regions. This indicates that demographic factors have a stronger influence than other variables in the model. However, this uniform dominance suggests that inter-regional variations are more evident in the magnitude of influence than in the type of dominant variable. Overall, the results of the GWR analysis indicate that the influence of traffic accident factors is not uniform across regions. Although the direction of the variables' influence tends to be consistent, there is variation in the magnitude of influence, indicating spatial heterogeneity. Furthermore, the results also show that the influence of variables can differ depending on traffic conditions, such as the time of day and lighting conditions. Thus, spatial approaches such as GWR provide a more comprehensive understanding than global models, particularly in identifying local variations that cannot be captured by OLS models.

4.2. Discussion

4.2.1. Relevance of the Findings to Theory and Previous Research

Based on the results of the ordinary least squares (OLS) regression analysis, a number of variables were found to significantly influence road traffic accidents across the various combinations of conditions analysed. These variables encompass aspects of the road network, demographics and land use.

Table 7. Overview of Significant Variables, Effect Direction, and Supporting References from the OLS Regression Analysis

Variables	Direction	Interpretation	Source
Motorway traffic density	+	High volume and high speed increase the likelihood of accidents	Abdel-Aty and Radwan (2000)
Arterial road traffic density	+	Numerous junctions and activities lead to increased traffic conflicts	Xie and Yan, (2013)
Collector road traffic density	+	Traffic flow transitions result in frequent manoeuvres, increasing the risk of accidents	Lee and Abdel-Aty (2005)
Junction traffic density	+	Increases the number of vehicle conflict points	Cheng et al. (2017)
	–	If controlled (signals/design), conflicts can be reduced	El-Basyouny and Sayed (2009)
Traffic density at signalised junctions	–	Traffic light control reduces direct conflicts	El-Basyouny and Sayed (2009)
Population density	+	Increased activity and exposure raise the likelihood of accidents	Wang and Kockelman (2013)
Proportion of men	+	More aggressive driving behaviour increases the risk of accidents	Islam and Mannering (2006)

Variables	**Direction**	**Interpretation**	**Source**
	–	The effect may be insignificant if environmental factors are more dominant	Islam and Mannering (2006)
Youth	+	Younger drivers are more aggressive and at higher risk	Williams and Shabanova (2003)
	–	More adaptable to traffic conditions	Williams and Shabanova (2003)
Adulthood	–	Tend to be more cautious and stable when driving	Kim and Yamashita (2007)
Residential land	+	Local activities increase traffic interactions	Wier et al. (2009)
	–	Low speeds in residential areas reduce accidents	Dumbaugh et al. (2009)
Commercial land	+	High economic activity increases vehicle conflicts	Pulugurtha et al. (2013)
	–	Good infrastructure can improve safety	Hadayeghi et al. (2003)
Industrial land	+	Heavy vehicles increase the risk of accidents	Miaou and Lum (1993)
Transport and utilities land	+	Transport hub activity increases traffic interactions	Cheng et al. (2017)
POI	+	Concentration of activity increases vehicle movement	Yu and Abdel-Aty (2014)
	–	Supporting facilities can reduce the risk of accidents	Yu and Abdel-Aty (2014)
TTI	+	Congestion increases vehicle interactions	Zheng et al. (2010)
	–	Low speeds reduce the severity of accidents	Zheng et al. (2010)

As summarised in Table 7, in general, the research findings indicate that variables relating to road network characteristics, such as junction density, have a positive influence on traffic accidents. This is consistent with traffic conflict theory, which posits that an increase in the number of points where vehicle flows converge increases the potential for accidents. From a demographic perspective, the proportion of males shows a consistent positive influence across various conditions, indicating that an increase in the proportion of the male population is associated with an increased risk of accidents. Conversely, the proportion of the adult population variable shows a negative influence, reflecting a tendency towards more stable and cautious driving behaviour.

Regarding land use, the commercial land ratio variable shows a predominantly negative influence in this study. This finding differs from some previous studies, which indicated a positive influence. This discrepancy suggests that the influence of commercial land is highly dependent on the regional context, such as infrastructure quality and traffic management. Overall, the results of this study are consistent with most previous research, although there are some differences indicating that accident factors are context-dependent.

4.2.2. Spatial Perspective

The results of the GWR analysis indicate that the influence of variables is not uniform across regions. Although the direction of the variables' influence tends to be consistent, there are variations in the magnitude of the influence, indicating spatial heterogeneity. These variations suggest that the same factors may have different levels of influence in different regions. Consequently, a global regression approach is insufficient to fully explain the phenomenon of road traffic accidents. The GWR approach is capable of providing more detailed information regarding local variations, thereby enabling the identification of regions with differing levels of sensitivity to each variable.

4.2.3. The Effect of Traffic Conditions on Variations in Results

The analysis was conducted across various combinations of conditions, including accident severity, time periods (peak and off-peak), and lighting conditions (daylight and nighttime). The results indicate that the influence of variables is determined not only by spatial characteristics but also by traffic operational conditions. Generally, during peak periods, variables related to traffic intensity tend to exhibit a stronger influence compared to off-peak periods. Furthermore, in dark conditions, some variables show changes in the magnitude of their influence, indicating the role of visibility factors. Although the detailed discussion focuses on representative conditions, the patterns found are generally consistent across other conditions, with the main difference lying in the strength of the variables' influence.

4.2.4. Implications for Transport Safety Planning

The research findings indicate that the factors influencing road traffic accidents are contextual and vary across regions. Therefore, transport safety planning approaches must take into account the local characteristics of each region. In areas strongly influenced by junction density, improvements in traffic management are required, such as junction regulation and improvements to road geometric design. Meanwhile, in areas where demographic factors have a strong influence, behaviour-based approaches are required, such as driver safety education. Furthermore, variations in influence based on traffic conditions indicate that safety policies need to be adapted to operational conditions, such as differences between peak and off-peak periods and between daylight and night-time conditions. Thus, the results of this study underscore the importance of a spatially-based approach in transport safety planning.

4.2.5. Policy Implications

Based on the results of the analysis conducted, a number of variables influencing road traffic accidents have been identified. Each variable has distinct characteristics of influence, thus requiring a specific, region-based policy approach. Policy implications have been formulated with reference to the direction of the variables' influence and are supported by previous research.

Table 8. Policy Implications for Road Traffic Accident Reduction by Variable Category, Effect Direction, and Literature Support

Variables	Direction of Influence	Policy Implications	Source
Motorway traffic density	+	Speed control, enhanced surveillance, and the implementation of intelligent transport systems (ITS) to minimise the risk of accidents	Abdel-Aty and Radwan (2000)
Arterial road traffic density	+	Road access management, junction management, and the mitigation of traffic conflicts through safer road design	Xie and Yan (2013)

Variables	Direction of Influence	Policy Implications	Source
Collector road traffic density	+	Management of traffic flow transitions and the improvement of safety facilities such as road signs and markings	(Lee and Abdel-Aty, 2005)
Junction traffic density	+	Optimisation of junction design, installation of traffic signals, and improved vehicle conflict management;	Cheng et al. (2017)
Signalised junction traffic density	–	Improved signal coordination and adaptive traffic control to reduce vehicle conflicts;	El-Basyouny and Sayed (2009)
Population density	+	Provision of pedestrian safety facilities, control of road activities, and improvements to public transport;	Wang and Kockelman (2013)
Proportion of males	+	Targeted driver safety education, safe driving behaviour campaigns, and traffic law enforcement;	Islam and Mannering (2006)
Young people	+	Driving education and training programmes for young drivers and restrictions on access for certain vehicles	Williams and Shabanova (2003)
Adults	–	Strengthening community-based safety programmes to maintain safe driving behaviour	Kim and Yamashita (2007)
Residential land	+	Implementation of traffic calming measures, speed limits, and management of local activities	Wier et al. (2009)
Commercial land	±	Vehicle access management, provision of organised parking facilities, and area traffic management	Pulugurtha et al. (2013)
Industrial land	+	Regulation of heavy goods vehicles, restrictions on operating hours, and enhanced traffic monitoring	Miaou and Lum (1993)
Transport and utility land	+	Integration of transport hubs, regulation of vehicle movements, and improved safety facilities	Cheng et al. (2017)
POI	±	Regulation of local activities, enhancement of supporting facilities, and traffic flow management	Yu and Abdel-Aty (2014)
TTI	±	Congestion management, adaptive traffic control, and increased road capacity	Zheng et al. (2010)

Table 8 shows that each influencing variable requires a different policy approach. Variables relating to the road network, such as junction density and arterial roads, require infrastructure-based and traffic management interventions. Meanwhile, demographic variables such as the proportion of males and age highlight the importance of behaviour-based approaches in improving road safety.

In terms of land use, the policies required are more focused on managing activities and regulating vehicle movement in specific areas. Furthermore, traffic condition variables such as TTI indicate that congestion management also plays a role in reducing the risk of accidents. Thus, the results of this study show that transport safety policies cannot be uniform, but must be tailored to the characteristics of the influencing variables and local conditions.

5. Conclusion

Road traffic accidents in Jakarta are a multidimensional phenomenon influenced by the interaction between the road network, demographics, land use, and traffic conditions. This directly addresses the first research objective, which sought to analyse the influence of multiple factor categories on accident frequency. The density of junctions, arterial roads, collector roads, and motorways generally has a positive effect as it increases complexity and the number of conflict points. Demographically, a higher proportion of male drivers increases the risk of accidents due to aggressive driving behaviour, whilst adult residents tend to be more stable and cautious. In terms of land use, commercial areas may exhibit a negative influence under certain conditions due to better infrastructure and traffic management. Traffic congestion levels (TTI) also have a dual effect: they increase vehicle interactions whilst reducing accident severity at low speeds.

Although not all models showed statistically significant spatial heterogeneity, GWR results revealed variations in coefficient magnitudes across sub-districts, indicating different regional sensitivities to accident factors. This addresses the second research objective, confirming that spatial heterogeneity exists in the influence of accident factors, albeit varying in strength across conditions. No single variable is consistently dominant across all regions and conditions; the influence of variables varies according to severity level, time period (peak vs. off-peak), and lighting conditions (daylight vs. nighttime). Thus, differences between regions are more reflected in the strength of influence than in the type of dominant variable. With regard to the third objective, the proportion of male drivers was identified as the most spatially dominant factor, though dominance remains context-dependent across regions and conditions. The implication is that transport safety planning must be area-specific, rather than uniform, given the differing characteristics and levels of sensitivity across sub-districts.

6. References

Abdel-Aty, M. A., & Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, *32*(5), 633–642. https://doi.org/10.1016/s0001-4575(99)00094-9

Aguero-Valverde, J., & Jovanis, P. P. (2006). Spatial analysis of fatal and injury crashes in Pennsylvania. *Accident Analysis & Prevention*, *38*(3), 618–625. https://doi.org/10.1016/j.aap.2005.12.006

Anselin, L., & Arribas-Bel, D. (2013). Spatial fixed effects and spatial dependence in a single cross-section. *Papers in Regional Science*, *92*(1), 3–17. https://doi.org/10.1111/j.1435-5957.2012.00480.x

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (2002). Geographically weighted summary statistics—a framework for localised exploratory data analysis. *Computers, Environment and Urban Systems*, *26*(6), 501–524. https://doi.org/10.1016/s0198-9715(01)00009-6

Cheng, W., Gill, G. S., Dasu, R., Xie, M., Jia, X., & Zhou, J. (2017). Comparison of Multivariate Poisson lognormal spatial and temporal crash models to identify hot spots of intersections based on crash types. *Accident Analysis & Prevention*, *99*, 330–341. https://doi.org/10.1016/j.aap.2016.11.022

Chiou, Y.-C., & Fu, C. (2015). Modeling crash frequency and severity with spatiotemporal dependence. *Analytic Methods in Accident Research*, *5*, 43–58. https://doi.org/10.1016/j.amar.2015.03.002

Dumbaugh, E., Rae, R., & Wunneberger, D. (2009). *Examining the relationship between community design and crash incidence*. https://rosap.ntl.bts.gov/view/dot/16799

El-Basyouny, K., & Sayed, T. (2009). Accident prediction models with random corridor parameters. *Accident Analysis & Prevention*, *41*(5), 1118–1123. https://doi.org/10.1016/j.aap.2009.06.025

Erdogan, S. (2009). Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. *Journal of Safety Research*, *40*(5), 341–351. https://doi.org/10.1016/j.jsr.2009.07.006

Ewing, R., & Dumbaugh, E. (2009). The built environment and traffic safety: a review of empirical evidence. *Journal of Planning Literature*, *23*(4), 347–367. https://doi.org/10.1177/0885412209335553

Gujarati, D., & Porter, D. C. (2010). *Dasar - Dasar Ekonometrika* (5th ed.). Salemba.

Hadayeghi, A., Shalaby, A. S., & Persaud, B. (2003). Macrolevel accident prediction models for evaluating safety of urban transportation systems. *Transportation Research Record*, *1840*(1), 87–95. https://doi.org/10.3141/1840-10

Islam, S., & Mannering, F. (2006). Driver aging and its effect on male and female single-vehicle accident injuries: Some additional evidence. *Journal of Safety Research*, *37*(3), 267–276. https://doi.org/10.1016/j.jsr.2006.04.003

Kim, K., & Yamashita, E. Y. (2007). Attitudes of commercial motor vehicle drivers towards safety belts. *Accident Analysis & Prevention*, *39*(6), 1097–1106. https://doi.org/10.1016/j.aap.2007.02.007

Lee, C., & Abdel-Aty, M. (2005). Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. *Accident Analysis & Prevention*, *37*(4), 775–786. https://doi.org/10.1016/j.aap.2005.03.019

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, *44*(5), 291–305. https://doi.org/10.1016/j.tra.2010.02.001

Miaou, S.-P., & Lum, H. (1993). Modeling vehicle accidents and highway geometric design relationships. *Accident Analysis & Prevention*, *25*(6), 689–709. https://doi.org/10.1016/0001-4575(93)90034-t

Pulugurtha, S. S., Krishnakumar, V. K., & Nambisan, S. S. (2007). New methods to identify and rank high pedestrian crash zones: An illustration. *Accident Analysis & Prevention*, *39*(4), 800–811. https://doi.org/10.1016/j.aap.2006.12.001

Wang, Y., & Kockelman, K. M. (2013). A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. *Accident Analysis & Prevention*, *60*, 71–84. https://doi.org/10.1016/j.aap.2013.07.030

Wier, M., Weintraub, J., Humphreys, E. H., Seto, E., & Bhatia, R. (2009). An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. *Accident Analysis & Prevention*, *41*(1), 137–145. https://doi.org/10.1016/j.aap.2008.10.001

Williams, A. F., & Shabanova, V. I. (2003). Responsibility of drivers, by age and gender, for motor-vehicle crash deaths. *Journal of Safety Research*, *34*(5), 527–531. https://doi.org/10.1016/j.jsr.2003.03.001

Wooldridge, J. M. (2016). *Introductory econometrics a modern approach*. South-Western cengage learning.

Xie, Z., & Yan, J. (2013). Detecting traffic accident clusters with network kernel density estimation and local spatial statistics: an integrated approach. *Journal of Transport Geography*, *31*, 64–71. https://doi.org/10.1016/j.jtrangeo.2013.05.009

Yin, R. K. (2003). Designing case studies. *Qualitative Research Methods*, *5*(14), 359–386.

Yu, R., & Abdel-Aty, M. (2014). Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. *Safety Science*, *63*, 50–56. https://doi.org/10.1016/j.ssci.2013.10.012

Zheng, Z., Ahn, S., & Monsere, C. M. (2010). Impact of traffic oscillations on freeway crash occurrences. *Accident Analysis & Prevention*, *42*(2), 626–636. https://doi.org/10.1016/j.aap.2009.10.009