

SPATIAL PATTERN ANALYSIS OF BUILT-UP LAND USE AND ZONING CONFORMITY USING POLYGON-BASED STATISTICS: A CASE STUDY OF TEGAL CITY, INDONESIA

Rifqiya LUTHFIYANI, Parluhutan MANURUNG, Adi WIBOWO

Geodesy and Cartography

<https://doi.org/10.3846/gac.2026.25842>

This file gives the mathematical formulations, weighting scheme and computation settings referred to in Section 3.1 of the main article.

S1. Nearest Neighbour Index (NNI)

For n centroid points in a study area of size A , the NNI compares the observed mean nearest-neighbour distance with the value expected under complete spatial randomness (Clark & Evans, 1954):

$$R = D_{\text{obs}} / D_{\text{exp}}, \quad D_{\text{exp}} = 0.5 / \sqrt{(n/A)}, \quad z = (D_{\text{obs}} - D_{\text{exp}}) / SE, \quad SE = 0.26136 / \sqrt{(n^2/A)}.$$

$R < 1$ indicates clustering, $R \approx 1$ a random pattern and $R > 1$ dispersion.

S2. Ripley's K- and L-functions

$$K(r) = (A / n^2) \sum_i \sum_{j \neq i} I(d_{ij} \leq r) / w_{ij}, \quad L(r) = \sqrt{(K(r) / \pi)},$$

where d_{ij} is the distance between centroids i and j , $I(\cdot)$ is the indicator function and w_{ij} is the isotropic edge-correction weight (the proportion of the circumference of the circle centred on i and passing through j that lies inside the study area). $L(r) - r > 0$ indicates clustering at distance r , and $L(r) - r < 0$ indicates regularity (dispersion). Significance was assessed with 999 Monte Carlo simulations of complete spatial randomness; observed values outside the 95% simulation envelope (2.5th and 97.5th percentiles) were considered significant (Diggle, 2014; Kiskowski et al., 2009). Distances were evaluated up to 2,000 m in steps of 50 m.

S3. Spatial weights and Global Moran's I

Neighbourhood relations were defined by first-order Queen contiguity: $w_{ij} > 0$ if polygons i and j share an edge or a vertex, and $w_{ij} = 0$ otherwise. The formulas below are written for general weights w_{ij} ; when the weights are row-standardised, I equals the slope of the Moran scatterplot.

$$I = (n / S_0) \cdot \sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x}) / \sum_i (x_i - \bar{x})^2, \quad S_0 = \sum_i \sum_j w_{ij},$$

where x_i is the built-up land area (ha) of polygon i and $\bar{x}$ is the mean. Under the null hypothesis of no spatial autocorrelation, $E[I] = -1/(n - 1)$. Significance was evaluated with the standardised z-score, $z = (I - E[I]) / \sqrt{\text{Var}[I]}$, at $p < 0.05$. Positive I indicates clustering of similar values and negative I indicates dispersion.

S4. Local Moran's I (LISA)

$$I_i = z_i \sum_j w_{ij} z_j, \quad z_i = (x_i - \bar{x}) / s,$$

where s is the standard deviation of x . For row-standardised weights the mean of the local statistics equals the global I (Anselin, 1995). Significance of each I_i was assessed with a conditional permutation test (pseudo p-value < 0.05). Significant polygons were classified by the quadrant of the Moran scatterplot, using z_i and the spatial lag $\sum_j w_{ij} z_j$:

High-high (HH): $z_i > 0$ and $\text{lag} > 0$; Low-low (LL): $z_i < 0$ and $\text{lag} < 0$; High-low (HL): $z_i > 0$ and $\text{lag} < 0$; Low-high (LH): $z_i < 0$ and $\text{lag} > 0$; polygons with $p \geq 0.05$ were classed as not significant (NS).

S5. Conformity indicators

Conformity with the RTRW was assessed by overlaying built-up polygons on the zoning map. Two indicators are reported in the main article:

$CI_{\text{polygon}} = N_{\text{conforming}} / N_{\text{total}} \times 100, \quad CI_{\text{area}} = A_{\text{conforming}} / A_{\text{total}} \times 100,$

where N is the number of built-up polygons and A the built-up area whose land-use category matches the designated zoning category.

Table S1. Software and computation settings

Parameter	Value
Software	ArcGIS Pro 3.2; GeoDa 1.20
Coordinate system	WGS 84 / UTM Zone 49S (EPSG:32749)
Ripley's K/L: maximum radius; step	2,000 m; 50 m
Ripley's K/L: edge correction	Isotropic
Ripley's K/L: significance envelope	999 Monte Carlo simulations under complete spatial randomness
Spatial weights (Moran's I, LISA)	First-order Queen contiguity
Significance level	95% ($p < 0.05$)
Variable (Moran's I, LISA)	Built-up land area per polygon (ha)