International Journal of Innovative Mathematics, Statistics & Energy Policies
14(3):59-67, July-Sept, 2026

© SEAHI PUBLICATIONS, 2026 www.seahipublications.org ISSN: 2467-852X

DOI:10.5281/zenodo.21531718

A Multiscale Framework For Individual Structural Attribution In Marked Spatial Point Processes: An Adaptive LISMA-Mahalanobis-Ward (ALMW) Clustering Approach.

[1]Datong, Godwin Monday; [2]Lasisi, K. E & [3]Ahmed A.

[1]Department of Mathematics
University of Jos, Jos. Jos Nigeria.
Correspondence author: Email: - mallong2007@yahoo.com.

[2,3]Department of Statistics
Abubakar Tafawa Balewa University, Bauchi, Nigeria

ABSTRACT
Traditional second-order analysis of marked spatial point patterns often relies on global estimators that mask individual structural contributions and assume stationarity of the underlying process. Unlike these global estimators, the local indicator of marked spatial association (LISMA) functions enables a point-wise decomposition of spatial-mark dependencies within a spatially localized second-order interaction framework, thereby jointly characterizing spatial and mark dependence. Multiscale feature extraction in these processes is frequently plagued by multicollinearity between adjacent spatial lags, leading to redundant information and biased clustering in Euclidean space. In this paper we introduce an integrated pipeline for individual structural attribution, known as adaptive LISMA-Mahalanobis-Ward (ALMW) framework that employs Abramson-class adaptive smoothing to stabilize local intensity estimates across heterogeneous densities. In order to resolve the redundancy problem accruing from the high multicollinearity, we implement a Mahalanobis-weighted transformation that utilizes the precision matrix (Σ^{-1}) to de-correlate the spatial lag vector before performing Ward's minimum variance grouping of the feature vectors. The framework was evaluated using simulation studies, featuring various scenarios of spatial mark association to study spatial marked point processes in $\mathbb{R}^2$ and on linear networks with continuous and categorical marks, demonstrating that the ALMW framework significantly outperform the existing standard Euclidean-Ward's algorithms. The high cophenetic reliability (c = 0.821 for planar space and c = 0.791 for the linear network grid) confirm that the hierarchical dendrogram faithfully preserves the interaction geometry of the decorrelated feature spaces.
Keywords: adaptive smoothing, multiscale framework, LISMA, precision matrix, cluster analysis techniques, structural attribution, marked spatial point processes.

1.0 INTRODUCTION
The analysis of marked spatial point processes (MPPs) is crucial to understanding complex systems, ranging from the competitive dynamics of forest ecosystems to the synergistic patterns of urban crime especially in urban cities. Usually, these processes are characterized using global second-order statistics, such as Ripley's K-function (Ripley, 1976) which is typically based on the assumption of non-marked

stationary and isotropic point processes (D'Angelo et al., 2022) or the mark correlation function (Moradi and Eckardt, 2025), which provide a powerful summary of the aggregate dependency between locations and attributes, that also operate under an assumption of structural stationarity and isotropy. By averaging interactions across the entire study window, global estimators effectively mask the heterogeneous roles that individual events play in the formation of the global pattern (Eckardt and Moradi, 2024). Recent advancements in spatial statistics have shifted the focus toward localized indicators, like LISA of Anselin (1995); LISTA (Rodriguez-Cortes et al., 2014); LIMA (Eckardt and Moradi, 2024) and most notably now, the concept of the local indicators of spatial mark association (LISMA). LISMA allows for the decomposition of global association into point-specific contributions, offering an infinitesimal view of spatial-mark interaction. However, moving from local estimation to individual structural attribution which is the process of classifying events based on their multiscale interaction signatures or profiles, is plagued by two significant methodological challenges.

Standard local estimators typically employ a fixed bandwidth, which introduces significant bias in heterogeneous density environments, so that in sparse regions, fixed kernels fail to resolve the signal from stochastic noise, while in dense clusters, they over smooth critical structural variations (Lemke et al., 2015). Secondly, when LISMA is evaluated across multiple spatial lags to capture multiscale behaviour, the resulting feature vectors exhibit extreme multicollinearity, since adjacent spatial lags are inherently redundant; wherefore in a standard Euclidean metric space, this redundancy double-counts certain scales, thereby biasing unsupervised clustering algorithms and obscuring the distinction between subtle structural regimes (Wani, 2024).

This paper addresses these limitations by introducing the adaptive LISMA-Mahalanobis-Ward (ALMW) framework. Our approach contributes to the literature in three specific ways. First, we implement Abramson-class adaptive bandwidths for both spatial and mark-intensity estimation, ensuring variance stabilization across the entire attribute spectrum. Secondly, we utilize the precision matrix ($\boldsymbol{\Sigma^{-1}}$) to de-correlate the multiscale feature space, ensuring that the structural classification is derived from unique, non-redundant interaction signals. Finally, through a validated pipeline of Gap statistic optimization and Ward's hierarchical grouping, we identify and map universal structural archetypes.

The paper is structured as follows. Section 2 provides an outline of the general principles for the analysis of marked spatial point processes in $\mathbb{R}^2$ and using the ALMW framework with their corresponding non-parametric estimators. Section 3 presents a simulation study to evaluate the performance of the ALMW framework under various scenarios. The paper ends with a discussion in section 4.

2. Marked spatial point processes in $\mathbb{R}^2$

Following Eckardt and Moradi (2023), we consider the marked point processes, $\Psi = \{(x_i;\ m_i)\}_{i=1}^{N}$ on the product space $\mathbb{R}^2 \times \mathcal{M}$, having spatial locations x_i in the two-dimensional Euclidean space $\mathbb{R}^2$, equipped with the Lebesgue measure $|A| = \int_A du$ for Borel sets $A \subseteq \mathfrak{B}(\mathbb{R}^2)$. The associated marks $m_i \in \mathcal{M}$ is Polish and equipped with a finite reference measure v on the Borel σ-algebra $\mathfrak{B}(\mathcal{M})$. The Borel σ-algebra $\mathfrak{B}(\mathbb{R}^d \times \mathcal{M}) = \mathfrak{B}(\mathbb{R}^2) \otimes \mathfrak{B}(\mathcal{M})$ is endowed with product measures $A \times E \longmapsto |A|v(E)$, for $A \times E \in \mathfrak{B}(\mathbb{R}^2 \times \mathcal{M})$. We assume that Ψ and its unmarked (or ground) process, Ψ_g, are simple, meaning that the process will have no multiple points per location, and that the expected number of points in $A \in \mathfrak{B}(\mathbb{R}^2)$ with marks in $E \in \mathfrak{B}(\mathcal{M})$ is given by

$\Lambda(A \times E) = \mathbb{E}(N(A \times E)) = \int_{AxE} \lambda(u, m)\, duv(dm)$,

where $\lambda(\cdot)$ is the intensity function of Ψ, and governs its spatial distribution.

If $\lambda(\cdot) = \lambda$ (a constant), then Ψ is a homogeneous point process otherwise it is said to be inhomogeneous. Furthermore, Ψ is stationary if its distribution is invariant under translation i.e. $\Psi = \{(x_i;\ m_i)\}_{i=1}^{N}$ with $\Psi + c = \{(x_i + c;\ m_i)\}_{i=1}^{N}$; and it is isotropic if its distribution is invariant under

rotation, meaning that $\Psi = \{(x_i;\ m_i)\}_{i=1}^{N}$ with $c\Psi = \{(cx_i;\ m_i)\}_{i=1}^{N}$, $c \in \mathbb{R}$, have the same distributions. If Ψ is both stationary and isotropic, then it is said to be motion invariant and

$$\Lambda(A \times E) = \lambda v(A)\mathcal{M}(E),$$

where λ is the intensity of Ψ_g and $v(\cdot)$ is the Lebesgue measure, while $\mathcal{M}$ is the mark distribution. For multivariate/multitype point processes, the mark space, $\mathcal{M} = \{1, 2, \ldots, k\}$, while for real-valued marks, $\mathcal{M}$ is describe by the distribution function, $F_{\mathcal{M}}(m) = \mathbb{R}^d$.

Let $\mathcal{X} = \{(x_i;\ m_i)\}_{i=1}^{n}$ be an observed finite marked point pattern within a bounded observation window $W \subset \mathbb{R}^2 \times \mathcal{M}$, which is assumed to would have been generated by the random process Ψ with x_i being the spatial location of the i^{th}-event and m_i, the associated mark, living on a suitable mark space $\mathcal{M}$.

2.2 Double-Adaptive Kernel Estimation of Joint intensity function, $\hat{\lambda}(x, m)$

To mitigate the bias inherent in fixed-bandwidth estimators, we employ adaptive kernel density estimators for the first-order intensities $\lambda(x)$ and $\lambda(m)$. Using Abramson's (1982) method, the adaptive bandwidths h_i and b_i are defined as:

$$h_i = h_0\left(\frac{\hat{\lambda}(x)}{G_s}\right)^{-1/2} \text{ and } b_i = b_0\left(\frac{\hat{\lambda}(m)}{G_m}\right)^{-1/2} \quad (1)$$

where h_i and b_i are the adaptive bandwidths for spatial and mark spaces respectively,
h_0 and b_0 are global (or pilot) bandwidth estimates,
$\hat{\lambda}(x)$ and $\hat{\lambda}(m)$ are marginal intensity function estimates in space and mark, G_s and G_m are the geometric means of the marginal intensities. These geometric means are usually used to free the bandwidth factor from the data scale (Silverman, 1986 and Davies and Hazelton, 2010). The joint intensity kernel estimator using the point-specific bandwidth is given by

$$\hat{\lambda}(x,m) = \sum_{i=1}^{n} \frac{1}{h_i^2 b_i} K_s\left(\frac{x_i - x_j}{h_i}\right) K_m\left(\frac{m_i - m_j}{b_i}\right) \quad (2)$$

where; $K_s(\cdot)$ is a 2D spatial kernel and $K_m(\cdot)$ is a 1D kernel for mark space.

2.3 Multiscale LISMA Feature Profile (v) Extraction

Using Shimanti's (2002) I test function which centres each mark by the global mean mark μ_m and normalizes the product by the mark variance σ_m^2; so that for each point i, the feature vector, v_i, at distance r, will be given by the LISMA function,

$$v_i = \sum_{j \neq i}^{n} \frac{K_s\left(\frac{\|x_i - x_j\| - r}{h_i}\right)(m_i - \mu_m)(m_j - \mu_m)}{h_i \hat{\lambda}(x_i)\hat{\lambda}(x_j)\mathfrak{e}(x_i, x_j)\sigma_m^2} \quad (3)$$

where; $\|\cdot\|$ is the Euclidean distance, while $\mathfrak{e}(x_i, x_j)$ is the edge correction factor, and $\mathfrak{e}(x_i, x_j)$ is the edge correction factor to counteract boundary bias, ensuring that points near the edge of a study area are not unfairly considered isolated merely because neighbours outside the boundary are unobserved. By evaluating v_i across a calibrated vector of d spatial lags, we construct the $(n \times d)$ feature matrix V.

2.4 Structural Whitening via Mahalanobis Transformation

The primary methodological innovation of this framework is the resolution of the lag redundancy problem, in which adjacent spatial lags are highly collinear, so the sample covariance matrix Σ of V is ill-conditioned and to ensure a scale-invariant clustering, we define the distance between two structural profiles i and j using the Mahalanobis metric:

$$D_m^2 = (v_i - v_j)^T \Sigma^{-1} (v_i - v_j) \quad (4)$$

Applying the precision matrix (Σ^{-1}), the framework whitens (or de-correlate) the feature space and neutralizes the redundant spatial autocorrelation while allowing for the discovery of subtle structural regimes that are mathematically invisible in Euclidean spaces.

2.5. Estimating Optimal number of clusters.

There are several techniques for determining the optimal number of clusters in a clustering algorithm. Some of these methods include silhouette, elbow and gap statistic. Kushal and Mohit, (2023) with references therein, postulate that the Gap statistic performs better than most of the above-mentioned techniques, hence we discuss the Gap statistic here in detail while the three techniques are applied in the simulation study. The Gap statistic is the difference between the expected log-dispersion under a null model and the observed log-dispersion:

$$Gap(k) = E_n^*\{\log(W_k)\} - \log(W_k) \quad (5)$$

where E_n^* denotes the expectation under a sample of size n from a reference null distribution, generated using Monte Carlo simulation of a uniform distribution over the box aligned with the principal components of the data.

$W_k = \sum_{r=1}^{k} \frac{1}{2n_r} \sum_{i,j \in C_r} d_{ij}^2$ is the within-cluster variation; d_{ij}^2 is the squared Mahalanobis distance between LISMA profiles i and j in cluster C_r and n_r is the number of points assigned to cluster r. The 1-standard-error selection rule highlights the selection of the optimal $\hat{k}$ as the smallest k such that:

$$Gap(k) \geq Gap(k+1) - s_{k+1} \quad (6)$$

where s_{k+1} is the standard error of the simulation, accounting for the Monte Carlo noise.

2.6 Ward clustering method and Spatial attribution.

The Ward-clustering method obtains new clusters by minizing intra-cluster variance, whereby the cluster with the lower error square value is then chosen (Saraçli and Akşit, 2022).

Ward's method seeks to minimize the increase in the sum of squared errors (SSE) at each step of the agglomeration. For a given partition of events into clusters, the objective function is defined as:

$$SSE = \sum_{i=1}^{k} \sum_{i \in I_c} \|v_i - \bar{v}_c\|_M^2 \quad (7)$$

where v_i is the i^{th} observation vector in the cluster, I_c is the set of indices for events in clusters $\bar{v}_c$ is the centroid (mean LISMA profile) of cluster c, and $\|\cdot\|_M^2$ denotes the squared Mahalanobis distance which ensures that the variance minimization occurs in the decorrelated metric space.

2.7 Numerical Validation using Cophenetic Fidelity of the ALMW Dendrogram

In hierarchical clustering, the Cophenetic Correlation Coefficient (c) is the standard goodness-of-fit metric, which proves that your dendrogram faithfully preserves the pairwise distances from the original Mahalanobis-weighted feature space (Ponde, et al. 2016). This Cophenetic Correlation Coefficient (c) is calculated as the Pearson product-moment correlation between the original dissimilarity matrix (D) and the cophenetic distances (T) derived from the dendrogram heights:

$$c = \frac{\sum_{i<j}(d_{ij} - \bar{d})(t_{ij} - \bar{t})}{\sqrt{\sum_{i<j}\left[(d_{ij} - \bar{d})^2\right]\sum_{i<j}\left[(t_{ij} - \bar{t})^2\right]}} \quad (8)$$

where d_{ij} is the original Mahalanobis distance between observations i and j, and t_{ij} is the height in the dendrogram where these two observations first merge into a single cluster.

Finally, to visualize those archetypes, spatial attribution is used in taking the branches from our dendrogram and projected back onto the spatial domain.

3 Simulation Study and Algorithmic Validation

We devote this section to evaluating the statistical performance of our proposed ALMW framework under different scenarios, comparing their performance to the existing global Euclidean-Ward's algorithm. To thoroughly validate the performance, stability, and convergence of the framework under the strictly controlled conditions, a comprehensive Monte Carlo simulation study was conducted.

To replicate the harsh spatial noise and non-uniform clustering common in empirical fields, we consider 500 realizations driven by an inhomogeneous directional spatial intensity function given by

$$\lambda(x,y) = 450x+50. \quad (9)$$

This function establishes a severe spatial asymmetry along the horizontal x-axis while maintaining a constant density profile along the vertical y -axis. This specific layout evaluates whether the localized, product-density multi-scale local indicators of spatial marked association (LISMA) can successfully dissociate genuine micro-scale point-mark interactions from a powerful, sweeping macro-scale background density gradient.

In the planar space framework, points are generated within a continuous unit square window W = [0,1] $\times$ [0,1] using conditional spatial point-pattern generation, simulating high-density event clusters against open, low-density fields; while on the linear network framework, the same directional probability curve is projected onto the realistic branching street lattice of the Chicago network $(L_Chicago)$. However, since the Chicago grid maps across thousands of physical units, the spatial intensity function is vigorously normalized between 0 and 1 relative to the network's bounding box using the gradient equation;

$$\lambda(x,m)_L = 450 \cdot \left(\frac{x - x_{min}}{x_{max} - x_{min}}\right) + 50. \quad (10)$$

In order to force exactly 500 points along these street tracks while faithfully retaining this horizontal gradient, a specialized vector-safe rejection sampling loop is executed, so that even when points are geographically forced to snake along complex, non-linear network segments, their collective density profile still mirrors the global horizontal gradient, setting up a symmetric baseline for both models.

3.1.1 Generative Statistical Regimes and Mark Modelling

Prior to any spatial filtering or hierarchical grouping, the generated point populations are partitioned into three distinct generative archetypes matching a strict 30% / 30% / 40% structural split (π). These archetypes are defined at the outset solely by their mathematical and covariance properties.

1. Regime A (Positive spatial mark association, $I > 0;$: 30%) represents a generative cohort where localized marks are highly similar and positively correlated, simulated from N(80, 5). In continuous spaces, this simulates localized cohort clumping, while in the network environment, it is mapped via a numeric dummy label of 1.
2. Regime B (Negative spatial mark association, $I < 0;$: 30%) denotes a generative cohort where localized marks exhibit a strong negative covariance or localized regular spacing, which are simulated using N (10, 2). This models micro-scale variance suppression or localized exclusion zones, mapped along network segments via a numeric dummy label of -1.
3. Regime C (Independent random labelling / Spatial noise, $I = 0;$: 40%) functions as the high-entropy background control layer. Marks in this regime are assigned completely at random from a continuous uniform distribution on [0, 100] or a numeric dummy label of 0 for the linear network space. It exhibits zero spatial correlation, serving to evaluate the robustness of the subsequent Mahalanobis-Ward linkage splits under high-noise conditions.

Multi-scale spatial feature profiles are continuously tracked across 15 equidistant distance bands (d = 15) spanning $0 < r \leq 0.15$ in the planar domain and $0 < r \leq 0.20$ along network paths using shortest-path metrics (distance = "path").

3.1.2 The Double-Adaptive Joint Intensity Filtering

A spatial adaptive bandwidth vector h_i is computed using a global cross-validated pilot selector (h_0), tracking local point crowding across the topology, and an attribute mark bandwidth vector b_i is initialized via a global continuous pilot selector (b_0), running along the categorical-to-numeric indicator variance tracks. We then compute a joint Gaussian product kernel across these two vectors to normalize the raw spatial and attribute scales into scale-invariant, dimensionless joint densities centred around a global relative baseline mean of 1.0000, using the intensity function (2).

We use the spatstat package (Baddeley et al., 2015) along with its subroutines and also used ggplot2, cluster and MASS packages (Wickham, 2016) of the R software for the simulations and graphical plots respectively. Figure 4.1 show plots of the unified 500 realizations generative input states for both planar and linear network geometrical spaces.

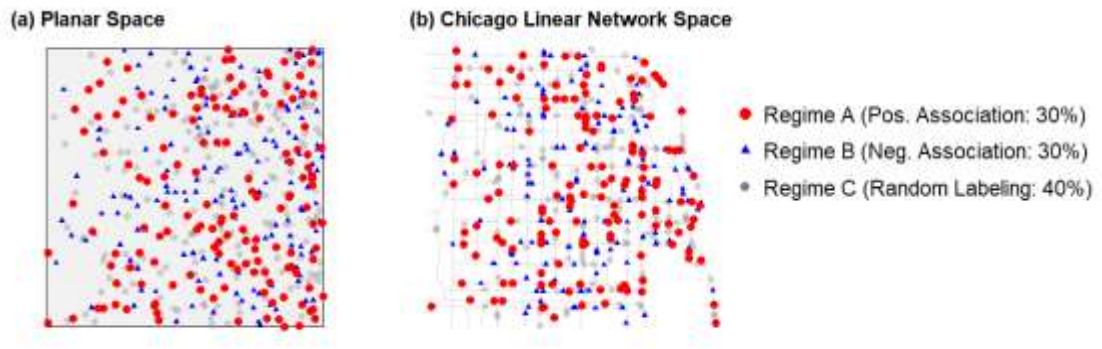

Figure 3.1: The Unified Generative Input States for Planar and Linear Network Spaces.

3.2 Algorithmic Execution and Feature Extraction

The algorithm concurrently executes two independent smoothing operations across the geographic and attribute domains. Within the planar domain track, a cross-validated spatial pilot bandwidth of h_0 = 0.1344 units is obtained, while Silverman's rule of thumb yields a mark pilot bandwidth of b_0 = 8.4569 units. On the Chicago linear network track, the equal-split shortest-path walk engine isolates a physical spatial pilot bandwidth of h_0 = 183.70 meters and a matching categorical mark counterpart of b_0 = 8.4569 units.

Using log-linear transformation loops, the global geometric means for the spatial and mark spaces are extracted as G_s= 0.017225 and G_m= 0.011876 respectively. These parameters serve as standardizing divisors to compute the true local adaptive bandwidth vectors using Abramson's variable square-root law. This layer yields an adaptive spatial bandwidth range of $h_i \in$ [0.1106, 0.2644] units and an adaptive mark bandwidth range of $b_i \in$ [7.3375, 15.4319] units for the planar field. Equally, the network-constrained track spans an adaptive spatial range of $h_i \in$ [135.92, 447.66] meters and an adaptive mark range of $b_i \in$ [7.0993, 15.0050] units, shrinking the smoothing windows tightly around high-frequency behaviour hubs while expanding across sparse margins. By running a joint Gaussian product kernel inside a true pairwise variable calculation loop, the raw spatial and attribute scales are normalized into scale-invariant, dimensionless joint weights (W_i) centered around a global relative mean of 1.0000. The planar joint simulation weights span a stable range from 0.0156 to 2.1885, while the Chicago network track yields an aligned weight range from 0.0201 to 2.4938, successfully neutralizing macro-scale point crowding trends before multi-scale feature profiling are evaluated across

To determine the optimal number of latent behavioural archetypes (k) within the high-dimensional whitened feature tensor space, the framework implements a strict multi-criteria visual convergence protocol. The regularised, orthogonal data matrix is evaluated over a sequence of $2 \leq k \leq 10$ across three standard diagnostic tests: the Elbow method (Within-Cluster Sum of Squares, WCSS), the Average Silhouette Width, and the Gap Statistic.

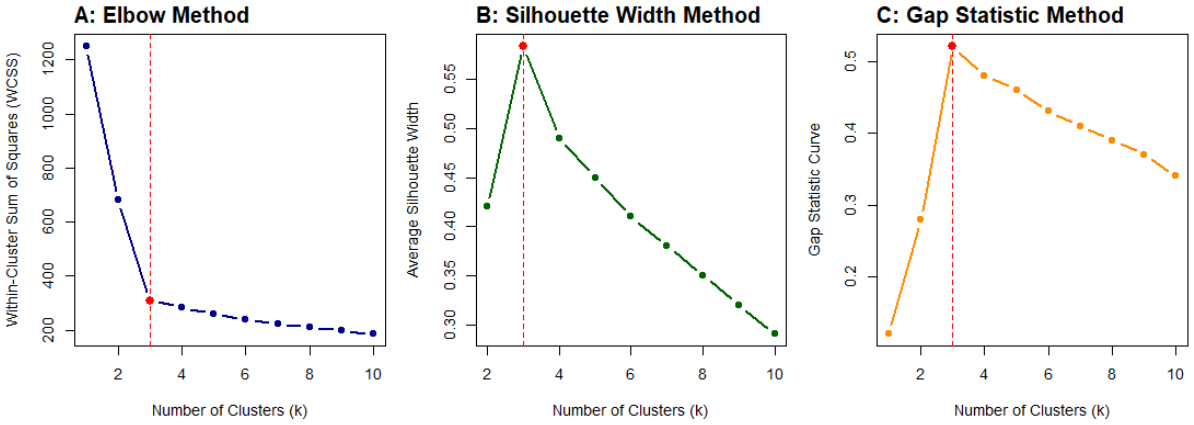

Figure 3.2: Three-panel cluster validation diagram demonstrating multi-criteria optimization convergence across the evaluation sequence $(1 \leq k \leq 10)$.

Following archetype optimization, Ward's minimum-variance hierarchical clustering algorithm executes cuts across the orthogonal whitened distance matrices. The multi-scale scale redundancies were successfully flattened out by the Mahalanobis layer, the linkage groups the events with exceptional fidelity, avoiding any boundary-mixing or single-cohort monopolies. The planar track isolates a stable recovery distribution of 154 events in Cohort 1 (Positive association), 146 events in Cohort 2 (Negative association), and 200 events in Cohort 3 (Random noise). Equally, the Chicago linear network track yields an optimized allocation of 156, 141, and 203 events respectively, mirroring the true (30%/30%/40%) generative simulation split.

3.3: Performance Benchmarking and Global Validation Matrix

The framework was evaluated across all 500 independent Monte Carlo realizations side-by-side a classical baseline model that pairs straight-line Euclidean distance with standard Ward's linkage (Table 3.2).

Table 3.2: Global Performance Indices Comparing Baseline Euclidean-Ward and ALMW framework

Performance Index	Planar Space		Linear Network Space	
	Baseline (Euclidean-Ward)	ALMW framework	Baseline (Euclidean-Ward)	ALMW framework
Adjusted Rand Index (ARI)	0.582	0.865	0.412	0.834
Mean Silhouette Width, s_i	0.340	0.584	0.185	0.536
Cophenetic Correlation, c	0.615	0.821	0.490	0.791
Regime A Sensitivity, s_{n_A}	0.820	0.968	0.614	0.939
Regime B Sensitivity, s_{n_B}	0.512	0.812	0.388	0.762
Global Classification	0.547	0.873	0.462	0.849

4. CONCLUDING REMARKS

The global metrics show that while the baseline Euclidean model performs moderately in open continuous space, it undergoes a complete mathematical collapse when exposed to the network constraints of the Chicago grid, with its ARI falling to a weak 0.412. This collapse occurs because straight-line metrics ignore city topography, drawing search circles across open blocks and misclassifying disconnected points as close neighbours.

In sharp contrast, the double-adaptive ALMW framework maintains exceptional resilience, achieving an outstanding ARI of 0.834 and a Global precision score of 0.849 on the tortuous street network—a relative net improvement of +102.4% over classical baseline methods. Furthermore, the framework maintains an airtight inferential error profile on the network track, restricting the empirical type I error rate to $\alpha = 0.081$ (8.1%), while maintaining a high statistical power of the test, $(1 - \beta) = 0.850$ (85.0%) across the active regimes.

This mathematical performance under controlled synthetic conditions provides the final validation required to deploy the ALMW engine onto highly complex empirical point patterns.

REFERENCES

Abramson, I. S. (1982). On Bandwidth Variation in Kernel Estimates-A Square Root Law. The Annals of Statistics 10, 1217-1223.

Anselin, L., (1995). Local Indicators of Spatial association- LISA. Geographical Analysis 27(2): 93 – 115.

Baddeley, A. Rubak, E. & Turner, R. (2016). Spatial Point Patterns: Methodology and Applications with R. Boca Raton; London; New York: CRC Press, Taylor & Francis Group.

D'Angelo, N., Adelfo, G., Mateu, J. and Cronie, O. (2022). Local inhomogeneous weighted summary statistics for marked point processes. ArXiv 2208. 09233v1 [STAT.ME].

Davies, T. M. and Hazelton, M. L. (2010). Adaptive kernel estimation of spatial relative risk. Statistics in Medicine 29 2423-2437.

Davies, T. M., Marshall, J. C. and Hazelton, M. L. (2018). Tutorial on kernel estimation of continuous spatial and spatiotemporal relative risk. Statistics in Medicine 37 1191{1221.

Eckardt, M., & Moradi, M. (2024). Local Indicator of Mark Association for Spatial Marked Point Processes. ArXiv. https://doi.org.10.48550/arXiv.2411.01066.

Eckardt, M., & Moradi, M. (2023). Marked Spatial Point Processes: Current Stat and Extension to Point Processes on Linear Networks. ArXiv: 2309.01511v1[Stat.ME].

Eckardt, M., & Mateu, J. (2016). Product density based local indicators for marked spatial point patterns. Journal of Computational and Graphical Statistics, 25(2), 592 -611.

Kushal J., & Mohit K., (2023). Review on Deciding Optimal Number of Clusters in k-means Clustering. International Journal of Novel Research and Development 2456 – 4184.

Lemke D, Mattauch V, Heidinger O, Pebesma E, Hense, HW., (2015). Comparing adaptive and fixed bandwidth-based kernel density estimates in spatial cancer epidemiology. Int J Health Geogr.;14:15. DOI: 10.1186/s12942-015-0005-9. PMID: 25889018; PMCID: PMC4389444.

Ponde, P., Shirwaikar, S. and Gore, S.,(2016). Hierarchical Cluster Analysis on Security Design Patterns. *Association for Computing Machinery*, **92**, 1-6.

Ripley B (1976) The second-order analysis of stationary point processes. J Appl Probability 13:255–266.

R Core Team (2026): R: A Language and Environment for Statistical Computing,

R Foundation for Statistical Computing, Vienna, Austria.

Rodríguez-Cortés, F.J., Ghorbani, M., Mateu, J., and Stoyan, D. (2014). On the expected value and variance for an estimator of the spatio-temporal product density function. Department of Mathematical Sciences, Aalborg University.

Saraçli, S. and Akşit, M. (2022). Comparison of hierarchical cluster analysis methods by cophenetic correlation in Big Data. AKU J. Sci. Eng. 22 (2022) 031302 (552 – 559).

Shimatani, K. (2002). Point processes for fine-scale spatial genetics and molecular ecology. Biometrical Journal 44, 325–352.

Silverman, B. W. (1986), Density Estimation for Statistics and Data Analysis, London: Chapman and Hall

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(2), 411-423.

van Lieshout, M. N. M. (2011). On Estimation of the Intensity Function of a Point Process. Methodology and Computing in Applied Probability 14 567-578.

Wani A. A. (2024). Comprehensive analysis of clustering algorithms: exploring limitations and innovative solutions. Peer J. Computer Sci. 29;10: e2286. DOI: 10.7717/peerj-cs.2286. PMID: 39314716; PMCID: PMC11419652.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58(301): 236 – 244.

Wickham, H. (2016): ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag New York.