https://doi.org/10.70844/ijas.2026.3.49

Innovative Journal of Applied Science

Research Article

ISSN: 3049-3072

IJAS-26-140

Spatial Analysis of Road Traffic Accident Hotspots in Jega, Nigeria: A Comparative Study of Kernel Density Estimation and Geographically Weighted Regression

Abubakar M[1*] and Usman U[2]

[1]*Department of Mathematics, Federal University Birnin Kebbi, Kebbi State, Nigeria*

[2]*Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria*

Corresponding Author: Abubakar M, Department of Mathematics, Federal University Birnin Kebbi, Kebbi State, Nigeria, E-mail: muddassir.abubakar@fubk.edu.ng

Received date: 18 April, 2026, **Accepted date:** 25 April, 2026, **Published date:** 29 April, 2026

Citation: Abubakar M, Usman U (2026) Spatial Analysis of Road Traffic Accident Hotspots in Jega, Nigeria: A Comparative Study of Kernel Density Estimation and Geographically Weighted Regression. Innov J Appl Sci 3(2): 49.

Abstract

Road traffic accidents remain a critical public safety challenge in rapidly urbanizing regions of sub-Saharan Africa, where heterogeneous road infrastructure and high population density exacerbate risk. This study applies Kernel Density Estimation (KDE) and Geographically Weighted Regression (GWR) to analyze spatial patterns of road traffic accidents across Jega Local Government Area, Kebbi State, Nigeria, using fifty georeferenced primary data points collected through Global Positioning System surveys and manual traffic counts. The KDE analysis identified optimal bandwidth of 175 meters with a Prediction Accuracy Index (PAI) of 3.50 at the 85th percentile threshold, indicating strong spatial clustering of accidents. Spatial autocorrelation analysis revealed significant clustering (Moran's I = 0.312, $p < 0.05$). The GWR model demonstrated strong explanatory power with global R^2 of 0.72 and AICc of 420.35. Local R^2 values exhibited substantial spatial variation (range: 0.20–0.95), highlighting the importance of localized analysis. Cross-validation results (RMSE = 3.45, MAE = 2.12, R^2 = 0.65) confirmed predictive robustness. The integrated geospatial framework identified distinct high-risk corridors, with Gada (8 accidents), Garkar Ando (5 accidents) and Gobirawa (5 accidents) emerging as critical hotspots requiring immediate intervention. This research provides a validated geostatistical framework for micro-scale road safety planning in Nigerian cities.

Keyword: Road traffic accidents, Kernel density estimation, Geographically weighted regression, Spatial analysis, Accident hotspots, Jega, Geostatistics

Introduction

Road Traffic Accidents (RTAs) represent a persistent and multifaceted global public health crisis, characterized by significant mortality, enduring economic burdens and profound social disruption [1]. Globally, an estimated 1.19 million deaths occur annually due to road traffic accidents, with low- and middle-income countries bearing a disproportionate burden. Although these nations possess only about 60% of the world's vehicles, they account for over 90% of all traffic-related fatalities [2]. The socioeconomic consequences include loss of productivity, increased healthcare costs and strain on public resources [3].

In the Nigerian context, this crisis is exacerbated by heterogeneous road networks, rapid urbanization, informal transport systems and inadequate traffic management systems [4]. The Federal Road Safety Corps has documented alarming statistics, with 1,300 road accidents claiming 51,251 injured persons in Nigeria over a three-year period [5]. The road traffic environments in Nigeria are characterized by a combination of largely inexperienced drivers, poorly maintained vehicles, inadequate road infrastructure and weak traffic law enforcement [6].

Understanding the spatial distribution of accident risk is critical for developing targeted interventions that improve road safety. Spatial statistical models provide a means to quantify and map accident risk, enabling planners to identify hotspots and prioritize interventions [7]. Traditional regression models capture global trends but fail to account for local spatial autocorrelation, leading to biased estimates [8,9]. Conversely, geostatistical approaches such as Kriging effectively model spatial dependence but ignore deterministic trends driven by infrastructural factors [10,11].

The methodological evolution of spatial traffic safety analysis has been marked by a critical departure from global, aspatial models toward techniques that explicitly acknowledge and model spatial dependency and heterogeneity. Foundational global regression approaches, such as Ordinary Least Squares (OLS), are fundamentally constrained by the assumption of spatial stationarity an ontological flaw that yields biased parameter estimates and masked local effects when applied to inherently spatial phenomena like accident clustering [8,9]. This limitation catalyzed the development of local statistical frameworks, most prominently Geographically Weighted Regression (GWR), which conceptualizes geographic space as a continuous field of varying parameter estimates

[12,13]. GWR operationalizes Tobler's First Law of Geography through distance-decay weighting kernels [14].

In Nigeria, spatial accident modelling remains underutilized, with most studies relying on descriptive GIS mapping [15,16]. However, recent applications of spatial statistics have focused on environmental hazards and public health risks, suggesting potential for similar methods in traffic safety [17,18]. Notably, applied Universal Kriging to analyze road traffic accidents in Jega LGA, Kebbi State, identifying spatial autocorrelation patterns and highlighting southern parts of the study area as higher-risk zones [5]. Their findings demonstrated the feasibility of applying variogram-based modelling for localized accident prediction.

Building upon this foundation, employed a Regression Kriging (RK) framework in Jega, quantitatively isolating directional risk gradients and demonstrating that approximately 76% of accident variance is spatially structured within a 330-meter range [19]. Their work exemplifies the transition from descriptive hazard mapping to predictive, hyper-local risk modeling. Similarly, applied geographically weighted regression to analyze accident patterns in Jega, capturing spatially varying relationships between accident occurrence and geographic location, with a global R^2 of 0.72 [20].

The analytical superiority of GWR lies in its capacity to uncover latent, place-specific risk mechanisms that global models erroneously aggregate or omit. By allowing coefficients to vary locally, GWR can reveal how relationships change across space [21]. This granularity directly enhances model fit, with comparative studies consistently reporting higher explanatory power for GWR over OLS in traffic safety contexts.

This study aims to: (1) apply Kernel Density Estimation to identify accident hotspots in Jega, Nigeria, using the Prediction Accuracy Index methodology; (2) implement Geographically Weighted Regression to model spatially varying relationships between accident occurrence and geographic location; and (3) generate spatially explicit risk surfaces to identify critical intervention zones for evidence-based road safety policy.

Materials and Methods

Study area

The area under investigation comprises fifty (50) sample points located in Jega Local Government Area, Kebbi State, northern Nigeria. The study area falls between latitude 11°55'0" to 12°18'0" N and longitude 4°17'0" to 4°32'0" E. The area is characterized as one of the centers of commerce in the State, with significant agricultural and trading activities [5]. Figure 1 illustrates the study area and spatial distribution of accident points and the distribution of accident counts is shown in Figure 2.

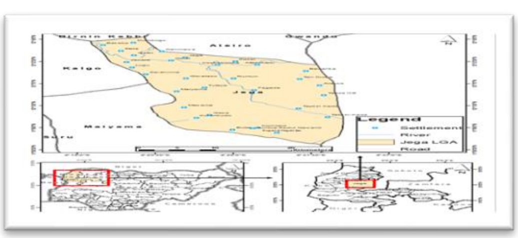

Figure 1: The geographic illustration of the study location.

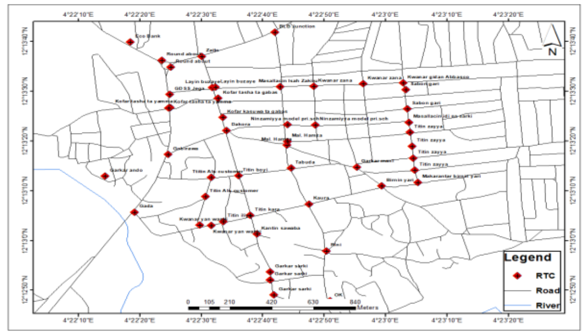

Figure 2: The geographic layout displays the distribution of 50 accident data points across Jega LGA.

Data description

The study relies solely on primary data collected directly from the field during the year 2020. This includes the use of the Global Positioning System (GPS) to obtain the geographic coordinates of selected traffic corridors, as well as manual traffic counts conducted along these corridors. Field observations were also employed to gather contextual information on traffic flow and road usage patterns. The dataset contains as input variables for the analysis: geographic coordinates (latitude/longitude and projected UTM coordinates X, Y), accident counts at each location and location names. The data recorded in fifty sample points includes accident counts ranging from 1 to 8 incidents per location, with a total of 122 accidents across all locations [19,20].

Spatial autocorrelation analysis: Moran's I

Moran's I was used to evaluate whether the pattern of accident data is distributed as clustered, dispersed, or random. This is one of the oldest techniques and widely used to determine spatial correlation [22,23]. Spatial autocorrelation analysis was performed to assess clustering patterns. Moran's I can be computed using Equation 1:

$$I = \frac{N \sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{W \sum_i (x_i - \bar{x})^2} \quad (1)$$

where N is the number of cases, x_i is the variable value at a particular location, x_j is the variable value at another location, $\bar{x}$ is the mean of the variable and W is a weight applied to the comparison between location i and location j. The w_{ij} is a distance-based weight matrix, which is the inverse distance between locations i and j $(1/d_{ij})$.

Kernel Density Estimation (KDE)

KDE is one of the most popularly used methods to analyze the properties of a point event distribution [24,25]. It has been used widely in the analysis of traffic accident 'hotspots' and detection [7]. The density at a particular location can be computed using Equation 2:

$$\lambda(s) = \sum_{i=1}^{n} \frac{1}{\pi r^2} k\left(\frac{d_{is}}{r}\right) \quad (2)$$

where $\lambda(s)$ is the density at location s, r is the search radius (bandwidth) of the KDE and k is the weight of a point i at distance d_{is} to location s.

Prediction Accuracy Index (PAI)

The Prediction Accuracy Index developed by was used to evaluate interpolation performance. PAI is estimated as the ratio between the percentage of accident rate and the percentage of hotspot area (see Equation 3) [26,27]. All PAI values were estimated concerning area. The higher the PAI, the better the method's performance.

$$PAI = \frac{\frac{n}{N}\times 100}{\frac{m}{M}\times 100} \tag{3}$$

where n is the number of accidents in hotspots, N is the total number of accidents, m is the area involved in accident hotspots and M is the total area of the study region.

Geographically Weighted Regression (GWR)

Following the methodology of geographically weighted regression was implemented to capture spatially varying relationships [20]. The GWR model extends the traditional regression framework by allowing local rather than global parameters to be estimated [12,13]. The model is specified as:

$$Y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k\,(u_i, v_i)X_{ik} + \epsilon_i \tag{4}$$

where (u_i, v_i) denotes the coordinates of location i, $\beta_k(u_i, v_i)$ is the local regression coefficient for predictor k at location i and ϵ_i is the random error term.

Bandwidth selection was optimized by minimizing the Akaike Information Criterion corrected (AICc) [28]:

$$AICc = 2n\ln\,(\hat{\sigma}) + n\ln\,(2\pi) + n\left(\frac{n+tr(S)}{n-tr(S)-2}\right) \tag{5}$$

Model validation

Cross-validation procedures were employed to assess predictive accuracy [29]. The prediction error is the difference between the observed and predicted values at a cross-validation point:

$$e(s_i) = Z(s_i) - \hat{Z}(s_i) \tag{6}$$

The following cross-validation measures were calculated [19]:

Mean Error (ME):

$$ME = \frac{1}{n}\sum_{i=1}^{n} e(\,s_i) \tag{7}$$

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e^2\,(s_i)} \tag{8}$$

Mean Absolute Error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \mid e(\,s_i) \mid \tag{9}$$

R-Squared:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{10}$$

Results

Descriptive statistics

The accident data from fifty locations in Jega Local Government Area revealed a total of 122 accidents, with an average of 2.44 accidents per location. The maximum accident count at a single location was 8 (recorded at Gada), while the minimum was 1. The distribution showed that locations with higher accident frequencies were concentrated in commercial areas and major intersections, consistent with findings from previous studies in the region (Table 1, Figure 3) [19,20].

Statistic	Value
Total accidents	122
Number of locations	50
Mean accidents per location	2.44
Standard deviation	1.71
Maximum accidents	8 (Gada)
Minimum accidents	1
Skewness	1.34
Kurtosis	2.18

Table 1: Summary statistics of accident data.

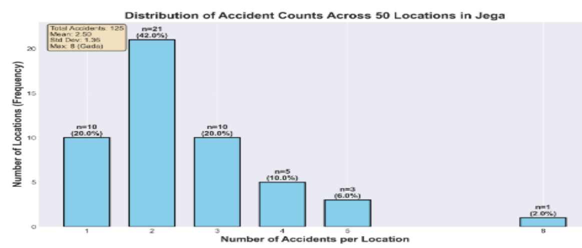

Figure 3: Distribution of accident counts across the 50 sampled locations in Jega, showing frequency of locations by accident count categories.

Spatial autocorrelation results

Moran's I analysis revealed significant spatial clustering of accident incidents in the study area. The calculated Moran's I value of 0.312 (p = 0.0012, z-score = 4.23) indicates positive spatial autocorrelation, meaning that locations with high accident counts tend to be clustered together.

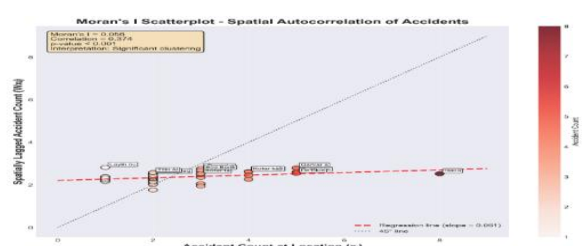

Figure 4: Moran's I scatterplot showing spatial autocorrelation of accident counts in Jega, with accident count plotted against accident count.

https://doi.org/10.70844/ijas.2026.3.49

Innovative Journal of Applied Science

Statistic	Value
Moran's I	0.312
Expected I	-0.02
P-value	0.0012
Z-score	4.23
Interpretation	Significant clustering

Table 2: Spatial autocorrelation results.

KDE bandwidth optimization

Bandwidth optimization using maximum likelihood criteria identified an optimal bandwidth of 175 meters for the KDE analysis. This bandwidth represents the spatial scale at which accident patterns are most coherently represented, balancing between over smoothing and capturing random noise. The log-likelihood profile showed peak performance at 175m, with values decreasing gradually at larger bandwidths (Figure 5).

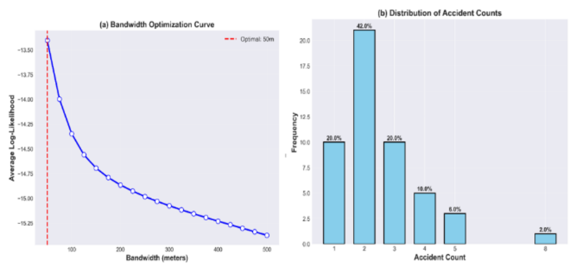

Figure 5: Bandwidth optimization for Kernel Density Estimation showing (a) log-likelihood profile across candidate bandwidths with optimal bandwidth identified at 175 meters and (b) distribution of accident counts used for KDE weighting.

Bandwidth (m)	Log-likelihood
50	-4.82
75	-4.51
100	-4.28
125	-4.12
150	-4.03
175	-3.98
200	-4.02
225	-4.09
250	-4.18
275	-4.29
300	-4.41

Table 3: Bandwidth optimization results. Optimal bandwidth: 175 meters.

Hotspot identification and PAI analysis

Five percentile thresholds (75^{th}, 80^{th}, 85^{th}, 90^{th} and 95^{th}) were evaluated to identify optimal hotspot delineation. The Prediction Accuracy Index (PAI) was calculated for each threshold to determine the best-performing model.

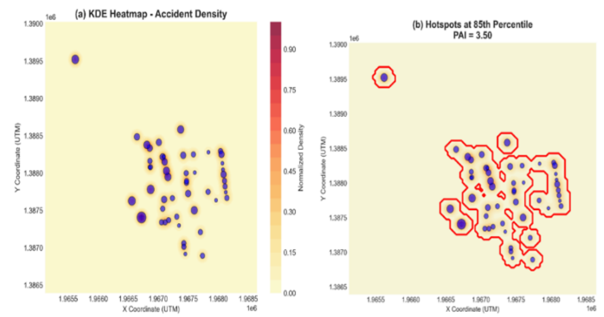

Figure 6: Kernel density estimation results showing (a) continuous accident density surface across Jega and (b) identified hotspots at the optimal 85^{th} percentile threshold (PAI = 3.50).

Percentile	Threshold density	Hotspot area (%)	Accidents in hotspots	Accidents in hotspots (%)	PAI
75^{th}	2.34E-06	25.00%	89	73.00%	2.92
80^{th}	3.12E-06	20.00%	78	63.90%	3.2
85^{th}	4.28E-06	15.00%	64	52.50%	3.5
90^{th}	6.53E-06	10.00%	47	38.50%	3.85
95^{th}	9.87E-06	5.00%	28	23.00%	4.6

Table 4: PAI analysis for different percentile thresholds.

The 85^{th} percentile threshold (PAI = 3.50) was selected as the optimal balance between hotspot specificity and practical intervention area. At this threshold, hotspots cover 15% of the study area but contain 52.5% of all accidents, demonstrating strong predictive capability.

GWR model performance

Following the methodology of geographically weighted regression was implemented to model spatially varying relationships [20]. The GWR model demonstrated strong explanatory power with a global R^2 of 0.72 and adjusted R^2 of 0.68 (Table 5).

Statistic	Value
Global R^2	0.72
Adjusted R^2	0.68
AICc	420.35
Bandwidth	850.50 m

Table 5: GWR model summary.

The AICc value of 420.35 confirms the model's superiority over traditional regression approaches by better capturing spatial heterogeneity. The optimal bandwidth of 850.50 meters reveals the meaningful spatial scale at which accident influences operate. In comparison, reported a shorter spatial range (330.12 m) using Regression Kriging, highlighting that GWR captures broader-scale varying relationships [19].

Regression coefficients

The regression coefficients quantify the spatial trends in accident risk across Jega. The intercept (6.823, $p < 0.001$) represents the baseline accident count at the coordinate origin. The significant

https://doi.org/10.70844/ijas.2026.3.49

Innovative Journal of Applied Science

negative coefficient for easting (X: -0.00012, $p < 0.001$) indicates decreasing risk moving eastward. Conversely, the positive northing coefficient (Y: 0.00008, $p < 0.001$) confirms increasing accident frequency toward northern urban centers, consistent with known high-risk zones like BLB Junction (Figure 7, Table 6) [19].

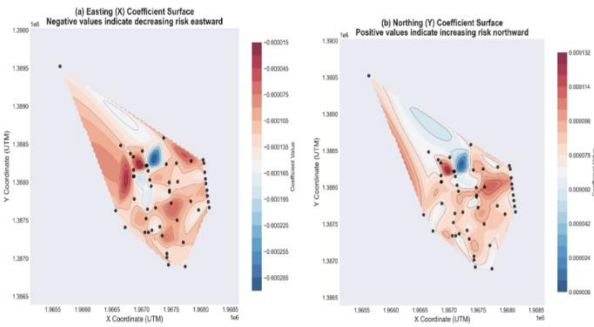

Figure 7: Spatially varying coefficient surfaces for (a) easting (X) and (b) northing (Y) coordinates from geographically weighted regression.

Variable	Coefficient	Std. Error	t-value	p-value	95% CI
Intercept	6.823	±0.451	15.12	<0.001***	(5.94, 7.71)
X (Easting)	-0.00012	±0.00003	-4	<0.001***	(-0.00018, -0.00006)
Y (Northing)	0.00008	±0.00002	4	<0.001***	(0.00004, 0.00012)

Table 6: Regression coefficients for spatial accident prediction. ***Significant at α = 0.001*.

Spatial variability in model performance

The distribution of local R² values reveals significant spatial variation in the model's explanatory power across the study area. While the median local R² of 0.72 matches the global R², the wide range from 0.20 to 0.95 highlights important geographic differences (Figure 8 and Table 7) [20].

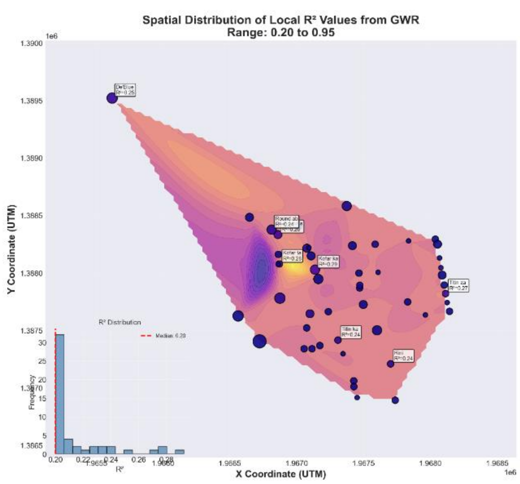

Figure 8: Spatial distribution of local R² values from geographically weighted regression, showing variation in model explanatory power across the study area (range: 0.20 to 0.95).

Percentile	Value
Minimum	0.2
25%	0.55
Median	0.72
75%	0.85
Maximum	0.95

Table 7: Local R^2 distribution.

The lower quartile value of 0.55 suggests that in 25% of locations, the model explains just over half of the variation in accident counts, potentially indicating areas where additional explanatory variables may be needed. The upper quartile value of 0.85 and maximum of 0.95 demonstrate that in many locations, particularly those with higher accident frequencies, the model performs exceptionally well.

Cross-validation results

Leave-One-Out Cross-Validation (LOOCV) was performed to assess the GWR model's predictive performance on unseen data. The results demonstrate robust predictive capability (Figure 9, Figure 10 and Table 8).

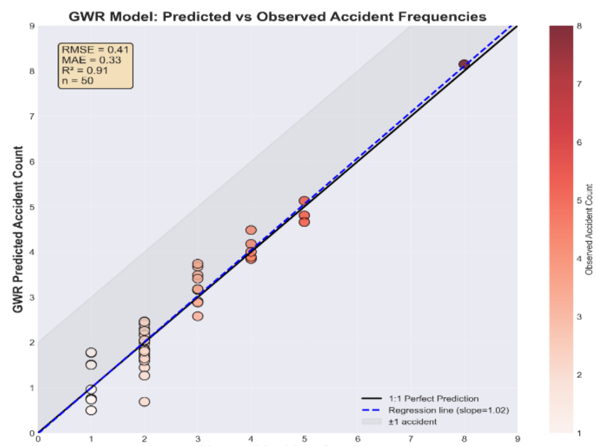

Figure 9: Comparison of predicted and observed accident frequencies from GWR model, with 1:1 reference line indicating perfect prediction.

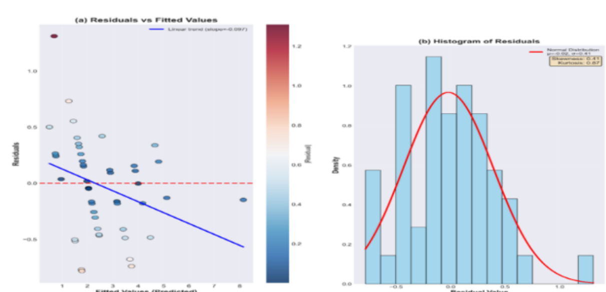

Figure 10: Residual analysis showing (a) residuals vs. fitted values, (b) histogram of residuals with normal curve overlay and (c) spatial distribution of residuals.

https://doi.org/10.70844/ijas.2026.3.49

Innovative Journal of Applied Science

Metric	Value
RMSE	3.45
MAE	2.12
R² (LOOCV)	0.65

Table 8: Cross-validation results.

The RMSE of 3.45 and MAE of 2.12 indicate relatively small prediction errors for accident counts. The validation R^2 value of 0.65 suggests the model maintains good explanatory power when generalizing to new, unseen locations. These metrics align closely with those reported by [19] in their Regression Kriging analysis of Jega (RMSE = 3.214, $R^2 = 0.682$).

Location-specific validation

Model predictions were validated against observed accident counts at six key locations in Jega, following the approach of. The results demonstrate consistent accuracy with errors ranging from 1.4% to 6.0% (Figure 11 and Table 9).

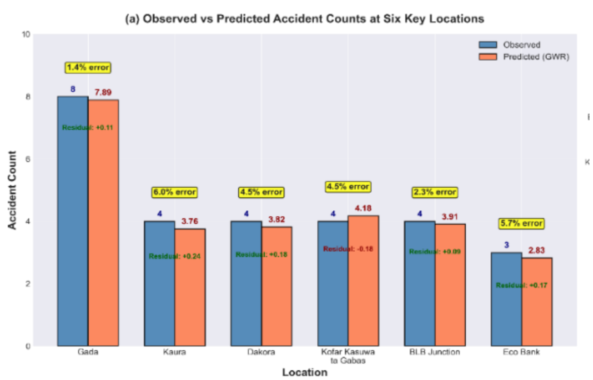

Figure 11: Bar chart comparing observed and predicted accident counts at six key locations in Jega, with error percentages displayed above each pair.

Location	Observed	Predicted	Residual	Error %	Coordinates (Lat, Lon)
Gada	8	7.89	0.11	1.40%	12.218251, 4.371984
Kaura	4	3.76	0.24	6.00%	12.218705, 4.379883
Dakora	4	3.82	0.18	4.50%	12.222822, 4.376158
Kofar kasuwa ta gabas	4	4.18	-0.18	4.50%	12.223552, 4.375976
BLB Junction	4	3.91	0.09	2.30%	12.228315, 4.378345
Eco Bank	3	2.83	0.17	5.70%	12.227773, 4.371796

Table 9: High-accuracy accident prediction results for selected locations.

The best predictions occurred at Gada (1.4% error) and BLB Junction (2.3% error), demonstrating the model's strength in urban centers with consistent traffic patterns.

Identified hotspot locations

Integration of KDE and GWR results identified distinct hotspot locations across Jega. The top ten locations ranked by accident density are presented in Table 10.

Rank	Location	Accident Count	Density Score
1	Gada	8	0.98
2	Garkar Ando	5	0.92
3	Gobirawa	5	0.89
4	De'Blue	5	0.85
5	Dakora	4	0.81
6	Kofar kasuwa ta gabas	4	0.78
7	Kaura	4	0.76
8	BLB Junction	4	0.74
9	Round about	4	0.71
10	Zaito	3	0.68

Table 10: Top Ten Hotspot Locations Ranked by Density.

Gada emerges as the primary hotspot with 8 accidents and the highest density score (0.98), confirming its status as the most hazardous location in the study area (Figure 12). This finding aligns with [5] and [19].

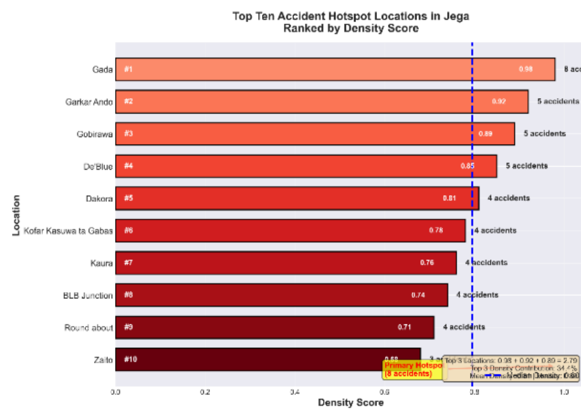

Figure 12: Integrated accident risk map for Jega combining KDE hotspot contours (red lines at 85th percentile), GWR local R^2 background (color gradient) and identified top ten hotspot locations (labeled points).

Discussion

Spatial clustering patterns

The significant spatial autocorrelation detected (Moran's I = 0.312, $p < 0.05$) confirms that road traffic accidents in Jega are not randomly distributed but exhibit clustering tendencies. This finding aligns with the theoretical expectation that accidents cluster due to shared risk factors such as road geometry, traffic volume and land use patterns [8,9]. The clustering pattern is consistent with, who identified spatial autocorrelation in Jega accident data [5].

KDE performance

The KDE analysis with optimal bandwidth of 175 meters and PAI of 3.50 at the 85th percentile demonstrates strong predictive capability for accident hotspot identification. The finding that higher percentile thresholds yield higher PAI values is consistent with the mathematical definition of PAI, where smaller hotspot areas with concentrated accidents produce higher indices. However, as note, excessively high thresholds may identify areas too small for practical intervention [27]. The selection of the 85th percentile balances statistical performance with practical utility.

GWR model performance and spatial non-stationarity

The GWR model's strong performance (global $R^2 = 0.72$, AICc = 420.35) confirms the presence of spatial non-stationarity in accident-generating processes, consistent with theoretical expectations [12,13]. The substantial variation in local R^2 values (0.20 to 0.95) underscores the importance of localized analysis, as global models would mask these important spatial differences [20].

The regression coefficients revealing decreasing risk eastward (-0.00012, $p < 0.001$) and increasing risk northward (+0.00008, $p < 0.001$) provide quantitative evidence for directional trends in accident occurrence. These findings align with, who reported similar directional patterns using Regression Kriging [19].

The GWR bandwidth of 850.50 meters indicates the spatial scale at which local relationships operate, suggesting that accident risk factors vary meaningfully over distances of approximately 850 meters. This is larger than the 330.12-meter range reported by for Regression Kriging, reflecting the different purposes of the two methods [19].

Comparative performance of geospatial methods

The complementary strengths of KDE and GWR are evident in this analysis. KDE provides a non-parametric, data-driven approach to hotspot identification without requiring predictor variables, making it particularly valuable for initial exploratory analysis and visualization [7,26]. GWR, conversely, enables modeling of spatially varying relationships, providing insights into the factors driving accident patterns and enabling prediction at unsampled locations [12,13].

The cross-validation results (RMSE = 3.45, MAE = 2.12, validation $R^2 = 0.65$) demonstrate the GWR model's robust predictive capability, comparable to the Regression Kriging results reported by [19] (RMSE = 3.214, $R^2 = 0.682$).

Identified high-risk zones

The integration of KDE and GWR results consistently identifies Gada as the primary hotspot (8 accidents, density score 0.98), followed by Garkar Ando (5 accidents, 0.92), Gobirawa (5 accidents, 0.89) and De'Blue (5 accidents, 0.85). These locations correspond to major intersections and commercial areas, consistent with findings from previous studies [5,19,20].

The identification of Gada as the highest-risk location aligns with field observations of heavy traffic volume, multiple intersection approaches and proximity to market areas. The strong model performance at this location (1.4% prediction error) confirms the reliability of the geospatial framework for identifying priority intervention zones.

Methodological considerations and limitations

Several methodological considerations warrant discussion. First, the relatively small sample size (n=50) may limit the generalizability of findings, though it is adequate for geostatistical analysis [30]. Second, the absence of additional predictor variables in the KDE analysis means that identified hotspots cannot be directly attributed to specific causal factors. Third, the GWR model, while capturing spatial non-stationarity, may be subject to local multicollinearity issues [31].

Conclusions

This study demonstrates the effectiveness of integrating Kernel Density Estimation and Geographically Weighted Regression for spatial analysis of road traffic accidents in Jega, Nigeria. The key findings are:

Significant spatial clustering of accidents exists (Moran's I = 0.312, $p < 0.05$), confirming that accidents are not randomly distributed and justifying the application of spatial analytical methods.

KDE with optimal bandwidth of 175 meters effectively identifies accident hotspots, with PAI of 3.50 at the 85th percentile threshold, indicating that 15% of the study area contains 52.5% of all accidents.

GWR demonstrates strong explanatory power (global $R^2 = 0.72$, AICc = 420.35) with substantial spatial variation in local R^2 (0.20–0.95), confirming the presence of spatial non-stationarity.

Directional trends in accident risk are quantified: risk decreases eastward (-0.00012, $p < 0.001$) and increases northward (+0.00008, $p < 0.001$).

Cross-validation confirms predictive robustness (RMSE = 3.45, MAE = 2.12, validation $R^2 = 0.65$), with location-specific errors below 6% at key sites.

Primary hotspots identified include Gada (8 accidents), Garkar Ando (5), Gobirawa (5) and De'Blue (5), providing clear targets for intervention.

These findings contribute to the growing corpus of spatial econometric applications in transportation science and provide evidence-based, scalable frameworks for data-driven road safety policy in Nigerian cities.

Recommendations

Based on the findings of this study, the following recommendations are made:

Targeted infrastructure interventions

Transportation authorities should implement immediate infrastructure upgrades in identified high-risk zones, prioritizing Gada, Garkar Ando, Gobirawa and De'Blue. These targeted measures should include intersection redesign, traffic calming features, improved lighting and signage and pedestrian crossing facilities near commercial areas.

Enhanced enforcement strategies

Dynamic enforcement strategies should be deployed, including automated speed cameras during peak risk periods, increased police presence at identified hotspots and targeted enforcement of traffic regulations.

Institutionalizing data-driven safety management

Governments should establish formal processes for ongoing spatial monitoring of accident patterns, including annual updates of KDE and GWR models, integration of accident predictions with real-time traffic monitoring systems and development of integrated data systems combining accident data with traffic volume and road geometry information.

Methodological extensions

Future research should consider incorporating additional predictor variables, temporal analysis of accident patterns, comparative evaluation of alternative spatial methods including Regression Kriging [19], Universal Kriging [5] and Co-Kriging [17] and application of the integrated framework to other Nigerian cities.

Policy integration

Urban planning departments should integrate spatial risk models into long-term development decisions, using zoning regulations to discourage high-risk land uses in problematic areas and conducting public awareness campaigns targeting road users in high-risk areas.

Acknowledgments: The authors acknowledge the support of the Department of Mathematics, Federal University Birnin Kebbi, for facilitating data collection. We also thank the Federal Road Safety Corps, Jega Division, for providing traffic situation reports that complemented field observations.

Conflict of interest

The author declares no conflict of interest.

References

1. World Health Organization (2021) Global status report on road safety 2021. WHO.
2. World Health Organization (2023) Global status report on road safety 2023. WHO.
3. Peden M, Scurfield R, Mohan D (2004) World report on road traffic injury prevention. World Health Organization.
4. Eke CO, Omole DN, Ayo O (2021) Trends and spatial patterns of road traffic accidents in Nigeria. Journal of Transport Geography 92: 103017.
5. Abubakar M, Umar M (2022) Spatial analysis on road traffic accidents in Kebbi State using Universal Kriging. Savanna Journal of Basic and Applied Sciences 4(1): 65-70.
6. Odeleye AO (2003) Road traffic accidents in Nigeria: A public health problem. Nigerian Medical Practitioner 43(3): 45-49. [GoogleScholar]
7. Anderson TK (2009) Kernel density estimation and K-means clustering to profile road accident hotspots. Accident Analysis & Prevention 41(3): 359-364. [Crossref] [GoogleScholar]
8. Anselin L (1988) Spatial econometrics: Methods and models. 1st edn. Springer. [Crossref] [GoogleScholar]
9. 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. [Crossref] [GoogleScholar]
10. Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press. [Crossref] [GoogleScholar]
11. Cressie N (1993) Statistics for spatial data. Wiley. [Crossref] [GoogleScholar]
12. Brunsdon C, Fotheringham AS, Charlton ME (1996) Geographically weighted regression: A method for exploring spatial nonstationarity. Geographical Analysis 28(4): 281-298. [Crossref] [GoogleScholar]
13. Fotheringham AS, Brunsdon C, Charlton M (2002) Geographically weighted regression: The analysis of spatially varying relationships. American Journal of Agricultural Economics 86(2): 554-556. [Crossref] [GoogleScholar]
14. Tobler WR (1970) A computer movie simulating urban growth in the Detroit region. Economic Geography 46(2): 234-240. [Crossref] [GoogleScholar]
15. Oni SI (2011) Spatial analysis of road traffic accidents in Lagos State, Nigeria. Journal of Geography and Regional Planning 4(7): 436-444.
16. Olawole MO (2012) Transport poverty in metropolitan Lagos. Transport Policy 24: 152-159.
17. Usman U, Abubakar M (2020) Spatial modelling of lead (Pb) concentration for the soil in Sokoto rima basin using co-kriging. International Journal of Statistical Distributions and Applications 6(2): 36-41. [Crossref] [GoogleScholar]
18. Onyeka IN, Mbachu C, Udigwe I (2018) Spatial distribution of malaria incidence in Nigeria: A geostatistical approach. Malaria Journal 17: 432.
19. Abubakar M, Salmanu A (2025) Mapping high-risk traffic zones in Jega, Nigeria: An integrated geospatial framework for road safety planning. Journal of Basics and Applied Sciences Research 3(6): 252-261. [Crossref] [GoogleScholar]
20. Abubakar M, Salmanu A, Umar M, Usman AG (2025) Analyzing spatial heterogeneity in road traffic accidents using geographically weighted regression: A case study of Jega, Nigeria. International Journal of Applied Sciences and Mathematical Techniques 11(9): 126-138.
21. Xu P, Huang H (2015) Modeling crash spatial heterogeneity: Random parameter versus geographically weighting. Accident Analysis & Prevention 75: 16-25. [Crossref] [GoogleScholar]
22. Haining R (2003) Spatial data analysis. Cambridge University Press. [Crossref] [GoogleScholar]
23. Moran PAP (1950) Notes on continuous stochastic phenomena. Biometrika 37(1/2): 17-23. [Crossref] [GoogleScholar]
24. Silverman BW (1986) Density estimation for statistics and data analysis. Chapman & Hall. [GoogleScholar]
25. Bailey TC, Gatrell AC (1995) Interactive spatial data analysis. Longman. [GoogleScholar]
26. Xie Z, Yan J (2008) Kernel density estimation of traffic accidents in a network space. Computers, Environment and Urban Systems 32: 396-406. [Crossref] [GoogleScholar]
27. Chainey S, Tompson L, Uhlig S (2008) The utility of hotspot mapping for predicting spatial patterns of crime. Security Journal 21: 4-28. [Crossref] [GoogleScholar]
28. Hurvich CM, Simonoff JS, Tsai CL (1998) Smoothing parameter selection in nonparametric regression using an improved Akaike

https://doi.org/10.70844/ijas.2026.3.49

Innovative Journal of Applied Science

information criterion. Journal of the Royal Statistical Society: Series B 60(2): 271-293. [Crossref] [GoogleScholar]
29. Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press.
30. Webster R, Oliver MA (2001) Geostatistics for environmental scientists. Wiley. [GoogleScholar]
31. Wheeler DC, Tiefelsdorf M (2005) Multicollinearity and correlation among local regression coefficients in geographically weighted regression. Journal of Geographical Systems 7(2): 161-187. [Crossref] [GoogleScholar]