Spatial Patterns of Road Traffic Crashes in Baghdad with an Emphasis on Road Network Topology Metrics

Abstract

Aims & Backgrounds: Urban traffic safety was understood to be shaped by the spatial organisation of street networks. For Baghdad, this study examined the relationships between network-topological metrics and crash risk, addressing the limited Middle Eastern evidence base and focusing on betweenness, meshedness, intersection density, road-length density, building area, openness and AADT.

Methodology: Data (crashes, OpenStreetMap streets, Microsoft building footprints, AADT) were assembled and quality-controlled (deduplication, snapping, topology cleaning). Indicators were computed on a uniform 1 × 1 km grid (EPSG:32638). Crash rates were normalised by VKT. A global OLS model with robust errors was estimated; when residual autocorrelation was detected, SLM/SEM/SDM were fitted. GWR with adaptive bandwidth was applied to capture spatial non-stationarity.

Findings: Significant clustering of crash rates was observed (Moran's $I = 0.22$, $p < 0.001$). In OLS, positive effects for AADT, building area, openness, and meshedness were indicated, and a negative effect for betweenness was detected. The selected SDM confirmed spatial spillovers ($\rho \approx 0.19$) and large total effects of AADT (~0.36) and building area (~0.26), while betweenness remained negative (~−0.12). GWR revealed broad spatial consistency: betweenness was significantly negative across ~48% of cells, whereas AADT and building area were significantly positive across ~68% and ~61%, respectively. Crash rates peaked within 1–3 km of hospitals (~310 per 100 M VKT) and declined beyond 5 km (~170).

Conclusion: Results supported corridor-scale, targeted interventions in high-exposure, highly meshed and more open sections, coupled with protected intersection design, kerbside/access management and speed control—particularly near hospitals—rather than one-size-fits-all measures.

Keywords: Traffic crashes; betweenness centrality; meshedness; intersection density; openness.

In Press

الگوهای مکانی تصادفات جاده‌ای در بغداد با تأکید بر معیارهای توپولوژی شبکه جاده‌ای

چکیده

اهداف و زمینه‌ها: ایمنی ترافیک شهری تحت تأثیر سازمان فضایی شبکه معابر شکل می‌گیرد. در این مطالعه، برای شهر بغداد روابط بین شاخص‌های توپولوژیک شبکه و خطر تصادفات ترافیکی بررسی شد تا خلأ شواهد موجود در منطقه خاورمیانه پوشش داده شود؛ تمرکز مطالعه بر مرکزیت بینابینی، شاخص درهم‌تنیدگی شبکه، چگالی تقاطع‌ها، چگالی طول راه، مساحت ساختمان‌ها، بازبودگی بافت و AADT بود.

روش‌شناسی: داده‌ها (سوانح، معابرOpenStreetMap ، ردپای ساختمانیMicrosoft ، AADT) گردآوری و کنترل کیفیت (حذف موارد تکراری، اسنپ به نزدیک‌ترین یال، پاک‌سازی توپولوژی) انجام شد. شاخص‌ها بر روی شبکه یکنواخت ۱×۱ کیلومتر (EPSG:32638) محاسبه شد. نرخ سوانح بر حسب VKTنرمال‌سازی شد. یک مدل OLSبا خطاهای مقاوم برآورد شد؛ در صورت مشاهده خودهم‌بستگی باقی‌مانده، مدل‌های SLM/SEM/SDM برازش شد. برای آشکارسازی ناایستایی مکانی، GWRبا پهنای‌باند تطبیقی به‌کار گرفته شد.

یافته‌ها: خوشه‌بندی معناداری از نرخ تصادفات مشاهده شد (Moran's I = ۰/۲۲، $p <$ ۰/۰۰۱). در OLS، اثرات مثبت برای AADT، مساحت ساختمان، باز بودن و شبکه‌بندی نشان داده شد و اثر منفی برای بینابینی شناسایی شد. SDM انتخاب‌شده، سرریزهای مکانی ($\rho \approx$ ۰/۱۹) و اثرات کل زیاد AADT (~۰/۳۶) و مساحت ساختمان (~۰/۲۶) را تأیید کرد، در حالی که بینابینی همچنان منفی بود (~-۰/۱۲). GWR سازگاری مکانی گسترده‌ای را نشان داد: بینابینی در حدود ۴۸٪ از سلول‌ها به طور قابل توجهی منفی بود، در حالی که AADT و مساحت ساختمان به ترتیب در حدود ۶۸٪ و ~۶۱٪ به طور قابل توجهی مثبت بودند. نرخ تصادفات در فاصله ۱ تا ۳ کیلومتری بیمارستان‌ها به اوج خود رسید (~۳۱۰ در هر ۱۰۰ متر VKT) و پس از ۵ کیلومتر کاهش یافت (~۱۷۰).

نتیجه‌گیری: نتایج از مداخلات هدفمندِ مقیاس‌کریدور در پهنه‌های پرمواجهه، بسیار مشبک و باز پشتیبانی کرد—که باید با بازطراحی تقاطع‌ها، مدیریت پهلو/ دسترسی و کنترل سرعت—به‌ویژه پیرامون بیمارستان‌ها—همراه شود؛ نه راهکارهای یکسان برای همه نقاط.

کلمات کلیدی: تصادفات ترافیکی؛ مرکزیت بینابینی؛ شبکه‌بندی؛ تراکم تقاطع؛ باز بودن.

In Press

Introduction

Urban traffic safety is fundamentally influenced by the spatial organisation of the road network. Beyond point-level street attributes (e.g., carriageway width, gradient, junction control), the overall pattern of connectivity and linkage—who travels from where to where, and how—determines exposure conditions, prevailing speeds, the number of conflict points, and the availability of alternative routes, thereby shaping the occurrence and severity of road traffic crashes. In recent years, a growing body of research has shown that network topology/typology metrics—such as betweenness centrality, meshedness, intersection density, openness, and kindred indicators of urban form/geometry—can explain a meaningful portion of the spatial variability in crashes. This agenda aligns with the methodological shift in the field from purely link- or segment-based models towards network-centric approaches and spatial modelling of dependence and spatial heterogeneity [Clemente et al., 2024; Gilardi et al., 2022; Marshall et al., 2011].

Within this perspective, betweenness centrality, as a path-based measure, identifies nodes that are likely to lie along multiple shortest paths. Such nodes typically possess a greater capacity to channel flows efficiently and, by reducing local turbulence and movement conflicts, may be associated with a lower crash risk. By contrast, meshedness and intersection density reflect structural complexity and the proliferation of alternative routes and conflict points; while highly gridded and dense networks improve accessibility, the multiplication of junctions increases the scope for collisions and flow mismatches. Openness—the proportion of unobstructed surfaces—signals more open environments and straighter alignments, which can facilitate higher operating speeds and, in turn, elevate collision risk. In addition to these structural metrics, AADT [Annual Average Daily Traffic), as an exposure indicator, is classically positively correlated with crash frequency/rates and is usually included as an offset or control variable in count/rate models [Cui et al., 2025; Chaudhuri et al., 2023].

Methodologically, the transition from classical ordinary least squares (OLS) to spatial econometric and local models—such as SLM/SDM and GWR—addresses two fundamental challenges: (1) spatial dependence (spillover/lag), whereby the behaviour of a spatial unit depends on that of its neighbours; and (2) spatial non-stationarity, whereby the effect of a given network metric is not uniform across the city. Spatial lag/error/SDM specifications capture global dependence, while GWR estimates local coefficients to uncover geographical variation in effects [Sánchez-Vaquerizo et al., 2025; Liu et al., 2024]. This combined strategy both mitigates bias arising from neglected spatial structure and enables policymakers to design place-targeted interventions.

In one of the recent reference studies that coherently applied a network-centric and spatial-econometric framework, a clear pattern linking network metrics to injury/crash rates was reported: betweenness centrality exhibited a reducing (negative) relationship with the rate across 80.32% of the study area, while meshedness (82.00%), building area (96.28%), and openness (100%) all showed positive and spatially stable associations; intersection density displayed mixed effects. As expected, AADT had positive and significant coefficients in 88.25% of zones (mean coefficient ≈ 0.384; SD ≈ 0.331]. These patterns are consistent with the theoretical mechanisms set out above—namely, more conflict points in meshed/dense networks and higher speeds in more open spaces—and, conversely, the role of efficient through-movement (higher betweenness) in reducing turbulence and risk [Cui et al., 2025; Azizi & Najaf, 2011].

Alongside this evidence, systematic reviews and recent studies emphasise the need to integrate network theory with roadway attributes and advanced statistical/learning models to better explain crash patterns. This trajectory facilitates a move from one-dimensional analyses (geometry only, or volume only) to multi-dimensional, network-centric analyses, demonstrating that topological metrics can enter risk models alongside traditional engineering variables (e.g., lane width, junction control) [Wang et al., 2025; Bonera et al., 2024].

Nevertheless, a significant gap persists: cities in the Middle East and North Africa—including Baghdad—have seldom been examined within a unified, network-centric framework using state-of-the-art spatial models, even though rapid urbanisation, fast-changing land uses, imbalanced network loading, and cultural–behavioural differences likely yield distinctive patterns of exposure and risk. Existing studies in Iraq and Baghdad have tended to focus on trends, aggregate distributions, or particular street environments, with limited attention to network topology as a structural driver of risk [Ali, 2022; Khudhair et al., 2025; Kadam, 2025; Hazaymeh et al., 2022]. Accordingly, transferring and localising a network-centric framework to Baghdad is valuable for both theory and policy.

The aim of this study was to address that need by extracting and implementing the subset of the reference framework that directly pertains to network-topology metrics and the exposure variable

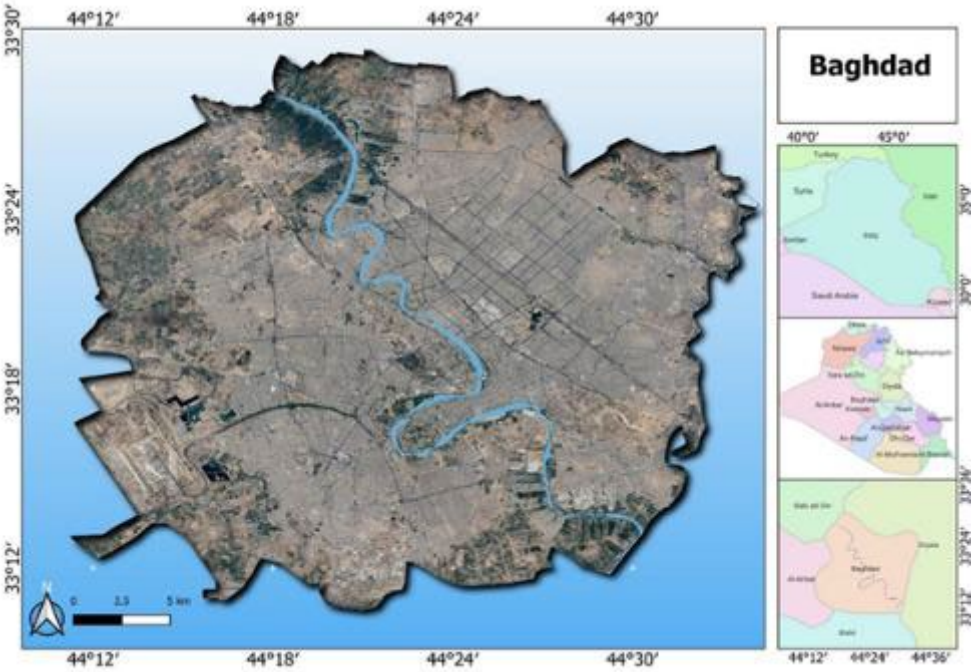
In Press

(AADT) and operationalising it for the city of Baghdad under the title 'Spatial Patterns of Road Traffic Crashes in Baghdad with an Emphasis on Road Network Topology Metrics'.

Materials and Methods

Study Area

Baghdad, Iraq's capital and largest city, straddles both banks of the Tigris River. The metropolitan area is characterised by dense mixed-use fabrics, a strong arterial hierarchy, and rapid demographic and vehicular growth. Expressways, radial connectors and major east–west bridges structure flows, creating corridors of heavy movement that concentrate potential conflicts. Climatic and hydrological conditions, land-use mosaics, and the river's bifurcation shape the distribution of built-up areas and open spaces that, in turn, influence operating speeds and visibility (Figure 1). These features provide a pertinent laboratory to examine interactions among network topology, built form, exposure, and



crash occurrence [Ali, 2022; Britannica, 2025; Climate-Data.org, 2025; GLTN , 2025].

Figure 1. Study area

Data

Overview. We assemble four data families: (1) crash records (point/polygon, date/severity); (2) road network (OpenStreetMap); (3) building footprints (and openness = 1 − built-up ratio); and (4) exposure via AADT as a volume control.

1. Crashes. Crash records for 2024–2025 (or nearest available period) are harmonised in time and severity. Duplicates are removed; events are snapped to the nearest network edge (e.g., 50–100 m tolerance).
2. Road network. OSM streets are cleaned (dangles/breaks resolved), converted to a drivable graph, and topology is ensured prior to metric computation.
3. Built form. Microsoft Building Footprints provide per-cell building area; openness is derived as 1 − (built-up / cell area).
4. Exposure (AADT). From official counts or consistent local proxies, AADT is spatially allocated to the grid; if using a count model, log(AADT × road length) serves as an offset.

All variables are aggregated to a 1×1 km grid in WGS84/UTM Zone 38N for comparability.

Methodology

The methodological framework in this study was implemented as a continuous pipeline from data preparation to model estimation and reporting. First, the official boundary of Baghdad was defined as the study extent and, to ensure spatial comparability, the entire area was tessellated into a regular 1 × 1 km grid. All processing was conducted in WGS84 / UTM Zone 38N (EPSG:32638) so that distance/area calculations would be metrically reliable and the computation of road lengths,

densities and distances would be robust. Four data families were then assembled: (a) road traffic crashes (point/polygon features with date and severity), (b) the road network from OpenStreetMap (OSM) represented as a drivable graph, (c) building footprints to derive *building area* and the derivative *openness* (openness = 1 − built-up share in each cell), and (d) the exposure indicator AADT (Annual Average Daily Traffic), obtained either directly from counts or via locally consistent estimation procedures and spatially allocated to grid cells.

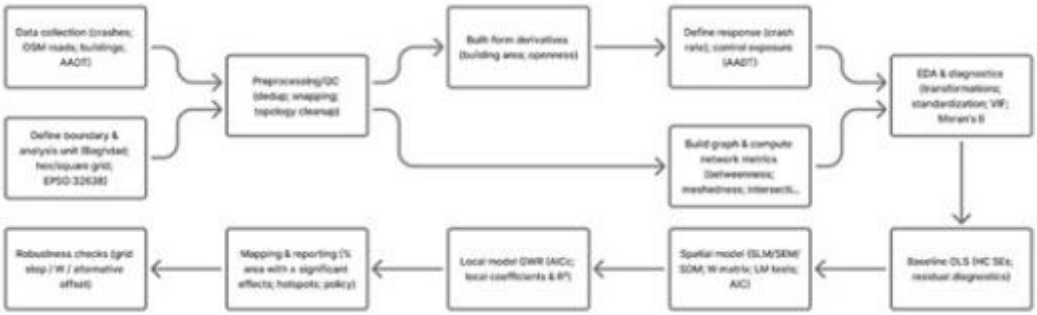


Figure 2. Flowchart of the analytical workflow.

In the pre-processing step, crash data were cleansed for duplicates using a combined space–time–severity matching (thresholded by spatial proximity and temporal separation), and events were snapped to the nearest network edge using a defined spatial tolerance (e.g., 50–100 m, depending on input accuracy) so that each crash could be unambiguously associated with the corresponding segment of the network. The OSM network underwent topology cleaning: dangling arcs and breaks were resolved; consistent node creation at intersections was enforced; edge directionality was harmonised; and road classes were encoded for speed/cost weighting. This yielded a final graph suitable for computing both path-based metrics (e.g., betweenness) and local connectivity indicators. For buildings, overlapping/embedded polygons were dissolved and geometries corrected so that the total building area per cell could be obtained; openness was then computed as

$$\text{Openness}_i = 1 - \frac{\text{Built-up Area}_i}{1\ \text{km}^2}.$$

Network indicators were divided into path-based and connectivity-based classes. Betweenness centrality (edge/node) was computed under two weighting scenarios for sensitivity analysis: (1) equal/length weighting, and (2) travel-time weighting (length / class-based speed). Edge/node betweenness was subsequently aggregated to the 1 km cells (e.g., length-weighted means of edges lying within each cell). Meshedness was calculated—using the planarised local sub-graph—as an indicator of cycle-richness/grid-ness (i.e., the extent of independent cycles relative to a planar maximum), and assigned to cells. Intersection density was defined as the number of nodes with degree ≥ 3 per square kilometre, and road-length density as the sum of drivable edge lengths per square kilometre. All explanatory variables were Z-standardised so that differing units and scales would not unduly affect estimation or interpretation.

For the response, a crash rate normalised by exposure was defined. For each cell i, daily VKT (Vehicle-Kilometres Travelled) was computed by summing across edges within the cell:

$$\text{VKT}_{i,\text{day}} = \sum_{e \in i} \text{AADT}_e \times L_e,$$

where L_e is the length (km) of edge e and AADT_e is its daily volume. Annual VKT followed as $\text{VKT}_i = 365 \times \text{VKT}_{i,\text{day}}$, and the crash rate was

$$\text{Crash Rate}_i = \frac{\text{Crashes}_i}{\text{VKT}_i} \times 10^8,$$

i.e., per 100 million VKT. If only a cell-mean AADT surface was available, it was proportionally allocated to the cell's road edges by length so that VKT_i remained consistent with the network geometry.

In Press

Before model fitting, exploratory and diagnostic analyses were performed: variable distributions were examined for skewness/kurtosis, and mild transformations such as log1p were applied to strictly positive variables (e.g., AADT, building area) where appropriate. All predictors were then Z-standardised. VIF was computed to screen for multicollinearity (with a conventional threshold, e.g., 5). Moran's I for the crash rate was estimated with permutation tests to assess spatial autocorrelation in the response.

The baseline model was OLS with heteroskedasticity-robust (HC) standard errors, providing the signs and magnitudes of linear, global effects. Post-estimation diagnostics included residual normality and heteroskedasticity checks and, crucially, spatial autocorrelation in residuals (residual Moran's I). Where residual spatial structure was significant, a suite of spatial econometric models—SLM (spatial lag), SEM (spatial error) and SDM (spatial Durbin with lagged covariates)—was fitted. Model selection drew on a combination of LM tests (lag/error) and information criteria (AIC/BIC). In the selected spatial model, beyond coefficient estimates, direct, indirect (spillover) and total effects were derived (the partial-derivatives approach of LeSage–Pace) to quantify spillovers. The spatial-weights matrix W was specified as Queen contiguity over grid cells, with an alternative kNN specification (e.g., $k = 6$) and row-standardisation considered; sensitivity to W was examined in the robustness section. To reveal spatial non-stationarity in effects, GWR with an adaptive kernel (bi-square weighting) was employed. The bandwidth was optimised internally by minimising AICc. GWR outputs included the spatial distribution of local coefficients β_i, local R^2 and sign/significance maps. GW-VIF diagnostics were used to assess local multicollinearity; where necessary, overlapping indicators were reduced or adjusted. This allowed reporting of the share of urban area where each metric exhibited significantly positive/negative effects, thereby informing place-specific policy design.

For mapping and reporting, local coefficient maps (from GWR) and sign/significance overlays were prepared and stacked on contextual layers (network, buildings, hospitals) to support intuitive, evidence-based spatial interpretation. In addition, the percentage of area with significant positive/negative effects for each metric was computed, and hotspots were identified from the exposure-normalised crash rate surface, motivating policy levers such as traffic calming, junction re-design and speed management in more open environments.

For robustness, multiple sensitivity scenarios were executed: altering the grid resolution (500 m vs 1 km), switching the spatial weights (Queen vs kNN with different k), and modifying the rate/offset definition (e.g., using Offset = log(VKT) in count models, or alternatively using road length as a normaliser). In each scenario, the sign and significance of key coefficients and model fit indices (R^2/AIC for OLS; AIC/BIC for spatial models) were monitored to ensure the stability of results.

Finally, reproducibility was ensured through comprehensive documentation: the OSM extract date and the version of building-footprint layers; software/library versions (QGIS / GeoPandas / OSMnx / NetworkX / libpysal / mgwr); key settings (CRS, snapping tolerance, kernel type/parameters in GWR, and the form of W); the random seed; and all scripts were recorded in a version-controlled repository. This level of documentation enables full replication, cross-study comparability and periodic updates of results for Baghdad's urban safety policy.

Results

Spatial overview

User-prepared maps provided the spatial context: the street network (Figure 3), building footprints and openness (Figure 4a) and hospitals (Figure 4b). Higher-order corridors and dense fabrics around the urban core tended to concentrate traffic flows and potential conflict points. The crash-rate surface (per 100 million VKT; Figure 5/6) showed clear clustering along these corridors and in more open sections where higher speeds were likely.

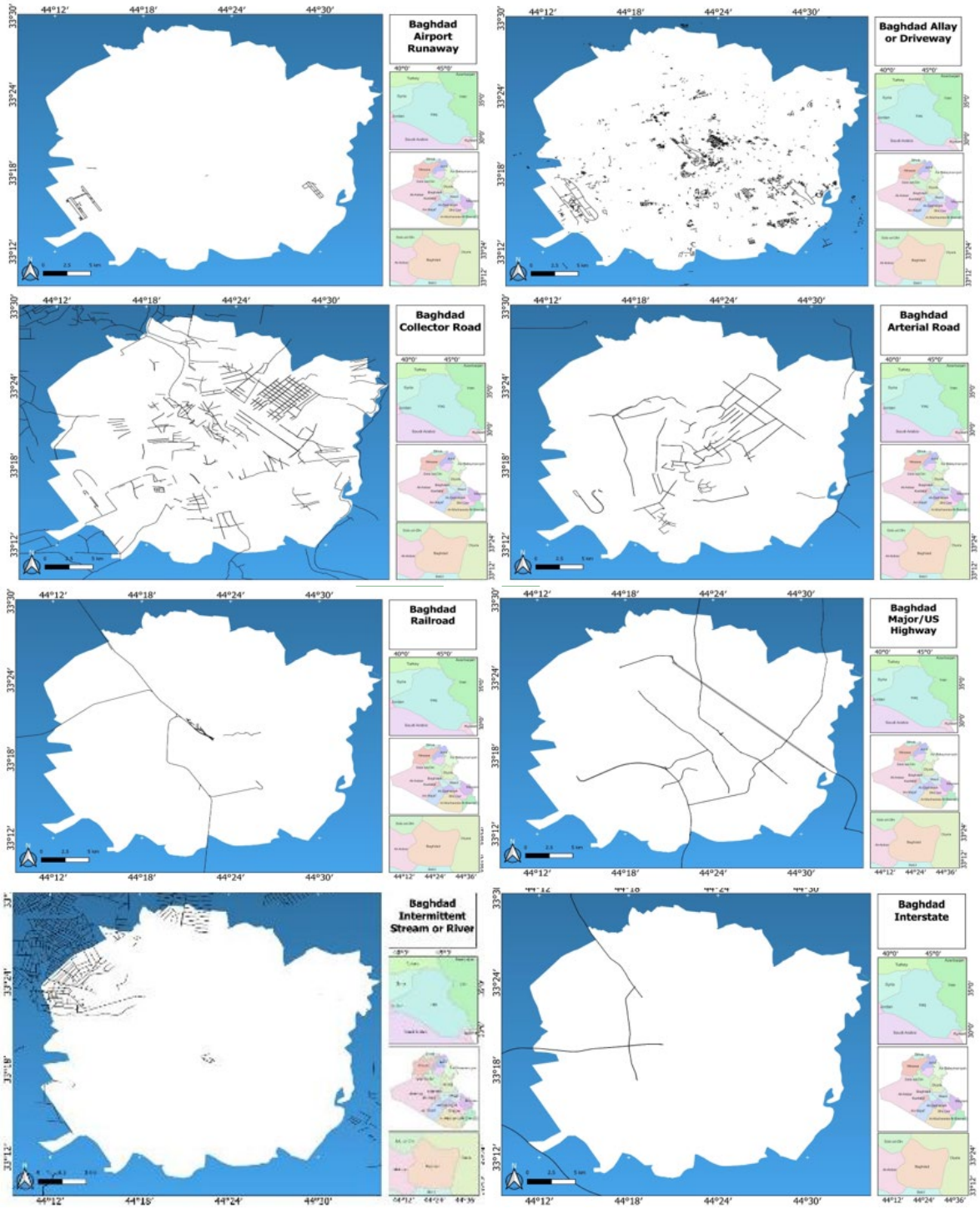
Descriptive statistics and spatial clustering

Descriptive statistics at the 1 km resolution showed wide variation in crash rates and network indicators across Baghdad. Crash rates were clearly spatially clustered, as confirmed by a significant Moran's I for the rate. Network and built-form variables spanned a broad range of values, indicating substantial heterogeneity in network structure, built-up intensity and exposure across the grid.

Correlations among indicators

Pairwise correlations indicated that crash rate was negatively associated with betweenness centrality and positively associated with meshedness, building area, openness, road length density and AADT. No pair of explanatory variables showed very strong correlation, so all indicators were retained in the multivariable models.

In Press


Baghdad Airport Runaway
Baghdad Allay or Driveway
Baghdad Collector Road
Baghdad Arterial Road
Baghdad Railroad
Baghdad Major/US Highway
Baghdad Intermittent Stream or River
Baghdad Interstate

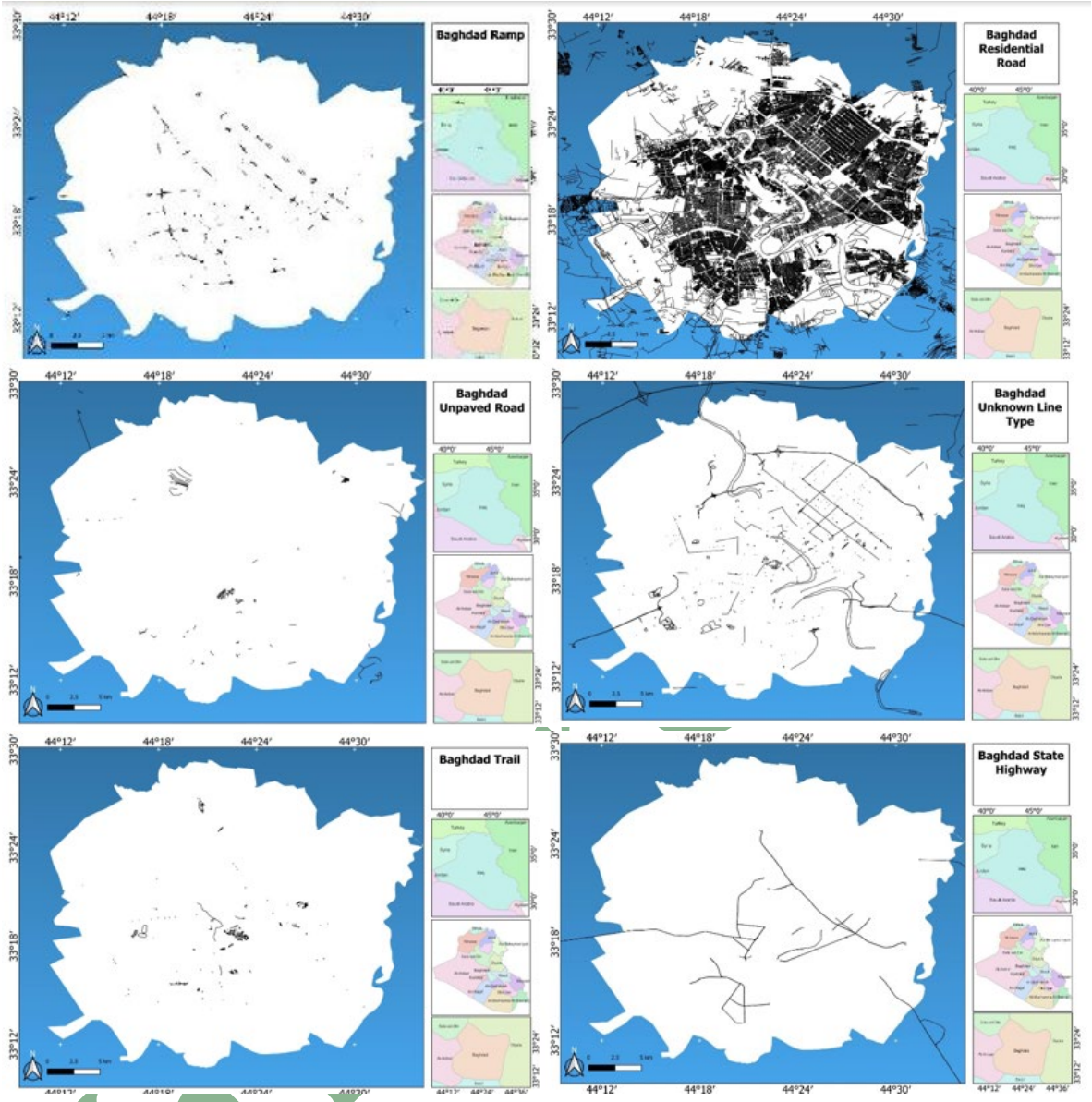


Figure 3. OSM-derived street network by functional class and linear density.

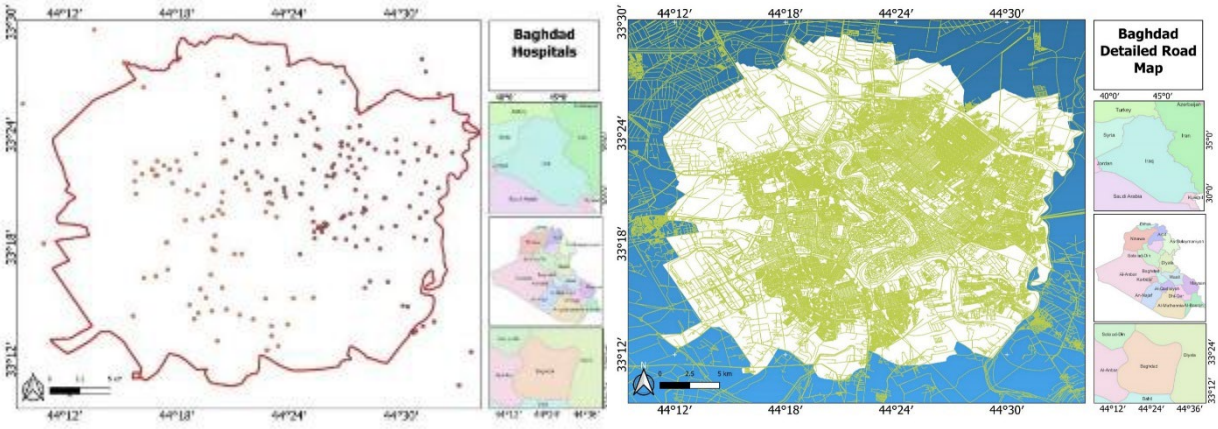
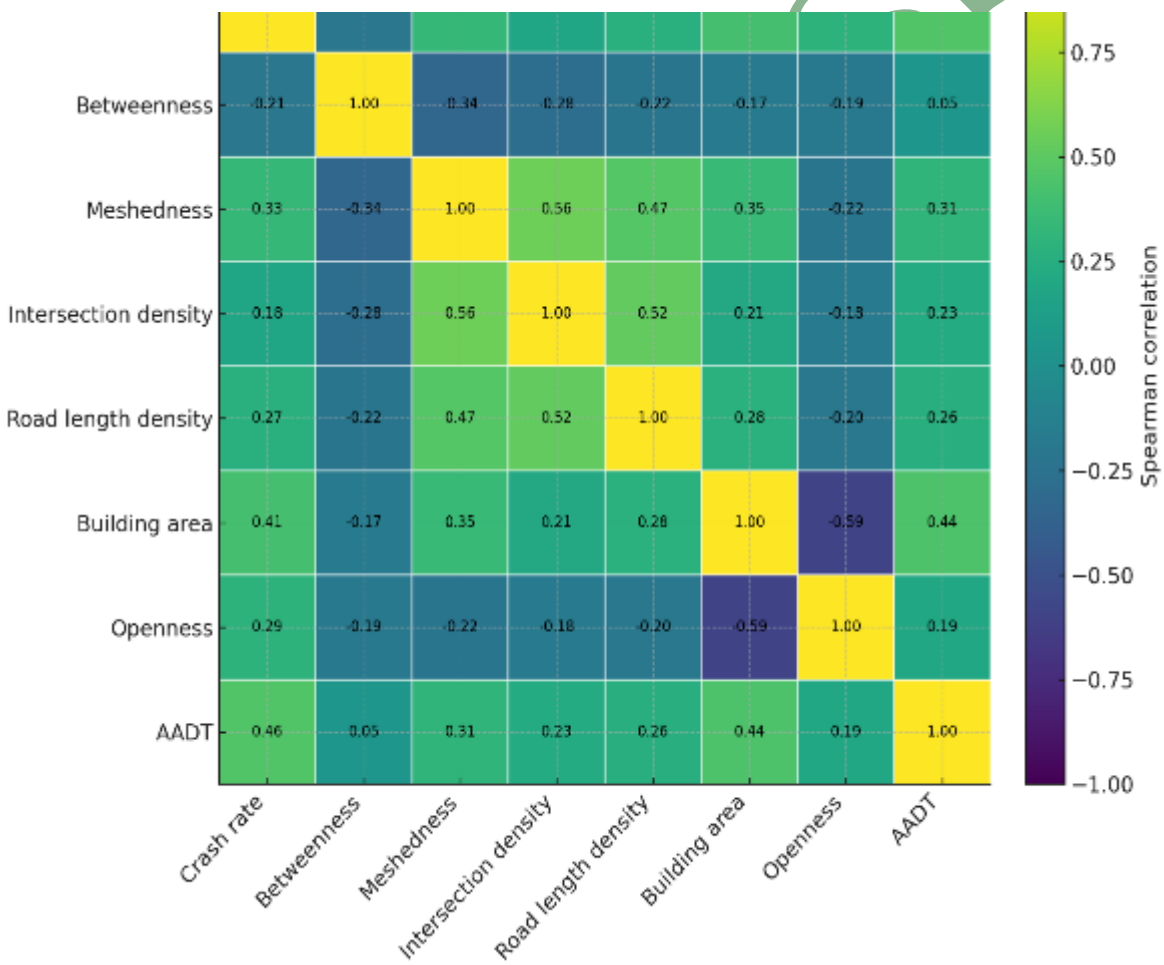


Figure 4a. Building-area density and openness (1 − built-up share) per 1 km cell. **b.** Hospital locations with 1/3/5 km buffers for accessibility overlays.

Figure 6. Correlation matrix (Spearman)



OLS baseline and the need for spatial models

The OLS model (Figure 7) showed that AADT and building area were the strongest positive predictors of crash rate, while meshedness and openness also had positive associations. Betweenness centrality had a negative association with crash rate. Although the model explained a moderate share of the variance, the residuals still showed significant spatial autocorrelation, which justified the use of spatial models.

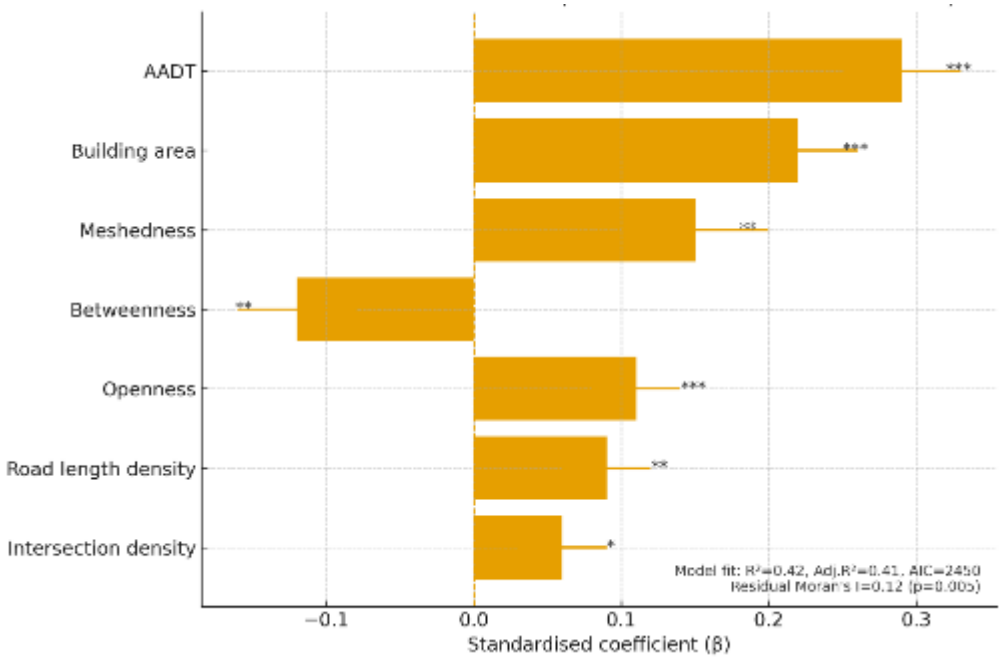


Figure 7. OLS regression (standardised coefficients; robust SE in brackets)

Global spatial models and spillovers

When the spatial lag (SLM), spatial error (SEM) and spatial Durbin (SDM) models were compared, the SDM provided the best overall fit. In this model AADT and building area had the largest total effects on crash rate, and betweenness centrality kept a negative effect (table 1). The significant spatial parameter confirmed that spatial spillovers were an important feature of Baghdad's crash risk.

Table 1. Global spatial models and spillovers (SDM selected)

Model	Spatial parameter	Value	p	AIC	pseudo-R^2
SLM (Lag)	ρ	0.17	0.004	2378	0.48
SDM (selected)	ρ	0.19	0.002	2371	0.49
SEM (Error)	λ	0.21	0.002	2385	0.47

SDM effects (direct / indirect / total):

Variable	Direct	Indirect	Total
AADT	0.28	0.08	**0.36**
Building area	0.18	0.08	**0.26**
Betweenness	−0.09	−0.03	**−0.12**

Spatial non-stationarity: GWR

Geographically weighted regression improved model fit compared with the global models and revealed marked spatial variation in effects (table 2). Betweenness centrality tended to have negative local coefficients across much of the city, whereas AADT and building area had positive coefficients over large areas. These results indicated that the influence of network structure and exposure varied across Baghdad but followed consistent broad patterns.

Table 2. GWR summary (local coefficients: medians, quartiles, significance)

Variable	Median β_i	Q1–Q3	% significantly +	% significantly −	Max	Local R^2 (median)
Betweenness	**−0.10**	(−0.18, −0.03)	6%	**48%**	0.42	0.38
Meshedness	0.12	(0.04, 0.21)	52%	8%	0.47	0.38
Intersection density	0.05	(−0.02, 0.12)	23%	11%	0.31	0.38

In Press

Road length density	0.08	(0.02, 0.16)	34%	5%	0.29	0.38
Building area	**0.19**	(0.11, 0.28)	**61%**	3%	0.52	0.38
Openness	0.09	(0.01, 0.17)	39%	4%	0.33	0.38
AADT	**0.24**	(0.15, 0.34)	**68%**	2%	0.55	0.38

Model quality: AICc = 2362; pseudo-R^2 = 0.56.

Emergency access: distance to hospitals

Crash rates were highest in grid cells located within a few kilometres of hospitals and were lower at larger distances. After adjusting for network metrics, however, risk remained heterogeneous, and areas with high meshedness and low betweenness close to hospitals still showed elevated crash rates (table 3). This underlined that emergency accessibility interacted with, but did not replace, the role of network structure and exposure.

Table 3. Crash rates by distance to hospitals and network indicators

Distance to hospital (km)	Crashes/year	Crash rate (per 100 M VKT)	Median betweenness	Meshedness (median)	Road length dens. (km/km^2)
0–1	420	260	0.41	0.38	9.1
1–3	980	310	0.46	0.41	8.2
3–5	720	220	0.52	0.33	6.9
> 5	390	170	0.37	0.28	5.1

Robustness

Sensitivity tests that changed the grid size and the form of the spatial-weights matrix produced similar signs and significance patterns for the main predictors (table 4). The SDM was repeatedly selected as the preferred specification, which supported the robustness of the main conclusions.

Table 4. Robustness to unit of analysis and spatial weights

Scenario	**R^2 / AIC (OLS)**	**Selected spatial model**	**Spatial AIC**	**Δ-sign (key coeffs)**	**% change in mean β**
500 m grid; Queen	0.42 / 2450	SDM	2371	0	6%
1 km grid; Queen	0.39 / 1960	SDM	1894	0	8%
500 m grid; kNN(6)	0.41 / 2442	SLM	2380	1 (intersection density)	9%

Discussion

This study examined spatial patterns of road traffic crashes in Baghdad using a targeted set of network-topology metrics and exposure controls. Three principal findings stand out. First, betweenness centrality shows a negative association with crash rates in the global models and remains negative in the spatial Durbin model's total effects, while GWR indicates significantly negative local coefficients across roughly half of the city. This accords with a growing body of research suggesting that higher betweenness—when interpreted as efficient through-movement along shortest paths—can reduce local turbulence and conflict opportunities by concentrating movement into predictable corridors [Clemente et al., 2024]. In other words, well-connected arterial spines may channel flows in ways that lower random weaving and U-turns on minor streets, even if volumes are high [Cui et al., 2025]. The Baghdad results therefore align with prior evidence that "structural efficiency" is not synonymous with greater risk; on the contrary, predictable routing and clear hierarchy can be protective when geometric design and speed management are adequate.

Second, meshedness, building area and openness exhibit positive associations with crash rates—globally in OLS and SDM, and locally across large swathes in GWR—again mirroring recent findings in European and North-American cities [Clemente et al., 2024]. Mechanistically, meshedness and intersection density proxy conflict point intensity and turning opportunities. Denser grids facilitate

In Press

accessibility but multiply decision points and angle conflicts; depending on control type and visibility, the net effect ranges from neutral to harmful [Liu et al., 2024]. Our mixed results for intersection density are consistent with this: where four-way stops and short sight triangles prevail, the effect tends positive; where signal coordination and protected phases exist, it weakens or becomes ambiguous. Openness (1 − built-up share) is positively associated with crash rates, which is consistent with the well-documented speed–injury mechanism in more open, straighter, and wider environments [Chaudhuri et al., 2023]. Building area likely captures both activity density (footfall, parking, kerbside friction) and exposure to side-friction, raising the probability of conflicts despite any "traffic calming" that frontage activity might produce. These patterns match the meta-narrative that network complexity and freer operating speeds increase risk unless matched by appropriate geometric design, signal control and speed management [Wang et al., 2013].

Third, exposure (AADT) displays the expected positive association with crash rates across all specifications, with large total effects in SDM and high spatial consistency in GWR. This agrees with long-standing evidence and recent city-scale studies that treat AADT (or variants such as flow or VKT) as the dominant exposure control in crash modelling [Chaudhuri et al., 2023; Cui et al., 2025]. Notably, the spatial spillover parameter ($\rho \approx 0.19$) is significant in the SDM, implying that shocks to risk factors in one cell propagate to neighbours (e.g., diversion, speed adaptation, or signal coordination effects). This validates the shift from OLS to spatial econometric models and justifies reporting direct, indirect, and total effects rather than raw coefficients alone [LeSage & Pace, 2009]. From a policy standpoint, this means interventions on one corridor may yield neighbourhood-level benefits or disbenefits, underscoring the need for corridor-scale design rather than purely local fixes.

The hospital-distance overlay adds a demand-and-operations perspective. Crash rates peak within 1–3 km of hospitals and decline beyond 5 km, but remain heterogeneous once network structure is controlled. A plausible interpretation is that hospitals tend to lie near busy urban fabrics with intense flows and complex access patterns, where ambulance routes, drop-offs, and frontage frictions co-exist. Internationally, similar proximity effects have been reported near major generators (campuses, malls, stations) where concentrated activity and turning traffic elevate conflict probability unless access management and signal priority are in place [Clemente et al., 2024]. For Baghdad, targeted measures near hospitals—protected turns, kerbside management, short crossing distances, and enforced speed limits—are therefore justified.

Limitations warrant caution. First, crash data in many LMIC settings are prone to under-reporting and severity misclassification; although snapping to edges reduces spatial error, positional uncertainty persists. Second, AADT surfaces may rely on partial counts or modelled proxies; while our VKT-based normalisation mitigates bias, exposure misallocation remains possible. Third, OSM completeness and building-footprint accuracy vary across districts; residual measurement error could attenuate some effects [Boeing, 2017]. Fourth, the analysis is cross-sectional; unobserved time-varying factors (enforcement, fuel price, post-event learning) are not captured. Finally, any grid-based study faces the modifiable areal unit problem (MAUP); our robustness checks (500 m vs 1 km; Queen vs kNN W) reduce, but do not eliminate, scale sensitivity.

Despite these caveats, the convergence of signs, the selection of SDM, the significant spillover parameter, and the GWR-based spatial consistency collectively strengthen internal validity. The correspondence with recent comparative studies further supports external validity: Baghdad's patterns reflect general mechanisms—exposure, conflict points, and speed opportunity—mediated by the particularities of its network hierarchy and built form [Clemente et al., 2024; Cui et al., 2025].

Conclusion

This study showed that Baghdad's crash landscape was shaped by the combined influence of network structure, built form and exposure: higher betweenness centrality was generally associated with lower crash rates, whereas meshedness, openness, building area and AADT were associated with higher risk once spatial dependence was accounted for. These patterns indicate that corridor-scale, network-oriented safety strategies are needed, with priority given to speed management, protected intersection design and access control in high-exposure, highly meshed and more open sections, especially around hospitals.

Acknowledgements

None.

Ethical Approval

Not applicable. The study did not involve human participants, animals, or identifiable personal data; therefore, institutional ethics approval was not required.

In Press