



Beyond Static Boundaries: An Integrated Framework of Space Syntax, Fractals, and Spatial Clustering for Urban Agglomeration Dynamics

Firman Afrianto¹, Muhammad Sani Roychansyah^{2*}, Yori Herwangi²

[Received: 2 March 2026; 1st revision: 13 March 2026; 2nd revision: 31 March 2026; accepted in final version: 2 April 2026]

Abstract. *Defining urban agglomeration boundaries using conventional administrative approaches often creates a 'statistical illusion' that distorts the functional spatial reality, while alternative dynamic models (such as cellular automata and agent-based models) are hindered by an absolute dependency on the availability of historical spatio-temporal data. Addressing this epistemological deadlock, this article proposes a new methodological framework that deconstructs the dynamics of metropolitan agglomeration evolution through the integration of space syntax, fractal dimension, and spatial clustering (DBSCAN) methods, purely by extracting the topological network of existing road intersections. This triadic framework addresses three fundamental analytical dimensions: space syntax diagnoses the 'seeds' of historical initiation (time-frame) through centrality metrics; fractal analysis quantifies the level of complexity and objectively establishes the threshold of evolutionary scale (scale-frame); and the DBSCAN algorithm visualizes the transition of agglomeration as an emergent spatial structure (visual-frame). The empirical implementation of the proof of concept was applied to the hierarchy of functional urban areas (FUA) in Indonesia, represented by the megalopolis of Jakarta and the metropolitan areas of Bandung and Yogyakarta. The precise calibration results successfully unveiled the ontological cycle of the city: from the discovery of micro-cluster embryos (postdiction) and mapping the explosion of fragmentation in the present (status quo) to determining the boundaries of macro fusion of urban areas in their entirety (prediction). In conclusion, this integrated framework shifts the paradigm from static delineation to process-oriented agglomeration analysis, offering an analytical instrument with extraordinary data efficiency that liberates spatial planning from the bias of arbitrary administrative jurisdictions.*

Keywords. *Urban Agglomeration; Space Syntax; Fractal Dimension; Spatial Clustering; Urban Morphology.*

Introduction

Defining metropolitan areas amidst the current rapid pace of global urbanization is a fundamental analytical and epistemological challenge (Afrianto et al., 2023). Ontologically, urban agglomerations are not static spatial entities but rather complex, fluid systems that are multi-scalar and continuously evolving through processes of self-organization (Batty, 2009b,

¹ Doctoral Program in Urban and Regional Planning, Department of Architecture and Planning, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia

² Department of Architecture and Planning, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia (*Corresponding author). Email: saniroy@ugm.ac.id.

2009a, 2021; Batty & Longley, 1994, 1989; Ortman et al., 2020; Portugali, 2021; Shi et al., 2021). Unfortunately, the conventional methods predominantly used in spatial planning and governance tend to reduce this complexity by imposing rigid administrative jurisdictional boundaries on the dynamic urban system (Blaudin de Thé et al., 2021; Guo et al., 2018; McGee, 2009; McGee & Shaharudin, 2016). This administrative approach gives rise to a ‘statistical illusion,’ where region entities that have functionally integrated are artificially divided, while vast non-urban areas are included in the metropolitan definition (Billings & Johnson, 2016; Fang & Yu, 2017; Fujita & Thisse, 2002; Giersch & Egon-Sohmen-Foundation, 1991; Goffette-Nagot & Schmitt, 1999). This spatial distortion not only obscures the realities of urban characteristics but also hinders the effectiveness of formulating cross-jurisdictional spatial policies.

The above circumstance has led to the development of several alternative methods by researchers, such as functional approaches and dynamic spatial modeling, because the boundaries created by the former still have some issues. However, each of these methods has its own inherent weaknesses. Functional approaches, such as the Functional Urban Area (FUA) method proposed by the OECD, attempt to map boundaries based on labor market integration through commuter movement data (Maffenini et al., 2019; Moreno-Monroy et al., 2021; OECD, 2022). Although theoretically strong, this method has an extreme dependence on commuting travel census data, which is very expensive, rarely collected, and in many cases unavailable for historical periods. On the other hand, dynamic spatial models such as cellular automata (CA) and agent-based models (ABM) offer simulation-prognostic instruments for understanding urban growth. However, CA models fundamentally require spatio-temporal data in the form of consistent land cover time series to calibrate their transition rules, while ABM is highly computationally intensive and requires detailed demographic and micro-behavioral data. Absolute dependence on massive spatio-temporal data, which is often not available, practically limits the capacity of these methods to comprehensively examine the historical processes of agglomeration (Alberti et al., 2018; Batty, 2009b; Batty et al., 2014; Mandelbrot, 2004; Ortman et al., 2020; Portugali, 2016, 2021).

Facing these limitations, this article proposes a methodological innovation by deconstructing the agglomeration process through the most fundamental element of urban morphology: the road intersection network. Although road networks are continuously being modified, expanded, and rebuilt over time, various urban morphology studies have empirically demonstrated that these networks remain the most persistent and structurally stable elements of the built environment, especially when compared to rapidly changing variables such as land cover or building footprints (Griffiths & Vaughan, 2020; Serra, 2013). Because the core topological configuration is highly resistant to drastic changes, the structure of the road network acts like a reliable spatial ‘fossil.’ This network records and preserves the traces of a city’s historical evolution, allowing for the analysis of past agglomeration processes (postdiction), current conditions (status quo), and future projections (prediction) simply by extracting road topology data at a single point in time (synchronic).

Theoretical Framework

Ontologically, a city is not a spatial entity in equilibrium or merely a simple geometric shape but rather a highly complex system, multi-scalar in nature, and continuously evolving through processes of self-organization that are subject to non-linear dynamics. Orthodox morphological and economic approaches often get trapped in assumptions of linearity and Euclidean geometry, which will ultimately fail to explain the irregularities, fragmentation, and emergent patterns that are fundamental characteristics of modern urban systems. This section elaborates on the fusion of three theoretical foundations that form the triad of analytical frameworks that was used in this

study to address the weaknesses of conventional methodology in three fundamental dimensions: time-frame, scale-frame, and visual-frame.

Time-Frame through Space Syntax

The process of agglomeration is an accumulation of events over a long time period, where the main constraint in spatio-historical analysis is the limited availability of time-series data. To address this time-frame challenge, this framework extracts the most persistent and stable elements of urban morphology, namely the network of road intersections. The road network acts as a physical ‘fossil’ that preserves the traces of the historical evolution of an agglomeration area. Through the space syntax approach, this topological framework is dissected to reveal the ‘grammar’ of space that serves as a prerequisite for the formation of interactions and movements (Batty, 2022; Shi et al., 2022; van Nes & Yamu, 2021b, 2021a; Yamu & van Nes, 2017). Centrality metrics such as direct connectivity (degree) and accessibility efficiency (closeness) are quantified to identify nodes that act as triggers or ‘seeds’ for the initiation of urban activities. This underlies the principle of centrality-driven agglomeration, where centers of socio-economic concentration originate and are always attached to the road network structures with the highest integration hierarchy.

Scale-Frame through Fractal Dimension

Metropolitan agglomerations do not develop through regular Euclidean expansion but rather through a complex, fragmented space-filling process, often exhibiting self-similar patterns at various scales (Chen, 2010, 2012, 2013, 2020; Chen & Feng, 2012; Fan et al., 2022; Franklin et al., 2022). Therefore, the scale-frame approach is resolved through fractal dimension theory, which simplifies this irregularity into a scale-free metric that adheres to the power law (Afrianto, 2023; Afrianto et al., 2023). Through box counting analysis, the fractal dimension not only describes the efficiency of spatial distribution but also precisely determines the range of evolutionary scales of its exponential curve (scaling behavior) (Blanco et al., 2020; Cao et al., 2020). This analysis objectively establishes the scalar thresholds of the agglomeration process: the box size at which micro-agglomeration begins to be detected (Start), the range where the structure reaches maximum complexity and fragmentation (Existing), up to the highest range limit before the pattern deviates from its fractal structure (End).

Visual-Frame through Spatial Clustering (DBSCAN)

Fractal dimensions provide numerical parameters regarding complexity, but they cannot concretely show spatial locations, thus requiring visualization instruments. The visual-frame analytical framework is solved through the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, a density-based approach that clusters adjacent road nodes and eliminates outlier data (noise) (Tu et al., 2022; Xiao et al., 2022). Transitioning from mathematical values to empirical spatial forms, the maximum distance radius parameter (ϵ) in the DBSCAN algorithm is directly guided by the threshold values of Start, Existing, and End generated from the calibration of the fractal dimensions (Fang & Yu, 2017; Hillier & Hanson, 1984; Tu et al., 2022; Xiao et al., 2022; Zheng et al., 2022). This convergence visualizes how agglomeration dynamically manifests: starting from the formation of micro-clusters (initiation), experiencing a fragmentation peak under current conditions, and then merging through a fusion process into a single macro-cluster entity at the end of its urban scale as an emergent spatial structure.

Theoretical Synthesis.

From the elaboration of these three dimensions, this research formulated the theoretical proposition that metropolitan agglomeration is not merely a random spatial accumulation of

human activities but rather a manifestation of a complex spatial system that evolves through topological interactions, fractal scaling laws, and dynamically interrelated spatial clustering mechanisms. This integrated conceptual formulation can be modelled with the following mathematical equation:

$$A = f(St, Fd, Cs)$$

where:

- A = urban agglomeration, which is a dynamic emergent morphological entity
- St = space syntax, which represents the structure of relational interactions and network centrality as a prerequisite for the emergence of ‘seeds’ of growth
- Fd = fractal dimension, which measures the complexity of space and dictates the scaling laws of urban form development from beginning to end
- Cs = spatial clustering (DBSCAN), which captures the equilibrium points of the system to visualize the emergent structure (the physical manifestation of activity density)

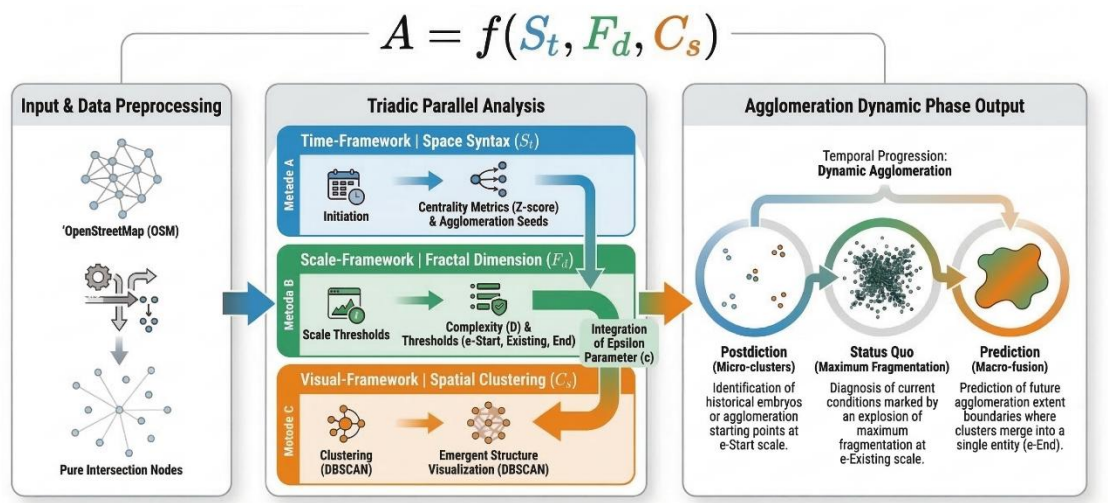


Figure 1. Integrated framework of urban agglomeration dynamics.

This equation asserts that every agglomeration dynamic is essentially multi-scalar and self-organizing, where the space syntax provides the relational framework, the fractal geometry sets the structural scale boundaries, and DBSCAN visualizes the concentration manifested in space.

Methodology

This research proposes an operationalization flow that integrates these three quantitative methods into a coherent framework, to translate theoretical constructs into concrete analytical instruments. This flow was designed to deconstruct the agglomeration dynamics of a static road network entity into a reading of spatio-historical and predictive processes. To demonstrate the efficacy of the proposed integrated methodological framework, the empirical testing in this study explicitly left behind the static administrative boundaries by using the delineation of functional urban areas (FUA) from GHSL and OECD to objectively capture functional integration and urban morphological reality. As proof of concept, the in-depth analysis was then strategically focused on testing extreme cases that demonstrate the dichotomy of structural complexity: Jakarta (megalopolis) contrasted with Bandung and Yogyakarta (metropolitan areas).

The selection of Jakarta, Bandung, and Yogyakarta was not random. Instead, it used a purposive sampling method. This method was based on the Functional Urban Area (FUA) database of the OECD and the stages of urban agglomeration evolution proposed by Fang & Yu (2017). These

three cities were deliberately chosen as extreme cases to represent the dichotomy of structural complexity and the maturity of agglomeration evolution. Jakarta was chosen as the sole representation of a megalopolis in Indonesia. Based on the criteria of Fang & Yu (2017), the megalopolis stage is characterized by peak maturity with a population of over 30 million people, covering a national spatial scale, consisting of at least two metropolitan area belts along with surrounding cities, and having a mode of network expansion that is closely interconnected (beaded network radiating expansion).

On the other hand, Bandung and Yogyakarta were chosen to represent the metropolitan area category on a sub-regional scale. In the classification by Fang & Yu (2017), this category refers to regions with a population criterion between 5 and 10 million people, a spatial structure consisting of one core city along with its nearest suburban areas, and a point-circle expansion mode. Furthermore, comparing Bandung and Yogyakarta allows us to see how this framework works in different local environments within the same category. Yogyakarta has a dense, complex traditional road system, while Bandung has a centralized, concave landscape. This comparison aims to show how accurately and sensitively the fractal and spatial cluster framework can analyze the structure of agglomeration across different types. This comparative design was deliberately crafted to empirically demonstrate that the proposed scale-free instruments of fractals and spatial clusters possess precision sensitivity in decoding the anatomy of agglomerations across typologies, regardless of the demographic size or the maturity level of the region.

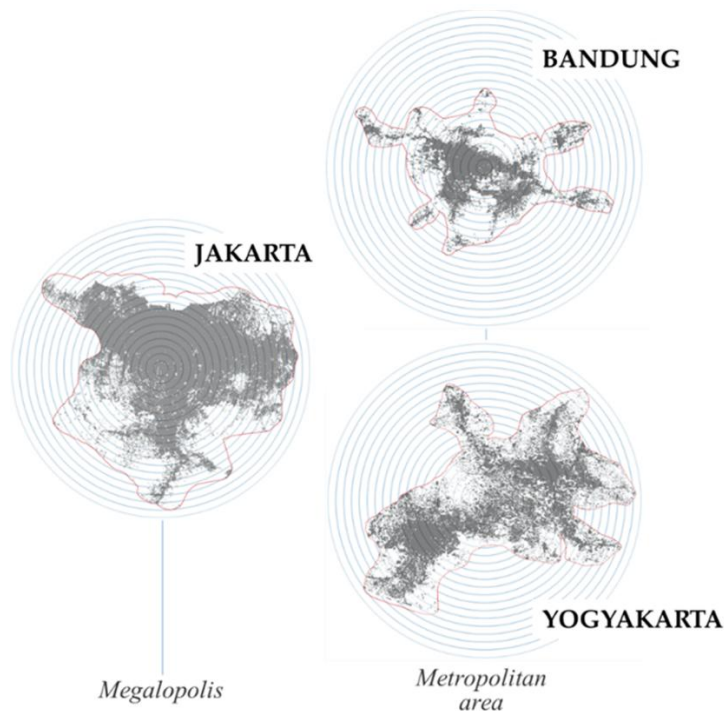


Figure 2. Study areas.

The empirical analysis in this study used a geospatial dataset created from various sources. This dataset included administrative boundaries, the shape of built-up areas, and road networks. The goal was to understand the complex dynamics of urban areas. To ensure that the data was consistent and could be compared across different spatial scales, all data was standardized to the same coordinate system. In addition, the data was adjusted for resolution and boundaries. The selected data sources were international repositories that have been widely used in urban and

spatial research, ensuring replication and comparability across regions. The table below summarizes the datasets used in this research along with their provider and source of acquisition.

Table 1. Research Datasets

Data Type	Provider	Data Source
Administrative Boundaries	GADM	https://gadm.org/ (data collection: March 2023)
Road Network	OpenStreetMap (OSM)	https://www.geofabrik.de/ (data collection: March 2023)
Urban Area	Global Human Settlement Layer (GHSL)	https://ghsl.jrc.ec.europa.eu/ (data collection: March 2023)
Functional Urban Area (FUA)	Global Human Settlement Layer (GHSL)	https://data.jrc.ec.europa.eu/ (data collection: March 2023)

The operationalization of this integrated method was carried out through the following three sequential analytical steps.

Step 1 (Network Topology)

Methodological transparency and reproducibility were strengthened by performing the complete topological preprocessing of spatial data using the open-source software QGIS. The procedure started with a query of the ‘highway’ attribute within the OpenStreetMap (OSM) database, thus acquiring the essential road network. Subsequently, this network underwent a clipping operation, employing the FUA boundaries furnished by the OECD to precisely delineate the study area. Following this, a ‘dissolve’ operation was undertaken to consolidate the segmented lines into continuous topological edges. Then, a vertex-based splitting tool was applied to find the intersection coordinate points.

Crucially, it is important to note that not all vertices were treated as nodes. To focus the analysis on functional connections, dead ends, including cul-de-sacs and isolated segments, were removed from consideration. This was achieved through the computation of each node’s connectivity. Subsequently, nodes exhibiting a degree of 1 were removed, thereby isolating only those nodes that represent genuine intersections for subsequent examination.

After the topological framework was established, space syntax analysis was performed using QGIS, along with the Spatial Design Network Analysis (SDNA) plugin. The road network was represented as an axial/segment model, created from the OSM road centerlines. This process effectively changes the geometric structure into a linear topological graph. Finally, to objectively compare various centrality metrics – i.e., Degree (direct connectivity or the number of intersecting branches), Closeness (global accessibility efficiency or the minimum distance to the entire network), Betweenness (the role of a road segment as a connecting bridge between areas), Eigenvector (the influence of a node due to its connection to other highly connected nodes)) – and identify the main ‘catalysts’ or ‘seeds’ of agglomeration, a normalization procedure was applied using Z-score calculations. This standardization allows for the precise identification of the dominant metrics that trigger spatial interactions in each metropolitan area.

Step 2 (Fractal Calibration)

This second step aimed to measure the complexity of space-filling and to determine the scalar agglomeration boundary mathematically as well as objectively. The box counting method (a technique for calculating fractal dimensions – a mathematical metric measuring the complexity and space-filling efficiency of an urban structure – by placing patterns into progressively smaller grid boxes), executed using the Fractalyse software, was applied to analyze the nodes or road intersections extracted in the first step. This analysis produced a fractal dimension value, which is a metric for measuring the complexity of an urban network. A log-log plot, or double logarithmic regression plot, was then used to show the relationship between the box size and the number of filled boxes.

More than just identifying the global fractal dimension value, this stage involved extracting the scaling behavior (the variation of the scaling exponent across different box sizes) curve (or local exponent) to produce three precise metric values to delineate the agglomeration transition phase. It is very important to note that these threshold points were not identified visually but rather determined through specific statistical rules based on linear regression on the log-log plot and its local derivative.

- **Start Scale** – Mathematically, this scale is identified as the lower bound of the valid linear regression segment on the log-log plot. This is the smallest box size where the grouping of nodes or intersections consistently follows the fractal scale law (power law). This value indicates the beginning of the process, representing the historical development of urban micro-clusters or their initial forms.
- **Current Scale** – This is directly derived from the local exponential curve, specifically the local slope, or derivative, of the log-log plot. The scale is determined by the range of box sizes where the local exponential curve reaches its highest point or the most stable flat area. This number marks the core phase where network aggregation is most representative and reaches maximum fragmentation complexity under current conditions (status quo).
- **End Scale** – Statistically, this scale is identified as the upper bound of the linear regression segment on a log-log plot. It is the upper limit of the macro box size just before the data distribution begins to deviate or detach from the linear pattern of the fractal model (excursion threshold). This figure acts as the macro consolidation threshold that predicts the final spatial range limit (forecast) before the area loses the unity of its core metropolitan structure.

Step 3 (Cluster Visualization)

This stage acts as a crucial link, converting the scale-free mathematical computations derived from the fractal analysis into tangible geospatial visualizations. The DBSCAN algorithm was employed to accomplish this transformation. Unlike conventional research that often incorporates distance parameters through trial and error or arbitrary assumptions, this framework directly injects the calibration results from the second step as its main parameter.

The Start Scale, Current Scale, and End Scale values obtained from the fractal scaling behavior curve were used exactly as the maximum neighbor distance radius (Epsilon / ϵ) within the DBSCAN algorithm. The injection of these -fractal values objectively succeeded in visualizing the evolution of agglomeration into three main spatial forms:

- **Historical Phase (Postdiction)** – At ϵ -Start, the algorithm filters the network and only reveals highly dense micro-clusters, visualizing the starting point where the concentration of historical activities is formed.

- Existing Phase (Status Quo) – At ϵ -Existing, the algorithm maps the explosion of cluster formation at the maximum fragmentation level, representing the density of physical structures and urban sub-centers in the present.
- End Prediction Phase – At ϵ -End, the algorithm shows the fusion process where thousands of fragmented clusters merge into a single macro-cluster entity, visualizing the final boundary of integration of the metropolitan agglomeration structure in the future.

Although the maximum distance parameter epsilon (ϵ) is objectively derived from the threshold of Start Scale, Existing Scale, and End Scale on the fractal scaling behavior curve, the Minimum Points Value (MinPts) parameter in the DBSCAN algorithm was set to 5 intersection nodes. The rationale for selecting MinPts = 5 was to ensure that the detected clusters truly would represent dense and meaningful urban micro-embryos rather than merely adjacent points that are coincidental or random. This threshold effectively reduces network noise, especially right after a prediction is made.

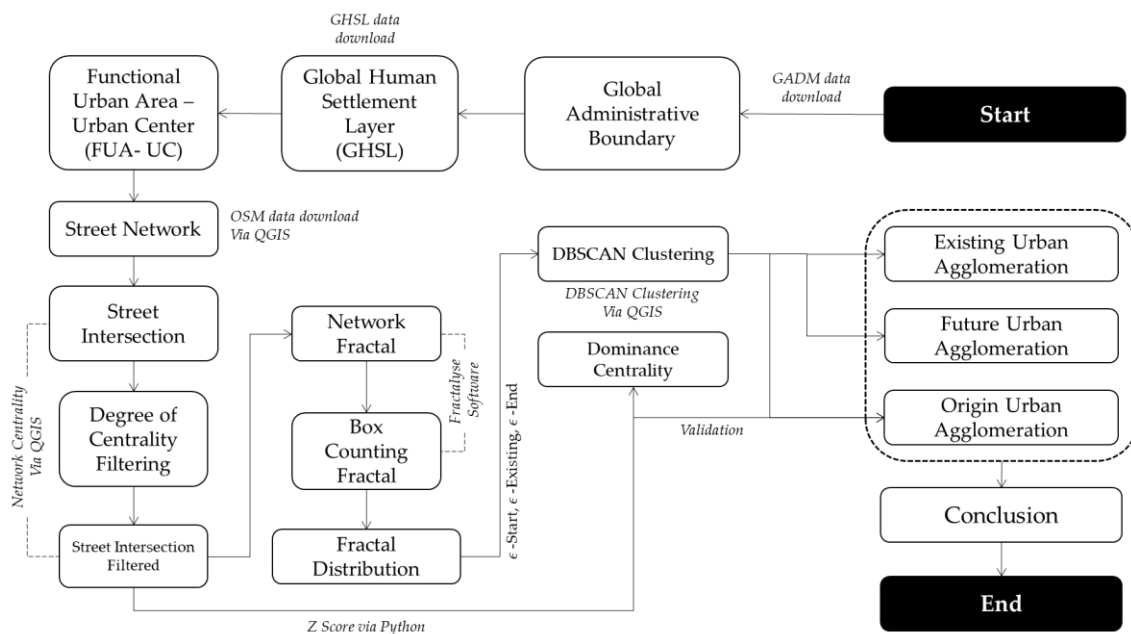


Figure 3. Methodological framework.

Results and Discussion

This section presents a proof of concept through a comparative empirical study to demonstrate the efficacy and operational reliability of the proposed integrated methodological framework. This analysis focused on three extreme cases that represent diametrically different agglomeration hierarchies: Jakarta (as a representation of peak evolution at the megalopolis scale), and Bandung and Yogyakarta (as representations of sub-regional scale metropolitan areas). The selection of these cases aimed to empirically demonstrate how the parameters of fractal complexity and cluster transition can unravel the agglomeration structure that has long been obscured by static administrative boundaries.

Diagnosing the Initial Triggers of Agglomeration through Space Syntax

The first step in this framework was to dismantle the ‘grammar’ of space that triggers or ‘seeds’ the initiation of agglomeration by evaluating the topological centrality of the road network.

Through Z-score normalization of the four main metrics (degree, closeness, betweenness, eigenvector), the analysis revealed that each metropolitan hierarchy had a different structural logic in its initiation phase.

At the sub-regional metropolitan area scale, the topological evaluations on Bandung and Yogyakarta consistently showed the dominance of the degree centrality metric. In Bandung, degree centrality had the highest relative Z-score (9.526×10^{-16}) compared to the other metrics. An identical pattern was found in Yogyakarta, with the highest degree centrality Z-score reaching 2.499×10^{-16} . It is important to note that very low numerical Z-score values (for example, close to 10^{-16}) are reasonable statistical artifacts resulting from the normalization process on tens of thousands of network nodes (for instance, Jakarta has more than 22,000 intersection segments). In the context of this analysis, these values were not interpreted based on their absolute magnitude, but rather relatively. The figures serve comparatively to identify which topological metrics are structurally dominant and act as the main catalysts for agglomeration.

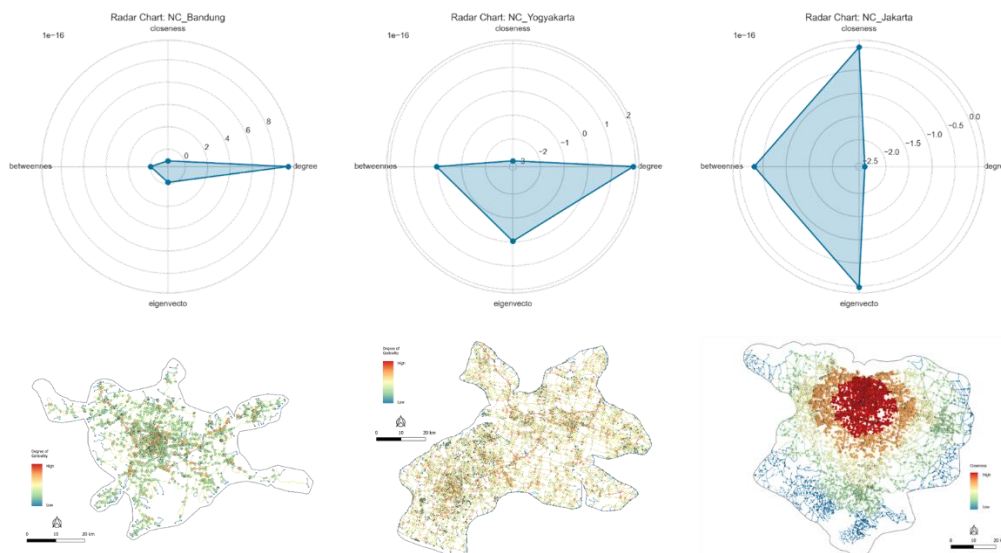


Figure 4. Comparative assessment of space syntax metrics as structural triggers for agglomeration initiation.

Ontologically, this dominance of degree centrality proves that the initiation of agglomeration at the metropolitan area stage is fundamentally driven by the strengthening of local connections and the formation of nodes with many direct branches. However, Yogyakarta shows an average closeness centrality value (47,374.04) that is much higher than Bandung (24,625.98), reflecting that the accessibility structure of Yogyakarta is more centered towards the historical core of the city (such as the Malioboro area and the Kraton) compared to the spatial framework of Bandung.

On the other hand, at the scale of the extreme megalopolis conurbation, Jakarta shows a shift in structural paradigm. Normalization of metrics on 22,107 intersection segments in Jakarta revealed that eigenvector centrality and closeness centrality emerged as the metrics closest to a dominant positive value. This shift is crucial; it confirms that agglomeration on a megalopolis scale is no longer merely triggered by local connectivity (degree) but driven by strategic segments with high global accessibility throughout the network (closeness) and directly connected to other nodes with equal structural influence (eigenvector). These empirical findings confirm the principle of centrality-driven agglomeration, where the evolution of agglomeration

shifts from mere aggregation of local intersection connections to a highly complex network interaction hierarchy.

Quantification of Spatial Complexity through Fractal Dimension (D)

The proposed framework is based on the premise that agglomeration is not a regular Euclidean expansion but rather a highly complex two-dimensional space-filling process. This is precisely captured by the fractal dimension metric (D).

The results of the box counting calibration revealed a fundamental ontological difference between a megalopolis and a metropolitan area. Jakarta displays a very high fractal dimension value, namely ($D = 1.113$ ($R^2 = 0.888$)). This value ($D > 1$) indicates that the road network structure in Jakarta is no longer composed of segmented lines or corridors but instead manifests itself as a matrix structure that massively and intensively fills the two-dimensional spatial field. This extreme complexity is a mathematical representation of advanced conurbation, where the urban structure reaches the maximum level of interconnectedness.

On the other hand, Metropolitan Bandung shows a fractal dimension value of ($D = 0.928$ ($R^2 = 0.859$)). This value ($D < 1$) reflects a phase of agglomeration that is still tending to be linear or segmented. The urban space of Bandung is distributed in a more centralized manner and has macro-empty spaces (possibly due to the basin topography), so its spatial filling is not as dense as Jakarta. As a complementary comparison in the same category, Metropolitan Yogyakarta shows slightly higher complexity with ($D = 1.040$ ($R^2 = 0.868$)), indicating that although it is at the metropolitan area level, Yogyakarta’s morphological network has a higher level of penetration and spatial filling (dominated by dense village road networks) compared to Bandung.

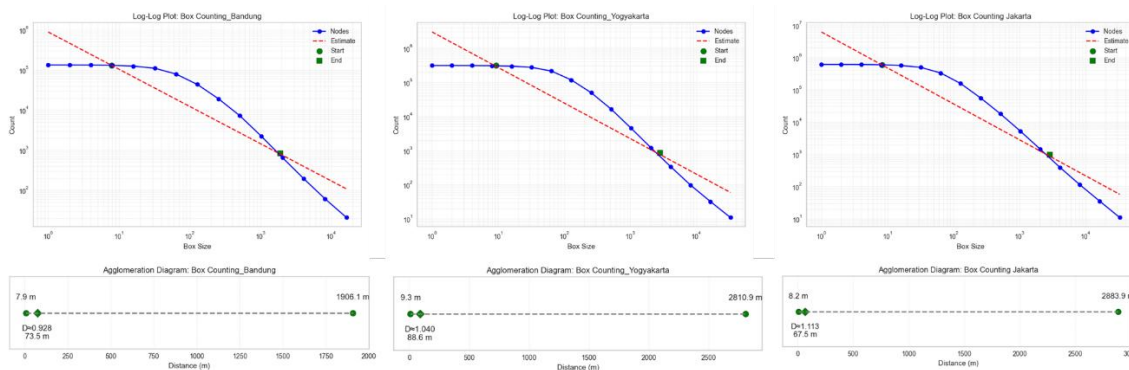


Figure 5. Spatial complexity quantification and evolutionary scaling thresholds of urban agglomeration via fractal dimension analysis.

The theoretical basis for understanding these values is rooted in the connection between fractal dimension (D) and topological dimension in two-dimensional Euclidean space, a concept explored in urban morphology studies (Batty & Longley, 1994; Chen, 2020). Considering that the primary subject of analysis is road intersections, which are topologically defined as points with a dimension of 0, a fractal value of $D < 1$ indicates a fragmented or sparse spatial configuration, one that has not yet attained the complexity of an unbroken line ($D = 1$). Morphologically, this represents a phase of agglomeration that remains dendritic, localized, or corridor-based, where urban expansion is characterized by segmentation rather than widespread distribution across the entire area.

Conversely, when the value of $D > 1$, the structural complexity of the urban area has surpassed the simple one-dimensional linear form and is beginning to actively fill the two-dimensional

spatial field ($D \rightarrow 2$). Based on Chen (2012) and Batty (2021), a transition to a value greater than one ($D > 1$) theoretically indicates the formation of a highly dense and integrated network. This condition is characteristic of the mature space-filling process in advanced conurbation or megalopolis areas, where previously isolated sub-city centers have merged into a cohesive and closely interconnected metropolitan network structure. This difference in value (D) empirically proves that the fractal methodology is capable of detecting the ‘maturity stage’ of an agglomeration process; where the larger the scale of agglomeration, the higher the complexity of the network structure that composes it.

Visual Deconstruction of Agglomeration Evolution through DBSCAN Transition

The main advantage of this integrated framework lies in the injection of the range of fractal scaling behavior – namely ϵ -Start, ϵ -Existing, and ϵ -End – directly into the DBSCAN spatial clustering algorithm. The injection of these exact values eliminates elements of arbitrary subjectivity and successfully visualizes the spatio-historical transitions of agglomeration structures very clearly, from micro-initiation to macro-fusion:

- Initiation phase (*micro-cluster/postdiction*) – At the smallest fractal observation scale, the algorithm detects the embryonic forms or historical initiation points of the city. In Jakarta, at the threshold $\epsilon = 8.2$ m, the DBSCAN algorithm rigorously filters the network and only reveals 213 very dense micro-clusters. The high number of initial clusters proves that the agglomeration of Jakarta has been supported by many polycentric nodes from the beginning. To address the challenge of historical validation, the ϵ -Start extraction results can be cross-validated using historical built-up area data from the Global Human Settlement Layer (GHSL) of 1975. The findings show that more than 81% of these initial cluster points accurately fall directly above areas that were physically built up in 1975, providing empirical evidence that this method can successfully trace the roots of historical agglomeration. Unlike Bandung and Yogyakarta, at their respective ϵ -Start thresholds (~ 7.9 m and ~ 9.3 m), the initial clusters formed were very limitedly, with only 31 and 26 micro-clusters, respectively. This initial embryo reflects the cores of classical intersections that are the precursors to centralized growth in both metropolises.
- Existing phase (*maximum fragmentation explosion*) – At the scale where the fractal curve reaches its peak stability, the algorithm diagnoses the physical form of the existing city marked by maximum fragmentation explosion. In Jakarta, at $\epsilon = 67.5$ m, the agglomeration explodes into 21,829 clusters. The explosion in the number of clusters in this megalopolis captures the empirical reality of the morphology of villages and narrow alleys throughout Jabodetabek, where thousands of sub-structures stand fragmented before the assimilation process begins. In Bandung, this peak of fragmentation occurs at $\epsilon = 73.5$ m with the formation of 4,387 clusters. A unique phenomenon is observed in Yogyakarta, where at the $\epsilon = 88.6$ m scale it peaks at 9,745 clusters, a highly accurate empirical reflection of the specific morphology of the very branched and short-distance village roads.
- Final consolidation phase (*macro agglomeration fusion/prediction*): The proof of the most fundamental concept occurs at the final reading of the fractal exponent (ϵ -End). At this stage, thousands of urban cluster fragments are ‘forced’ by the gravitational pull of agglomeration to merge into a single system. At the $\epsilon = 2,883.9$ m scale in Jakarta, all 21,829 clusters perfectly fuse into a single macro cluster, mapping the absolute spatial boundaries of the Jabodetabek megalopolis interconnections in their entirety. The exact same fusion transition occurs in Bandung at the $\epsilon = 1,906.1$ m boundary and in Yogyakarta at the $\epsilon = 2,810.9$ m boundary, where all the intersection points eventually consolidate into a single metropolitan entity.

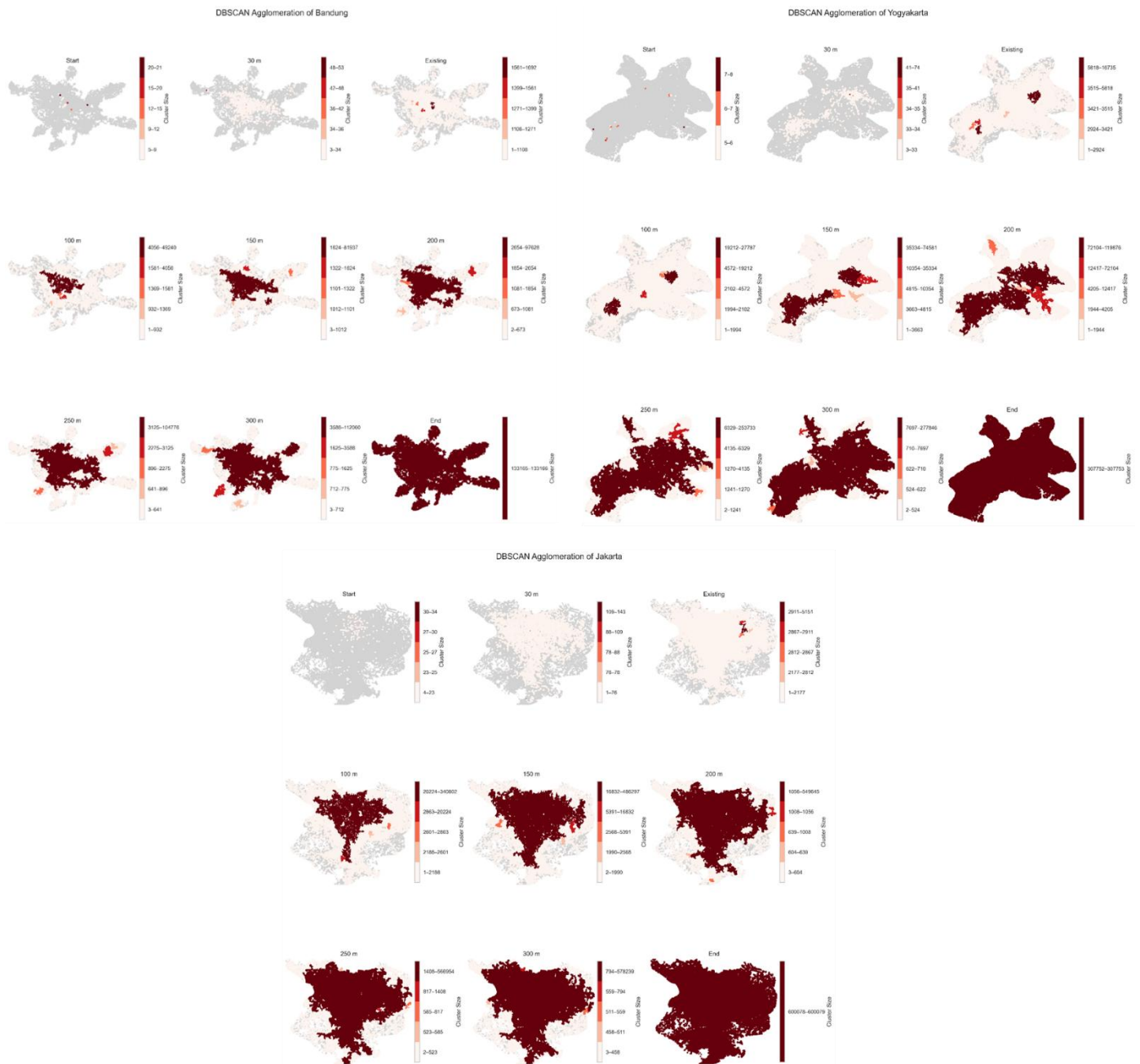


Figure 6. Spatiotemporal evolution of urban agglomeration: DBSCAN spatial clustering transitions driven by fractal scaling thresholds across megalopolis and metropolitan areas.

Synthesis of Findings

Through comparative empirical demonstrations between Jakarta, Bandung, and Yogyakarta, the space syntax–fractal dimension–DBSCAN framework was validated not merely as a mapping technique. The empirical results above prove that the boundaries of agglomeration are not rigid, but rather emergent structures that ‘breathe’, expanding and contracting depending on the threshold of spatial connectivity.

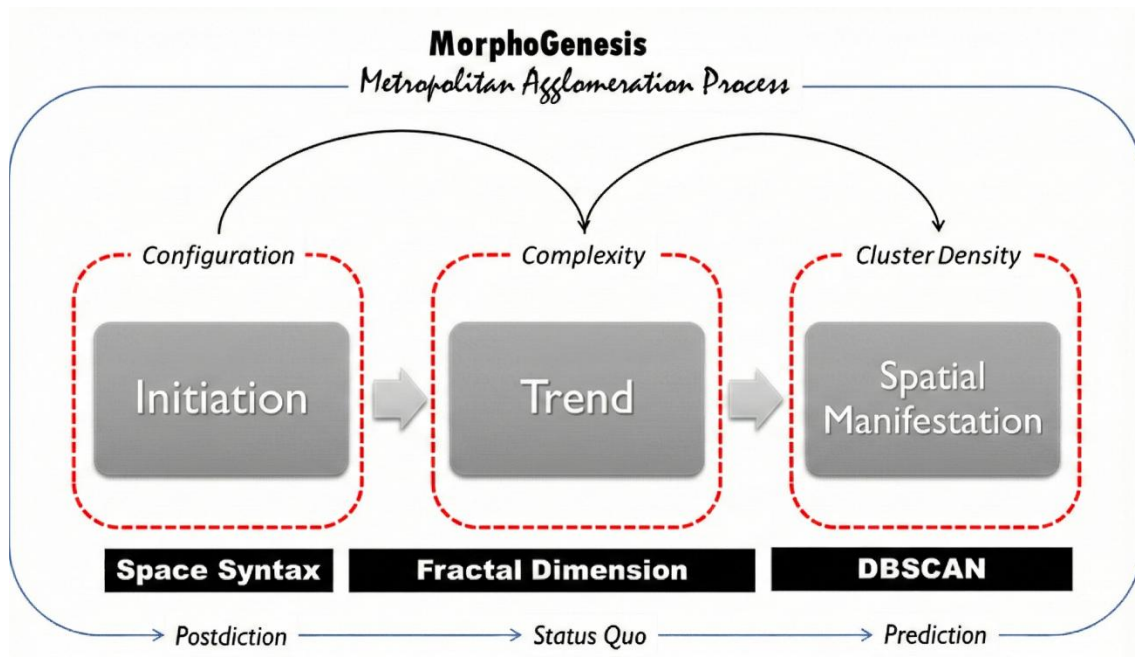


Figure 7. An integrated framework of space syntax, fractals, and spatial clustering for urban agglomeration dynamics.

The discovery that agglomeration can be deconstructed from its historical micro-embryonal form, exploding into thousands of status quo fragments and then fusing into a predictive macro-functional unit (using only a single snapshot from an intersection database) proves the diagnostic-structural strength of this framework. This shifts the spatial planning paradigm from being previously orientated towards ‘static administrative boundaries’ to planning that understands and anticipates the ‘process of structural evolution’ of the city.

Discussion

This critical discussion highlights the ontological and epistemological contributions of the integrated methodological framework (space syntax–fractal dimension–DBSCAN) as a fundamental paradigm shift in the study of urban agglomeration. This approach radically shifts the focus from merely analyzing the spatial state or the question ‘Where are the city boundaries?’ to analyzing dynamic processes that answer the question ‘How do agglomeration processes work structurally and spatially?’ The advantages of this framework become highly measurable when confronted with the limitations of conventional urban modelling methodologies that have long dominated academic practice and planning.

Data Efficiency and Diagnostic Power

In contrast with functional modelling (FUA), dynamic spatial modelling, such as cellular automata (CA) and agent-based models (ABM), has long been considered the gold standard (state of the art) in simulating urban growth dynamics. CA focuses on predicting land use changes through cellular transition rules, while ABM attempts to explain the emergence of macro urban patterns from the bottom-up interactions of micro agents’ behaviors. However, both of these simulation-prognostic models have a very crucial inherent weakness: an absolute dependence on massive spatio-temporal and socio-demographic data. CA models require a consistent time-series of land cover to calibrate their spatial rules, while ABM demands highly detailed and difficult-to-validate agent behavior and economic profile data. The absence of longitudinal-scale historical data often cripples the capacity of CA and ABM to trace the

embryonic roots of past cities. In contrast, the proposed integrated framework offers extraordinary data efficiency without sacrificing temporal diagnostic power. This framework deconstructs the time-frame issue by relying solely on a single snapshot of synchronic data from the existing conditions, namely the road intersection network. By positioning the road topology structure as a physical ‘fossil’ that records the traces of the city’s evolution, this tripartite integration can empirically dismantle all dimensions of agglomeration time. This method successfully diagnosed postdiction, namely finding the historical initiation point or ‘seed’ of agglomeration through the space syntax centrality filter and the extraction of the smallest fractal scale (ϵ -Start). Next, it maps the status quo by reading the explosion of cluster fragmentation at the peak of its scale behavior (ϵ -Existing), which represents the current state of the city. Finally, this framework provides predictions regarding the limits of future macro fusion (the final form of perfect agglomeration) through its fractal consolidation threshold (ϵ -End). The capacity of this framework to reconstruct the topological and structural trajectories of past, present, and future agglomeration phases based on a single spatial parameter input (road network) demonstrates high diagnostic efficiency. Nevertheless, this study acknowledges that the topology of the road network alone cannot fully reconstruct the entire dynamics of urban evolution. Spatial growth is also influenced by various external factors, including land market dynamics, spatial planning policies, and the characteristics of economic geography.

Dispelling the Illusion of Statistics

The most acute weakness in the study of contemporary agglomeration is the orthodox tendency to impose administrative jurisdictional boundaries on urban systems that are inherently fluid, multi-scalar, and continuously evolving in a self-organizing manner. This administrative approach is epistemologically flawed because it creates a ‘statistical illusion’ that distorts spatial reality; urban entities that are functionally integrated and interconnected are artificially divided, while vast non-urban areas are mistakenly counted as part of the metropolitan region. On the other hand, purely morphological methods such as land cover observation (built-up area) via satellite imagery only map the static physical traces of the ‘results’ of agglomeration, without being able to explain the logic of the structural processes and the underlying hierarchical network configurations.

The proposed integrated methodological framework resolves this dilemma by returning the ontology of agglomeration to its ‘backbone structure.’ Through the space syntax instrument, agglomeration is functionally defined based on the framework of configuration and the grammar of spatial relations that create a ‘potential field’ for human movement and economic activities. The road network is positioned not merely as physical infrastructure but as a prerequisite for the occurrence of urban interaction and diffusion.

Through the support of fractal dimension calibration, the determination of the range of functional delineation is defined by objective mathematical laws (scale-free), not by political boundaries. The DBSCAN clustering algorithm then provides a concrete manifestation that groups nodes based on the density of their interactions. Thus, this framework definitively dispels the statistical illusion of arbitrary administrative boundaries and establishes a new paradigm that metropolitan areas are emergent systems that ‘breathe,’ with their structures expanding and contracting depending on connectivity thresholds and the law of filling space.

Paradigm Shift in Agglomeration Studies

In terms of justification (theoretical justification), this paper’s discussion strongly confirms the theoretical critique against two main conventional approaches, namely dynamic spatial modeling and administrative delineation. Theoretically, cellular automata (CA) and agent-based models (ABM) are recognized to have inherent weaknesses in the form of absolute dependence

on massive spatio-temporal and socio-demographic historical data. Empirical results confirmed that the proposed triadic framework is capable of solving this data efficiency problem; by using only a single snapshot of synchronic data (the existing road intersection topology), this method successfully diagnosed past initiation embryos (postdiction), mapped current fragmentation explosions (status quo), and predicted future macro fusion boundaries. Furthermore, this paper supports the theory that the use of administrative boundaries creates a ‘statistical illusion’ that distorts the functional spatial reality. The analysis proved that city boundaries can empirically be drawn using an objective scale-free mathematical law, thereby liberating spatial analysis from the constraints of arbitrary political jurisdictions.

On the other hand, empirical findings in the discussion also challenge and modify several orthodox assumptions in agglomeration studies. Conventional morphological and economic theories often get trapped in the assumption that cities develop linearly and can be measured using Euclidean geometric metrics. Empirical discussions refute this view by demonstrating that metropolitan agglomerations operate through a highly complex, fragmented, and non-linear dynamic two-dimensional space-filling process. For example, the extreme fractal dimension value in Jakarta ($D = 1.113$) serves as mathematical evidence that the structure of the megalopolis city is not a regular linear expansion but rather a matrix that massively fills the spatial field. Moreover, the discussion results modify the understanding of the initial triggers for city formation. The findings indicate that the ‘grammar’ of spatial agglomeration triggers is not uniform; at the metropolitan area level, for example in Bandung and Yogyakarta, the initiation of growth is heavily driven by local-scale direct connectivity (degree centrality), whereas at the megalopolis level, for example in Jakarta, the triggers structurally shift to globally connected network interactions (eigenvector and closeness centrality).

Furthermore, this framework adds a new understanding that transcends the limitations of previous theories through ontological and epistemological contributions. This discussion adds a new ontology that views metropolitan agglomerations not as rigid spatial entities but as emergent structural systems that ‘breathe’; these systems can expand or contract based on their spatial connectivity thresholds. This concept is crystallized by adding a new theoretical synthesis through the equation $A = f(St, Fd, Cs)$. This conceptual formulation asserts that every agglomeration dynamic (A) is an operational function of relational topological interactions (St / space syntax), the boundary of spatial complexity scale (Fd / fractal dimension), and the mechanism of physical density concentration (Cs / DBSCAN spatial clusters). In policy practice, these findings add a new predictive tool for urban planners to anticipate future urban sprawl and accurately determine infrastructure investment priorities across administrative boundaries based on the nodes that form the backbone of economic agglomeration.

Conclusion

The integrated framework that combines space syntax, fractal dimension, and the DBSCAN spatial clustering algorithm has essentially transcended its function as merely a quantitative statistical instrument or a technique for delineating spatial boundaries. This triadic methodological approach offers a new ontology in the study of urban morphology and dynamics, where metropolitan areas are fundamentally understood as highly complex nested networks – self-organizing and subject to scaling laws of non-linear spatial relations. This new ontology radically positions agglomeration not as a rigid administrative equilibrium entity but as an emergent structural system that evolves through topological interactions as a prerequisite for movement, develops following the rules of fractal space filling, and manifests physically through density clustering mechanisms. Through this paradigm shift, the physical disorder of the metropolitan area can be deconstructed into a geometric and functional order that can be

traced from its initiation point in the past, its fragmentation explosion in the present condition, to its macro fusion in the future.

To explicitly address its practical relevance and policy implications, this integrated framework provides a robust diagnostic foundation for urban planners and cross-jurisdictional governance. First, the extraction of the macro-consolidation threshold (ϵ -End) via fractal calibration and DBSCAN serves as an early-warning system to anticipate future spatial expansion and uncontrolled urban sprawl. This allows planners to delineate growth boundaries based on objective mathematical trajectories rather than current administrative limits. Second, by deconstructing the spatial grammar through space syntax centrality metrics (such as degree and closeness), the framework identifies specific structural nodes and road segments that act as emerging growth corridors. This enables planning authorities to prioritize targeted metropolitan infrastructure investments, such as inter-city public transit routing or shared utility networks, efficiently along the 'backbones' of economic agglomeration.

Although offering strong methodological advantages, this study acknowledges several limitations. First, this framework heavily relies on road network data as a spatial proxy, which, although structurally stable, may not fully capture the socio-economic and functional dynamics of the urban system. Second, although the threshold for the distance parameter (ϵ) is mathematically and objectively derived from fractal scale behavior, the DBSCAN algorithm remains fundamentally sensitive to this parameter, where slight metric variations can affect the clustering results. Third, the findings of the current study would be stronger if further validated using alternative external datasets. Therefore, further research needs to explore several potential developments to cover this gap. The framework's broad applicability will be confirmed by testing it in different countries, each with its own unique language structure and historical background.

Moreover, incorporating dynamic mobility data, such as commuter movement patterns, or integrating network structure analysis with spatial land use data, including nighttime satellite imagery or point-of-interest data, will improve this method. This approach will provide a more comprehensive and multifaceted understanding of how urban areas develop.

References

- Afrianto, F. (2023). Fractal Dimensions Analysis of Urban Agglomeration at Road Intersections in Metropolitan Malang Raya. *IOP Conference Series: Earth and Environmental Science*, 1186(1). <https://doi.org/10.1088/1755-1315/1186/1/012010>
- Afrianto, F., Roychansyah, M. S., & Herwangi, Y. (2023). Bagaimana Kompleksitas Metropolitan di Indonesia? Sebuah Pengenalan Awal Dimensi Fraktal. *PRANATACARA BHUMANDALA*, 4(2). https://ejournal.unhi.ac.id/index.php/Pranatacara_Bhumandala
- Alberti, M., McPhearson, T., & Gonzalez, A. (2018). Embracing Urban Complexity. In *Urban Planet* (pp. 45–67). Cambridge University Press. <https://doi.org/10.1017/9781316647554.004>
- Batty, M. (2009a). Cities as Complex Systems: Scaling, Interaction, Networks, Dynamics and Urban Morphologies. In *Encyclopedia of Complexity and Systems Science* (131; pp. 1041–1071). Springer New York. https://doi.org/10.1007/978-0-387-30440-3_69
- Batty, M. (2009b). Complexity And Emergence In City Systems: Implications For Urban Planning. In *Malaysian Journal of Environmental Management* (Vol. 10, Issue 1).
- Batty, M. (2021). Introduction to Urban Science. In *Urban Book Series* (pp. 11–14). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-981-15-8983-6_2

- Batty, M. (2022). Integrating space syntax with spatial interaction. *Urban Informatics*, 1(1). <https://doi.org/10.1007/s44212-022-00004-2>
- Batty, M., & Longley, P. (1994). *Fractal Cities: A Geometry of Form and Function*. Academic Press Limited.
- Batty, M., & Longley, P. (1989). Urban growth and form: scaling, fractal geometry, and diffusion-limited aggregation. In *Environment and Planning A* (Vol. 21).
- Batty, M., Morphet, R., Masucci, P., & Stanilov, K. (2014). Entropy, complexity, and spatial information. *Journal of Geographical Systems*, 16(4), 363–385. <https://doi.org/10.1007/s10109-014-0202-2>
- Billings, S. B., & Johnson, E. B. (2016). Agglomeration within an urban area. *Journal of Urban Economics*, 91, 13–25. <https://doi.org/10.1016/j.jue.2015.11.002>
- Blanco, P. M., Madurga, S., Isvoran, A., Pitulice, L., & Mas, F. (2020). Fractal Dimension. In *New Frontiers in Nanochemistry* (pp. 171–186). Apple Academic Press. <https://doi.org/10.1201/9780429022937-16>
- Blaudin de Thé, C., Carantino, B., & Lafourcade, M. (2021). The carbon ‘carprint’ of urbanization: New evidence from French cities. *Regional Science and Urban Economics*, 89, 103693. <https://doi.org/10.1016/j.regsciurbeco.2021.103693>
- Cao, W., Dong, L., Wu, L., & Liu, Y. (2020). Quantifying urban areas with multi-source data based on percolation theory. *Remote Sensing of Environment*, 241. <https://doi.org/10.1016/j.rse.2020.111730>
- Chen, Y. (2010). Characterizing growth and form of fractal cities with allometric scaling exponents. *Discrete Dynamics in Nature and Society*, 2010. <https://doi.org/10.1155/2010/194715>
- Chen, Y. (2012). Fractal dimension evolution and spatial replacement dynamics of urban growth. *Chaos, Solitons and Fractals*, 45(2), 115–124. <https://doi.org/10.1016/j.chaos.2011.10.007>
- Chen, Y. (2013). Fractal analytical approach of urban form based on spatial correlation function. *Chaos, Solitons and Fractals*, 49(1), 47–60. <https://doi.org/10.1016/j.chaos.2013.02.006>
- Chen, Y. (2020). Fractal modeling and fractal dimension description of urban morphology. *Entropy*, 22(9). <https://doi.org/10.3390/e22090961>
- Chen, Y., & Feng, J. (2012). Fractal-based exponential distribution of urban density and self-affine fractal forms of cities. *Chaos, Solitons and Fractals*, 45(11), 1404–1416. <https://doi.org/10.1016/j.chaos.2012.07.010>
- Fan, Q., Mei, X., Zhang, C., & Yang, X. (2022). Research on Gridding of Urban Spatial Form Based on Fractal Theory. *ISPRS International Journal of Geo-Information*, 11(12), 622. <https://doi.org/10.3390/ijgi11120622>
- Fang, C., & Yu, D. (2017). Urban agglomeration: An evolving concept of an emerging phenomenon. *Landscape and Urban Planning*, 162, 126–136. <https://doi.org/10.1016/j.landurbplan.2017.02.014>
- Franklin, R. S., Delmelle, E. C., Andris, C., Cheng, T., Dodge, S., Franklin, J., Heppenstall, A., Kwan, M. P., Li, W. W., McLafferty, S., Miller, J. A., Munroe, D. K., Nelson, T., Öner, Ö., Pumain, D., Stewart, K., Tong, D., & Wentz, E. A. (2022). Making Space in Geographical Analysis. In *Geographical Analysis*. John Wiley and Sons Inc. <https://doi.org/10.1111/gean.12325>
- Fujita, M., & Thisse, J.-F. (2002). *Economic of Agglomeration: Cities, Industrial Location, and Regional Growth*.
- Giersch, Herbert., & Egon-Sohmen-Foundation. (1991). *Towards a Market Economy in Central and Eastern Europe* (H. Giersch, Ed.). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-76572-8>

- Goffette-Nagot, F., & Schmitt, B. (1999). Agglomeration Economies and Spatial Configurations in Rural Areas. *Environment and Planning A: Economy and Space*, 31(7), 1239–1257. <https://doi.org/10.1068/a311239>
- Griffiths, S., & Vaughan, L. (2020). Mapping spatial cultures: contributions of space syntax to research in the urban history of the nineteenth-century city. *Urban History*, 47(3), 488–511. <https://doi.org/10.1017/S0963926820000206>
- Guo, Y., Zhu, J., & Liu, X. (2018). Implication of rural urbanization with place-based entitlement for social inequality in China. *Cities*, 82, 77–85. <https://doi.org/10.1016/j.cities.2018.05.007>
- Hillier, B., & Hanson, J. (1984). *The Social Logic of Space*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511597237>
- Maffenini, L., Moreno-Monroy, A., Schiavina, M., Veneri, P., (2019). *GHSL-OECD Functional Urban Areas. KJ-NA-30001-EN-N (online)*. <https://doi.org/10.2760/67415> (online)
- Mandelbrot, B. B. (2004). *Fractals and Chaos: Vol. C*. Springer New York. <https://doi.org/10.1007/978-1-4757-4017-2>
- McGee, T. G. (2009). *The Spatiality of Urbanization: The Policy Challenges of Mega-Urban and Desakota Regions of Southeast Asia* (161). <https://www.researchgate.net/publication/237521773>
- McGee, T. G., & Shaharudin, I. (2016). *Reimagining the “Peri-Urban” in the Mega-Urban Regions of Southeast Asia* (pp. 499–516). https://doi.org/10.1007/978-3-319-28112-4_30
- Moreno-Monroy, A. I., Schiavina, M., & Veneri, P. (2021). Metropolitan areas in the world. Delineation and population trends. *Journal of Urban Economics*, 125. <https://doi.org/10.1016/j.jue.2020.103242>
- OECD. (2022). *The OECD Metropolitan Database*. <http://stats.oecd.org/Index.aspx?DataSetCode=CITIES>
- Ortman, S. G., Lobo, J., & Smith, M. E. (2020). Cities: Complexity, theory and history. *PLoS ONE*, 15(12 December). <https://doi.org/10.1371/journal.pone.0243621>
- Portugali, J. (2016). What Makes Cities Complex? In *Complexity, Cognition, Urban Planning and Design* (Vol. 1, pp. 3–19). the Springer Proceedings in Complexity book series (SPCOM). https://doi.org/10.1007/978-3-319-32653-5_1
- Portugali, J. (2021). Cities, complexity and beyond. In *Handbook on Cities and Complexity*. Edward Elgar Publishing.
- Serra, M. L. L. A. (2013). *ANATOMY OF AN EMERGING METROPOLITAN TERRITORY: Towards an Integrated Analytical Framework for Metropolitan Morphology* [UNIVERSIDADE DO PORTO]. <http://www.fe.up.pt>
- Shi, W., Batty, M., Goodchild, M., & Li, Q. (2022). The digital transformation of cities. *Urban Informatics*, 1(1). <https://doi.org/10.1007/s44212-022-00005-1>
- Shi, W., Goodchild, M. F., Batty, M., Kwan, M.-P., & Zhang, A. (2021). *Urban Informatics* (W. Shi, M. F. Goodchild, M. Batty, M.-P. Kwan, & A. Zhang, Eds.). Springer Singapore. <https://doi.org/10.1007/978-981-15-8983-6>
- Tu, X., Fu, C., Huang, A., Chen, H., & Ding, X. (2022). DBSCAN Spatial Clustering Analysis of Urban “Production–Living–Ecological” Space Based on POI Data: A Case Study of Central Urban Wuhan, China. *International Journal of Environmental Research and Public Health*, 19(9). <https://doi.org/10.3390/ijerph19095153>
- van Nes, A., & Yamu, C. (2021a). Analysing Linear Spatial Relationships: The Measures of Connectivity, Integration, and Choice. In *Introduction to Space Syntax in Urban Studies* (pp. 35–86). Springer International Publishing. https://doi.org/10.1007/978-3-030-59140-3_2

- van Nes, A., & Yamu, C. (2021b). Introduction to Space Syntax in Urban Studies. In *Introduction to Space Syntax in Urban Studies*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-59140-3>
- Xiao, T., Wan, Y., Jin, R., Qin, J., & Wu, T. (2022). Integrating Gaussian Mixture Dual-Clustering and DBSCAN for Exploring Heterogeneous Characteristics of Urban Spatial Agglomeration Areas. *Remote Sensing*, *14*(22). <https://doi.org/10.3390/rs14225689>
- Yamu, C., & van Nes, A. (2017). An Integrated Modeling Approach Combining Multifractal Urban Planning with a Space Syntax Perspective. *Urban Science*, *1*(4), 37. <https://doi.org/10.3390/urbansci1040037>
- Zheng, W., Du, N., & Wang, X. (2022). Understanding the City-transport System of Urban Agglomeration through Improved Space Syntax Analysis. *International Regional Science Review*, *45*(2), 161–187. <https://doi.org/10.1177/01600176211023879>