

COMPLEXITY-INFORMED ANALYSIS IN DATA-SCARCE CITIES: AN EXPLORATORY METHODOLOGICAL FRAMEWORK

Fabio Bayro Kaiser¹

Abstract

Urban change often unfolds in ways that are difficult to observe and explain, particularly in data-scarce cities. To address this, this paper presents a complexity-informed exploratory framework for spatial analysis using globally consistent Earth observation data. Here, data scarcity refers not to absence but to limitations in validation, temporal continuity, resolution, and proxy alignment. Guided by non-linearity, emergence, and distributed agency, the framework compares multiple indicators: built-up extent, density, vegetation, and land consumption across time and space, and examines their alignment and divergence. It applies explicit stopping rules to bound interpretation and treats ambiguity and non-detection as valid analytical outcomes; illustrated through an application to La Paz–El Alto. The framework establishes a disciplined basis for interpreting spatial evidence and supporting planning-relevant orientation under constrained observation.

Keywords:

Urban Complexity; Data-Scarce Cities; Exploratory Spatial Diagnostics; Earth Observation Indicators; Interpretive Planning Analysis.

¹ Chair and Institute for Urban Design, Faculty of Architecture, RWTH Aachen University.
E-mail: bayrokaiser@staedtebau.rwth-aachen.de. ORCID: <https://orcid.org/0000-0002-4839-5991>.

1. Introduction — Scarcity as constraint and opportunity

In many cities, urban change unfolds through decentralised decisions, uneven feedback, and non-linear dynamics that resist linear explanation. Complexity-inspired thinking has, therefore, become influential in urban studies, and provided conceptual resources for describing emergence, path dependence, and distributed agency in urban systems (Allen, 1997; Batty, 2005; Bettencourt, 2021; Portugali, 2011, 2016; Moroni, 2015). These ideas have informed debates on urban morphology, polycentricity, and self-organisation, and have shaped both theoretical reflection and empirical work (Moroni and Cozzolino, 2019; Pumain, 2018, 2024; Taubenböck et al., 2017; White and Engelen, 1997). Despite this influence, complexity-informed spatial analysis remains difficult to apply in practice, particularly in cities where empirical observation is constrained (Batty and Marshall 2012; Byrne 2003; Portugali, 2011; Pumain 2018).

This difficulty is often framed as a missing-data problem. However, such a framing is increasingly misleading. Over the past decade, Earth observation (EO) technologies have expanded spatial coverage and temporal frequency; enabling consistent observation of built-up form, land cover, and selected structural attributes across much of the globe (Brown et al., 2022; Taubenböck, 2025; Taubenböck et al., 2012; Zhu et al., 2022). Data scarcity today rarely denotes absence. Rather, it refers to an observational condition that is shaped by limited temporal depth, uneven validation, coarse thematic resolution, and imperfect alignment between spatial proxies and underlying urban processes (Kuffer et al., 2020; Taubenböck et al., 2017). These constraints fundamentally shape what spatial analysis can support, even when data is formally available.

Crucially, this form of scarcity is not a constraint but an analytical opportunity. Reliance on globally available EO-derived data makes large-scale and comparative urban analysis feasible (Angel et al., 2016; Taubenböck et al., 2014). Approaches that rely on dense, locally validated datasets typically extend only to a small number of cities. It follows, that scarcity enforces a shared observational baseline and with it, a degree of comparability that would otherwise be unattainable. At the same time, it demands analytical parsimony. Under constrained conditions, extensive indicator sets, complex models, or layered assumptions quickly exceed the (given) data's capacity (Angel et al., 2021). Scarcity shifts the analytical emphasis from completeness to sufficiency: what can be learned responsibly and transparently from the available data.

This shift exposes a persistent tension in complexity-informed empirical work. Non-linearity, asynchronous change, and distributed agency imply that urban dynamics cannot be inferred from a single indicator or a static snapshot (Batty, 2013; Moroni, 2015; Pumain, 2018). Nevertheless, much spatial analysis continues to rely on simplified metrics or aggregate trends; thereby often smoothing over indicator disagreement or temporal instability. Where multiple indicators are combined, interpretive decisions are frequently implicit, and uncertainty is treated as noise to be reduced rather than as an analytical outcome. Although this tension between conceptual ambition and analytical feasibility has been widely acknowledged (Byrne, 2003; Batty and Marshall, 2012), it remains unresolved because complexity is typically applied either at a high conceptual level or within data-rich modelling environments.

For planning-related research, this gap is consequential. Planning routinely engages with partial, provisional, and uneven evidence particularly in rapidly urbanising or informally governed contexts (Rauws, 2017; Colander and Kupers, 2014). In such settings, the primary analytical need is not explanation or prediction, but orientation: problem framing, situational differentiation, and early signal interpretation that can inform reflection and prioritisation without overstating what the evidence supports. Yet spatial analytical outputs are often read diagnostically by default, even when observational conditions do not warrant causal attribution or forecasting. The absence of explicit stopping rules encourages overinterpretation and obscures where inference should end.

This paper addresses this problem by proposing a complexity-informed, exploratory methodological framework for spatial analysis in data-scarce cities. Complexity is used as an analytical influence that disciplines interpretation through non-linearity, emergence, and distributed agency (Portugali, 2006). This entails comparing multiple indicators over time, examining their alignment and divergence, and avoiding linear extrapolation or single-indicator inference. Data scarcity is explicitly treated as an observational condition that shapes indicator selection, temporal comparisons, and interpretive scope. The framework relies on EO-derived data because their consistent spatial and temporal observation regimes are essential for interpreting longitudinal signals. In contrast, crowdsourced or administrative datasets often reflect contributor behaviour (e.g. uneven mapping activity or reporting practices) or institutional processes (e.g. changes in data collection

or classification), which introduce temporal inconsistency and complicate such analysis. Methodologically, the framework emphasises analytical economy: a limited number of interpretable indicators, simple temporal standardisation, and a small set of signal configurations. These choices prioritise legibility, transparency, and the speed of iteration – understood here as safeguards against overinterpretation amid uncertainty.

An instrumental application to the La Paz–El Alto metropolitan region illustrates how the framework behaves under constrained observational conditions (Bayro Kaiser, 2024). The case is suitable because it combines rapid urban change, topographic constraints, and limited validated data. It is used to demonstrate the framework’s methodological performance in structuring interpretation, not to explain urban development trajectories or to claim representativeness.

The paper is structured as follows. Section 2 situates complexity as an interpretive resource and clarifies its implications for reasoning under constraint. Section 3 conceptualises data scarcity as an observational condition and discusses why expanded EO coverage does not automatically translate into analytical adequacy. Section 4 introduces the exploratory methodological framework, and in so doing details its analytical logic, indicator selection, and stopping rules. Section 5 presents the instrumental case application and examines how the framework supports problem framing, situational differentiation, and early signal interpretation. Section 6 discusses what generalises from this approach and where its limits lie. Thereafter Section 7 concludes the paper with suggestions for directions of future research.

2. Complexity as analytical inspiration

Complexity thinking has become a central reference in urban research because it offers a coherent way by which to conceptualise cities shaped by decentralised interactions, feedbacks, and non-linear change (Portugali, 2012). From early work on self-organisation and urban evolution (Jacobs, 1961) to later formulations of cities as complex adaptive systems (Portugali, 2011), such literature has challenged linear and equilibrium-based accounts of urban dynamics (Allen, 1997; Allen and Sanglier, 1981; Batty, 2005, 2013; Bettencourt, 2021). Its conceptual influence is now well established across debates on urban morphology, scaling, and spatial organisation (Bettencourt, 2013). What remains challenging, however, is how these ideas can inform empirical spatial analysis when observation itself is constrained. Without a disciplined reframing, complexity-inspired analysis either remains conceptually abstract or slips into empirical claims that exceed what available evidence can support. Given this, this section positions complexity as an analytical inspiration and clarifies how selected concepts constrain the interpretation of spatial indicators and their temporal behaviour under observational limits.

A first concept with direct analytical relevance is non-linearity. In complex systems, change rarely proceeds smoothly; extended periods of apparent stability may be followed by rapid reconfiguration, while short-lived accelerations may dissipate without structural consequence (Batty, 2013; Scheffer et al., 2012). For spatial analysis, this undermines the plausibility of single-indicator narratives or short temporal windows as reliable evidence of structural change. Empirical studies of urban expansion and densification have repeatedly shown the presence of alternating phases of acceleration, stagnation, and redirection; shaped by the timing of infrastructure, regulatory shifts, or access constraints rather than by monotonic growth (Angel, 2012; Angel et al., 2021; Taubenböck et al., 2014; Seto et al., 2012). Given data scarcity, non-linearity imposes a concrete analytical requirement: urban change must be examined using multiple indicators over time, while linear extrapolation and premature identification of transitions must be avoided.

A second concept is emergence, which emphasises that higher-level urban patterns arise from local interactions, yet cannot be directly inferred from or reduced to them (Jacobs, 1961; Allen and Sanglier, 1981; Portugali, 2004; Moroni and Cozzolino, 2019). Empirically, this implies that no single indicator can be expected to capture such higher-level patterns in isolation. Structural change may become partially visible through alignment across signals, such as concurrent shifts in built-up extent, patch structure, and land consumption (Taubenböck et al., 2014; Gerten, Fina and Rusche, 2019; Zhu et al., 2022). However, such alignment remains indicative rather than conclusive. It follows, that emergence directs attention to synchrony and divergence across indicators as potential expressions of higher-level pattern formation. Importantly, this does not justify holistic claims about system behaviour. Rather, it frames interpretation around signal configurations that are consistent with but do not confirm emergent spatial patterns. This means that under constrained observation uncertainty is not residual error but an expected feature of complex urban systems that must be reported rather than resolved.

A third concept, distributed agency, further conditions analytical expectations. Urban change is produced through the cumulative actions of heterogeneous actors operating across scales – often with weak coordination and conflicting objectives (Moroni, 2015; Slaev et al., 2022). As a result, spatial change is rarely uniform (Reicher, Ostos-Prieto and Bayro Kaiser, 2025). Densification, expansion, and decline may unfold simultaneously in different parts of the same urban region, producing contrasting signals when aggregated (Hall and Pain, 2006; Meijers, 2008; Reicher et al., 2011). In data-scarce settings, distributed agency implies that divergence should be anticipated rather than treated as noise to be smoothed out. Spatial unevenness becomes analytically relevant only if interpretation remains bounded by observational limits.

These three concepts—non-linearity, emergence, and distributed agency—have been deliberately selected for this framework because they have immediate and direct implications for how spatial indicators behave and how they should be interpreted under constrained observation. Broader notions such as resilience, adaptation, or sustainability have been excluded from the framework because they often presuppose evaluative or normative criteria that cannot be empirically supported given data scarcity. Crucially, none of the selected concepts requires treating complexity as a testable theory. Empirical validation of complexity models would require levels of causal attribution, temporal depth, and data completeness that are rarely available outside data-rich settings (Byrne, 2003; Pumain, 2018).

Distinguishing inspiration from operationalisation is, therefore, central to the framework’s methodological logic. Much of the gap which exists between complexity theory and empirical practice stems from attempts to translate abstract concepts directly into indicators or models; a process that has often resulted in conceptual dilution or false precision (Batty and Marshall, 2012). A heuristic use of complexity accepts that non-linearity or emergence cannot be measured directly. Their analytical value lies in disciplining interpretation: shaping expectations about indicator behaviour, encouraging comparison across indicators and time, and legitimising ambiguity and non-detection as outcomes. At no point are these concepts treated as measurable properties or validated mechanisms.

Taken together, these concepts motivate a specific analytical stance. First, urban change should be approached through multiple indicators observed over time; second, alignment and divergence should be examined explicitly; and third, instability or non-detection should be treated as informative rather than as methodological failure. This stance neither resolves the challenges of observing complex urban systems, nor advances complexity theory itself. Rather, it provides a disciplined conceptual foundation for exploratory spatial analysis under constraint. Given this, the next section turns to data scarcity as an analytical condition and examines how observational limits shape indicator behaviour and why expanded EO coverage does not automatically translate into analytical adequacy.

3. Data scarcity as an observational condition

The term data-scarce is widely used in urban research, yet it is often interpreted narrowly as a lack of available data. In practice, many cities are now covered by an expanding range of spatial datasets, including global EO products, volunteered geographic information, and locally produced administrative or survey data. Data scarcity, therefore, does not primarily refer to the absence of spatial data, but to an observational condition in which limitations in resolution, temporal continuity, validation, and comparability constrain what can be inferred from multi-source urban evidence (Figure 1). This perspective shifts attention from data availability to the conditions under which spatial information can be interpreted, and frames scarcity as a defining analytical constraint rather than a residual deficiency.

Recent advances in EO have substantially expanded the empirical basis for urban analysis. Globally consistent datasets now provide information on built-up extent, land cover, vegetation dynamics, and selected structural proxies, often with annual or higher temporal frequency (Taubenböck et al., 2012; Brown et al., 2022; Zhu et al., 2022). These developments have transformed the feasibility of spatial analysis in contexts where systematic monitoring was previously limited (Taubenböck, 2025). EO-derived data, therefore, functions as a primary longitudinal observation regime for exploratory urban analysis. At the same time, expanded coverage does not eliminate observational constraints; instead, it reshapes them, as summarised schematically in Figure 1.

A first constraint concerns spatial resolution and thematic specificity. Global EO products are optimised for consistency across regions and sensors, which limits their capacity to capture fine-grained or functionally

differentiated urban processes. Incremental densification, informal infill, or small-scale vertical extension may be weakly detected or inconsistently represented across time. Comparable issues arise with volunteered geographic data, where coverage and detail often reflect contributor activity rather than uniform urban conditions. These limitations do not invalidate such datasets, but they do constrain the types of questions that they can support and the degree of confidence with which indicator behaviour can be interpreted (Kuffer et al., 2020; Taubenböck et al., 2017). Analytically, this shifts emphasis from absolute values to relative change, and from isolated indicators to configurations of signals.

Temporal continuity constitutes a second constraint. Although many spatial datasets provide nominal time series, continuity is frequently disrupted by sensor changes, algorithm updates, or shifts in data-processing practices. Apparent changes in indicator trajectories may, therefore, reflect methodological adjustments rather than urban dynamics. This complicates short-term trend analysis and renders the identification of structural change sensitive to processing choices, particularly in the absence of independent benchmarks (Seto et al., 2012). Under such conditions, ambiguity and instability are expected analytical outcomes rather than signs of methodological failure.

A third constraint relates to validation capacity. In many urban contexts, locally produced reference data such as recent censuses, building registries, or land-use surveys are incomplete, outdated, or inconsistent. This limits the calibration and verification of EO-derived indicators beyond plausibility checks or indirect comparison (Stevens et al., 2015). As a result, spatial indicators can often support descriptive and comparative analysis, but not precise quantification or causal attribution.

Finally, comparability across space and data modalities introduces further complexity. Indicators derived from EO, volunteered data, and administrative sources may align imperfectly across space and time due to differing definitions, coverage, and update cycles. Similar indicator values can, therefore, reflect different underlying conditions, while similar processes may produce divergent signals across contexts (Angel, 2012; Gerten et al., 2019). Without explicit attention to these misalignments, interpretation risks conflating measurement artefacts with substantive change.

Understanding data scarcity in this way clarifies the scope of the proposed framework. It does not aim to compensate for missing data, resolve measurement uncertainty, or establish causal explanations. Instead, it treats observational constraints as a defining analytical condition and focuses on how spatial indicators can be interpreted cautiously and coherently within those limits. This framing aligns with planning-related analysis in which engagement with partial and heterogeneous evidence is the norm. On this basis, the following section introduces an exploratory methodological framework that has been designed explicitly for such observational regimes.

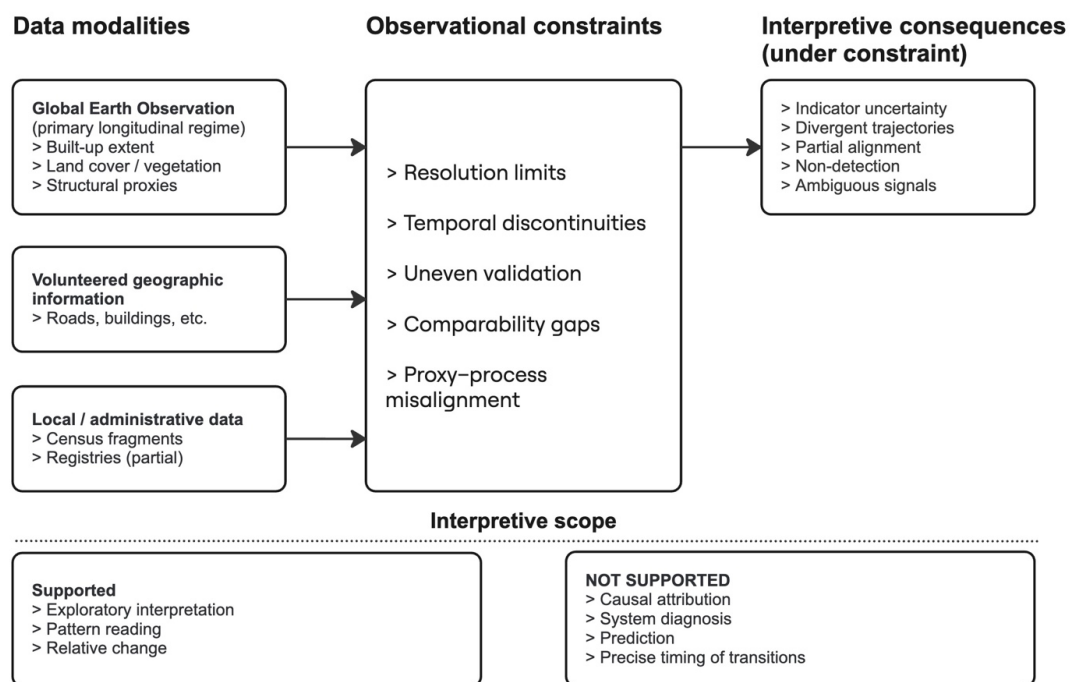


Figure 1: Data scarcity as an observational condition in multi-source urban analysis. Source: own representation.

4. An Exploratory Methodological Framework

The methodological framework proposed here is designed to support complexity-informed spatial orientation under explicit observational constraints. Its purpose is not to explain urban change, detect transitions, or validate complexity theory. Instead, it provides a disciplined structure for organising, comparing, and interpreting EO-derived spatial signals when inference is limited by resolution, temporal continuity, or validation capacity. The core idea of the framework is to formalise interpretive restraint as a methodological design feature, rather than treating uncertainty and non-detection as residual shortcomings to be minimised.

A substantial body of research has demonstrated the analytical potential of EO-derived indicators for describing urban expansion, densification, fragmentation, and structural change across scales (Taubenböck et al., 2012; Angel et al., 2016; Seto et al., 2012; Sirko et al., 2023; Zhu et al., 2022). Comparative approaches have further shown that consistent indicator sets can reveal broad morphological regularities across cities (Demuzere et al., 2022; Gerten et al., 2019; Fleischmann and Arribas-Bel, 2022). At the same time, this literature has repeatedly exposed persistent challenges: sensitivity to scale and delineation, temporal instability, proxy-process misalignment, and limited opportunities for validation, particularly in data-scarce contexts. While recent contributions have increasingly acknowledged uncertainty (Kuffer et al., 2020; Demuzere et al., 2022), indicators are still commonly treated as inputs to be refined or aggregated, often with the implicit aim of improving explanatory or predictive power.

The framework presented here takes a different stance. It treats observational constraints as a defining analytical condition and examines how spatial evidence can be interpreted coherently without exceeding what the data supports. Its contribution lies not in proposing new indicators or techniques, but in structuring interpretation through parsimony, explicit limits, and transparent decision points.

4.1 Purpose and positioning

The framework is positioned as an exploratory analytical scaffold. Complexity concepts inform expectations about indicator behaviour such as non-linearity, uneven spatial expression, and partial alignment. Data scarcity is addressed not by compensating for missing information, but by shaping how evidence is selected, compared, and interpreted. The intended output is orientation: a structured account of how spatial signals align, diverge, or remain ambiguous over time.

This positioning differentiates the framework from optimisation-oriented analytics, simulation-based modelling, and monitoring systems that are aimed at diagnosis or evaluation. For planning-related analysis, this form of orientation is particularly relevant in early analytical phases, when the task is to frame problems, spatially differentiate situations, and identify areas that require closer investigation, rather than supporting decisions or prescribing interventions.

4.2 Framework logic and analytical phases

The framework unfolds through four analytical phases which are understood as a logic of interpretation rather than as a data-processing pipeline. Figure 2 visualises this logic visually, emphasising decision points and stopping conditions.

The first phase defines the analytical context. This involves delineating the spatial extent, identifying a temporally stable observation window, and characterising dominant data constraints. Context definition establishes expectations regarding which indicator behaviours are plausible and which interpretations are untenable from the outset.

The second phase concerns inputs and uncertainty. Indicator families are fixed at the framework level, while specific indicators are selected, justified, or excluded at the case level based on availability, temporal stability, and plausibility under local conditions. The framework distinguishes between a broad data universe and the subset of datasets that actively structure inference in a given application. While multiple EO-derived and ancillary datasets may be available in principle, their inclusion is conditional on temporal coverage, proxy stability, and alignment with the observational regime. Exclusion at the case level is treated as a methodological

decision that is governed by the framework, not as a limitation or data deficit. Vertical signals are treated as conditional and may be excluded where EO-derived height proxies are unreliable or inconsistent across time.

The third phase focuses on relational diagnostics. Indicators are examined through their temporal trajectories and cross-family behaviour, rather than as isolated measures. Attention is directed to alignment, divergence, volatility, and lag across indicator families. Iteration is possible at this stage through indicator selection or exclusion but remains bounded and pre-interpretive.

The fourth phase introduces an explicit interpretive decision. As shown in Figure 2, signal coherence across indicator families may support a cautious analytical orientation, whereas persistent divergence or instability can lead to non-detection as a valid outcome. Crucially, both paths terminate at explicit stopping points, reinforcing the idea that inference does not extend beyond what the evidence can sustain.

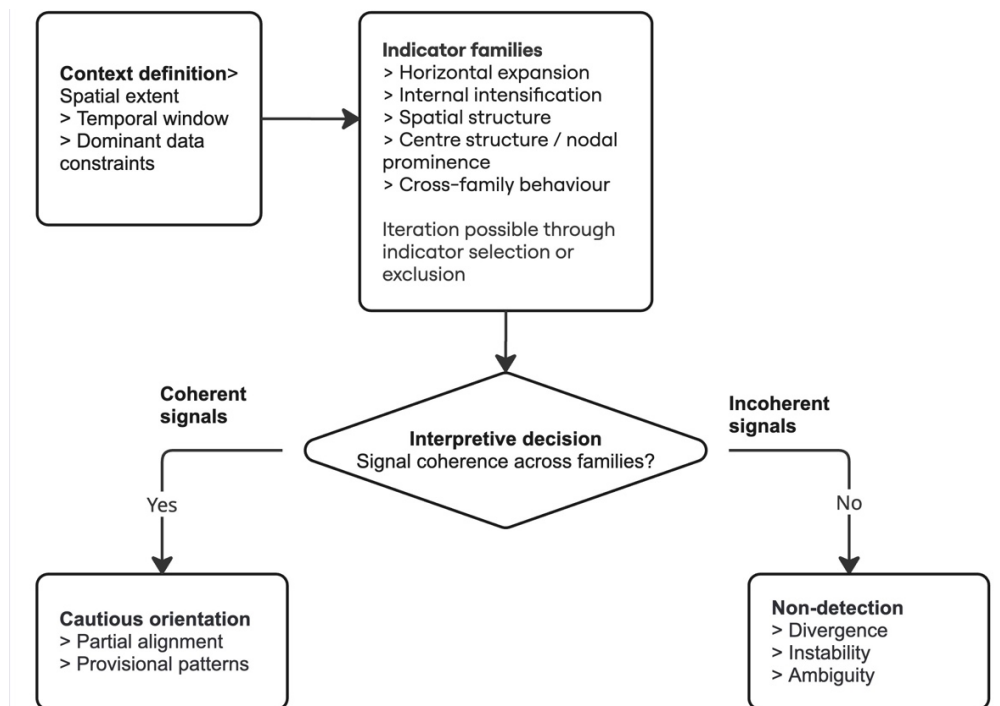


Figure 2: Exploratory methodological framework for spatial orientation under data scarcity. Source: Author's own representation.

4.3 Indicator families and signal configurations

The framework employs five indicator families. These are summarised in Table 1, and together capture analytically distinct dimensions of spatial change. These families are deliberately limited in number. They represent dimensions that: first, can be observed consistently using EO-derived data; second, respond differently under nonlinearity and distributed agency; and third, minimise redundancy under data scarcity. Expanding the framework with additional families such as functional land-use or socio-economic proxies would increase interpretive ambiguity without improving inferential capacity under the stated constraints.

Horizontal expansion indicators describe the outward growth of the built-up footprint and are widely used in global urban studies (Angel, 2012; Seto et al., 2012; Taubenböck et al., 2014). Within the framework, they are interpreted as signals which are consistent with decentralised expansion dynamics, rather than as measures of growth drivers. Internal intensification indicators capture changes within the (given) existing footprint, including density proxies and, where reliable, vertical signals. Their analytical value lies in highlighting reallocation within constrained space, not in identifying saturation thresholds.

Spatial structure indicators, such as dispersion or fragmentation metrics, reflect spatial unevenness and coherence. Prior work has shown that they are sensitive to scale and resolution (Dovey and Pafka, 2014), which limits their diagnostic utility. Within the framework, they are interpreted as expressions of uneven signal

distribution consistent with distributed agency. Centre-structure indicators address relative dominance and nodal patterns, and thus builds on research into morphological polycentricity (Meijers, 2008; Hall and Pain, 2006), while avoiding functional interpretations that would require flow or institutional data.

Analytical emphasis emerges from cross-family behaviour. Alignment across families may suggest coherent spatial reorganisation, while divergence or volatility signals ambiguity. For example, when built-up extent, land consumption, and NDVI shift in a consistent direction within the same area, this may indicate outward urban expansion; whereas if built-up extent increases while density and vegetation remain unstable or move in opposing directions, the signal remains ambiguous. In both cases, interpretation remains provisional and bounded.

Indicator family	Spatial dimension captured	Complexity principles engaged	What signal alignment may indicate	What cannot be inferred
Horizontal expansion	Outward change of built footprint	Non-linearity; distributed agency	Spatial extension consistent with decentralised growth dynamics	Growth drivers; permanence; causal mechanisms
Internal intensification	Reallocation within stable footprint (horizontal or vertical proxies)	Non-linearity; emergence	Increasing use of existing urban space	Saturation thresholds; adaptive strategies; socio-economic causes
Spatial structure	Distribution and continuity of built form	Distributed agency; emergence	Increasing or uneven dispersion of development	Functional fragmentation; governance dynamics
Centre structure / nodal prominence	Relative dominance across space	Emergence; feedback processes	Redistribution of spatial prominence	Functional polycentricity; flow reorganisation
Cross-family behaviour	Relational alignment across families	All principles	Coherence, divergence, volatility, or non-detection	System states; transitions; predictive claims

Table 1: Indicator families and interpretive scope. Source: Author's own representation.

4.4 Uncertainty, non-detection, and stopping rules

Uncertainty is not treated as residual but as a produced analytical outcome. The framework formalises conditions under which interpretation must stop. Interpretation is halted when indicator trajectories diverge persistently across families; results hinge on a single unstable indicator, or signal configurations change materially under minor parameter variations. In such cases, synthesis is withheld rather than forced.

Non-detection—the absence of stable or interpretable signal configurations—is considered a valid result. It indicates that, under the given observational conditions, spatial change cannot be confidently interpreted. By embedding stopping rules directly into the analytical logic, the framework clarifies not only what can be inferred but also when inference must end.

In summary, the framework builds on a substantial body of EO-based urban science while addressing its persistent interpretive gaps. It should be evaluated not by the richness of its outputs, but by the transparency with which it exposes uncertainty, limits, and interpretive choices. Table 1 provides the analytical matrix which links indicator families, complexity-informed expectations, and interpretive scope, while Figure 2 visualises how this logic is enacted in practice. The following section applies the framework instrumentally to the La Paz–El Alto metropolitan region, and demonstrates how it structures interpretation under constraint without exceeding what the data supports.

5. Instrumental Case Application: La Paz–El Alto

5.1 Instrumental positioning and observational regime

The La Paz–El Alto metropolitan region (GAMLP, 2017) is used instrumentally to demonstrate how the exploratory framework performs under constrained observational conditions (Figure 3). The case is not presented as being representative of a broader urban category. Neither is it analysed to explain development trajectories. Rather, its analytical value lies in integrating rapid demographic change, strong topographic constraints, and uneven institutional capacity within a setting in which longitudinal EO data constitutes the primary observational regime.

All indicators are derived from globally consistent EO-based datasets processed in Google Earth Engine and spatially aggregated in QGIS using a fixed hexagonal grid (Table 3, Appendix). The built areas for the period 2015–2025 were computed using data from Dynamic World (Brown et al., 2022). The population data were obtained from the WorldPop database (Bondarenko et al., 2025). The NDVI data were obtained from the Landsat-8 and Sentinel-2 missions. Temporal standardisation and Δz computation were performed in Excel, following the logic outlined in Section 4. The workflow, therefore, makes the analytical separation between computational processing, spatial aggregation, and relational interpretation explicit. The Appendix documents detailed technical steps; here, the focus remains on interpretive performance.

Four core indicators structure inference at the hex level:

- built-up extent (patch-derived built area per hex),
- built-up density proxy (population per built area),
- land consumption proxy (built area per capita),
- and mean NDVI per hex.

These indicators correspond directly to the families defined in Table 1 and are intentionally limited in number. Their role is not to capture the full complexity of urban dynamics, but to provide a parsimonious and interpretable basis for cross-family comparison.



Figure 3: The La Paz–El Alto metropolitan region. Source: Author, 2024.

5.2 Orientation logic and signal configurations

Orientation is derived from temporally standardised change. For each hexagon and each indicator, a z-score is computed per year, followed by Δz across a defined time window. A moderate deviation threshold ($\tau = 0.25$) is applied symmetrically to identify meaningful departures from baseline behaviour. The threshold was selected on the basis that, after sensitivity testing, it was sufficiently strict to avoid noise amplification while remaining responsive to moderate shifts.

The active signal count, denoted as k , measures the number of indicator families in which $|\Delta z| \geq \tau$. This metric operationalises cross-family reinforcement:

- $k = 0$ indicates no coherent deviation across families.
- $k = 1$ suggests isolated change without relational reinforcement.
- $k \geq 2$ indicates partial or multi-family alignment.
- Mixed sign deviations across families are treated as divergent configurations.

Importantly, k does not measure the magnitude of change, but relational coherence. This distinction is central to the framework's logic: the presence or absence of cross-family alignment is analytically more relevant than large movement in a single metric.

Table 2 summarises the orientation configurations used in the case application. The expanded matrix clarifies required indicator combinations and spatial expectations under distributed agency.

Configuration	Required Δz Combination	k	Spatial expectation under distributed agency	Interpretive orientation
Stable / low signal	No indicator exceeds $\pm\tau$	0	Micro-adjustments without synchronisation	No coherent reorganisation detectable
Single-family deviation	One indicator exceeds $\pm\tau$	1	Localised response or measurement sensitivity	Insufficient for cross-dimensional orientation
Partial alignment	Two indicators exceed $\pm\tau$ in same direction	2	Local reinforcement across dimensions	Emerging coherence
Multi-family alignment	≥ 3 indicators exceed $\pm\tau$ in same direction	≥ 3	Coordinated spatial shift	Stronger cross-dimensional signal
Divergent configuration	Indicators exceed $\pm\tau$ in opposing directions	≥ 2	Uneven change consistent with distributed agency	Ambiguous or conflicting orientation

Table 2: Orientation configurations and relational signal logic

5.3 Window comparison and temporal behaviour

To assess temporal sensitivity, Δz was computed across multiple window lengths. Three observations are analytically relevant.

First, the one-year window exhibited high volatility. Numerous hexagons registered isolated deviations ($k = 1$), while cross-family reinforcement remained rare. This behaviour was consistent with the non-linearity discussed in Section 2: short temporal windows amplify episodic fluctuation and sensor-level variability.

Second, the two-year window reduced volatility and increased partial alignment. A modest rise in $k \geq 2$ configurations suggests that certain areas exhibit relational reinforcement when short-term oscillations are smoothed. This window balances responsiveness and stability and was, therefore, adopted as the primary orientation window.

Third, extending to three-year windows increased smoothing but reduced sensitivity to more recent shifts. While coherence became more stable, interpretive responsiveness declined. Given the framework's exploratory purpose, the two-year window provided the most informative compromise.

This comparison demonstrates that signal coherence is not absolute but contingent on temporal framing. Rather than arbitrarily selecting a window, the framework treats window selection as part of the analytical logic and explicitly reports its implications.

5.4 Problem framing and situational differentiation

Across the two-year window (2023–2025), most hexagons fell into the stable/low signal category ($k = 0$) (Figure 4). Interpreted superficially, this could be read as urban stagnation. Within the framework's logic, however, stability is defined more precisely as the absence of cross-family reinforcement.

This macro-stability suggests three analytically relevant conditions:

- Spatial adjustments occurred in a distributed manner without synchronisation across dimensions.
- No dominant reorganisation pattern emerged at the metropolitan scale.
- Indicator behaviour remains within baseline variability under the defined observational regime.

In this sense, stability is not the absence of change but the absence of coherent cross-dimensional alignment. It reflects distributed micro-adjustments consistent with the principle of distributed agency.

Situational differentiation became visible in a minority of hexagons with $k \geq 2$. These areas clustered spatially along selected corridors and peripheral zones (Figure 5). In these locations, built-up expansion aligned with declining NDVI and increasing land consumption—indicating reinforced outward development signals. Other clusters exhibited intensified built-up areas alongside reduced per capita land consumption; consistent with selective internal intensification.

The spatial pattern that emerged was, therefore, not a city-wide transition but a differentiated mosaic: stable background conditions punctuated by localised zones of relational reinforcement.

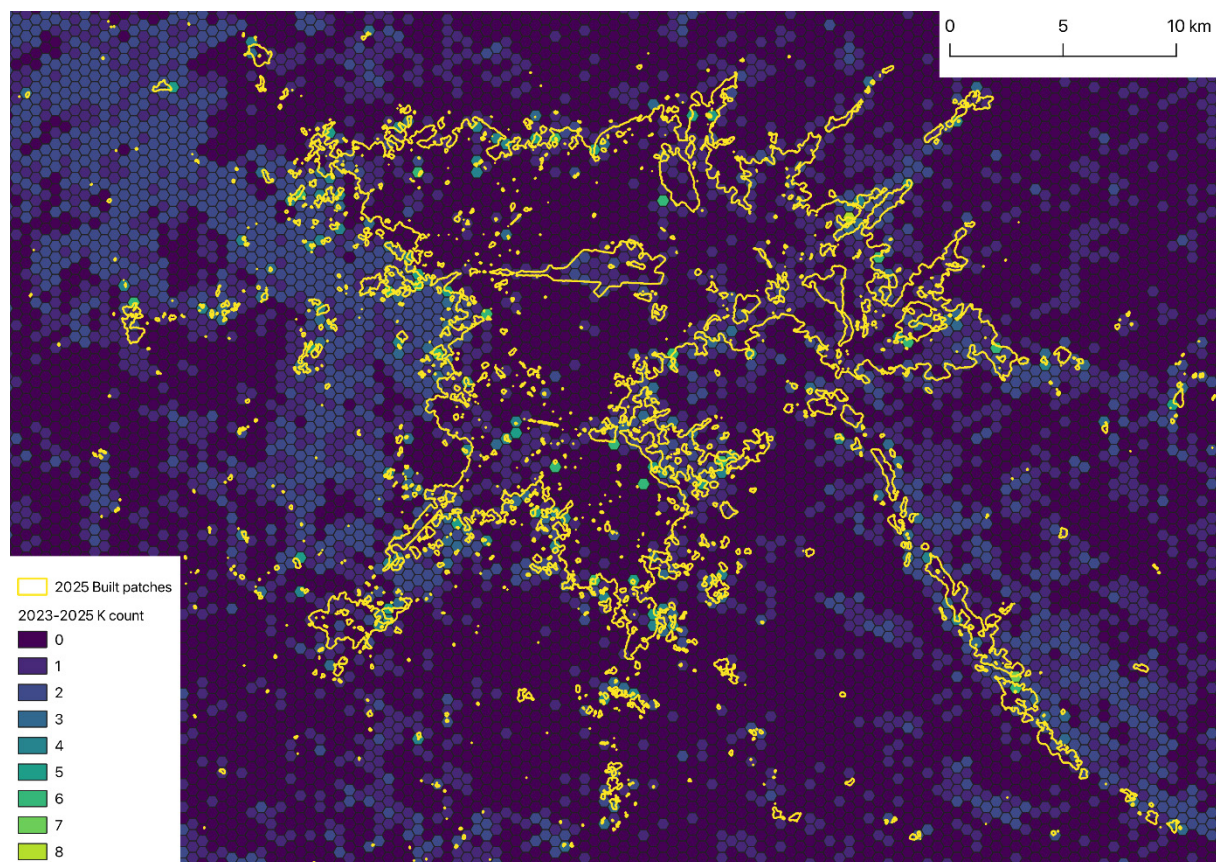


Figure 4: Active signal intensity (k). Hexagon map of the count of indicator deviations ($|\Delta z| \geq \tau$), indicating where signals reinforce or remain stable. Source: Own representation.

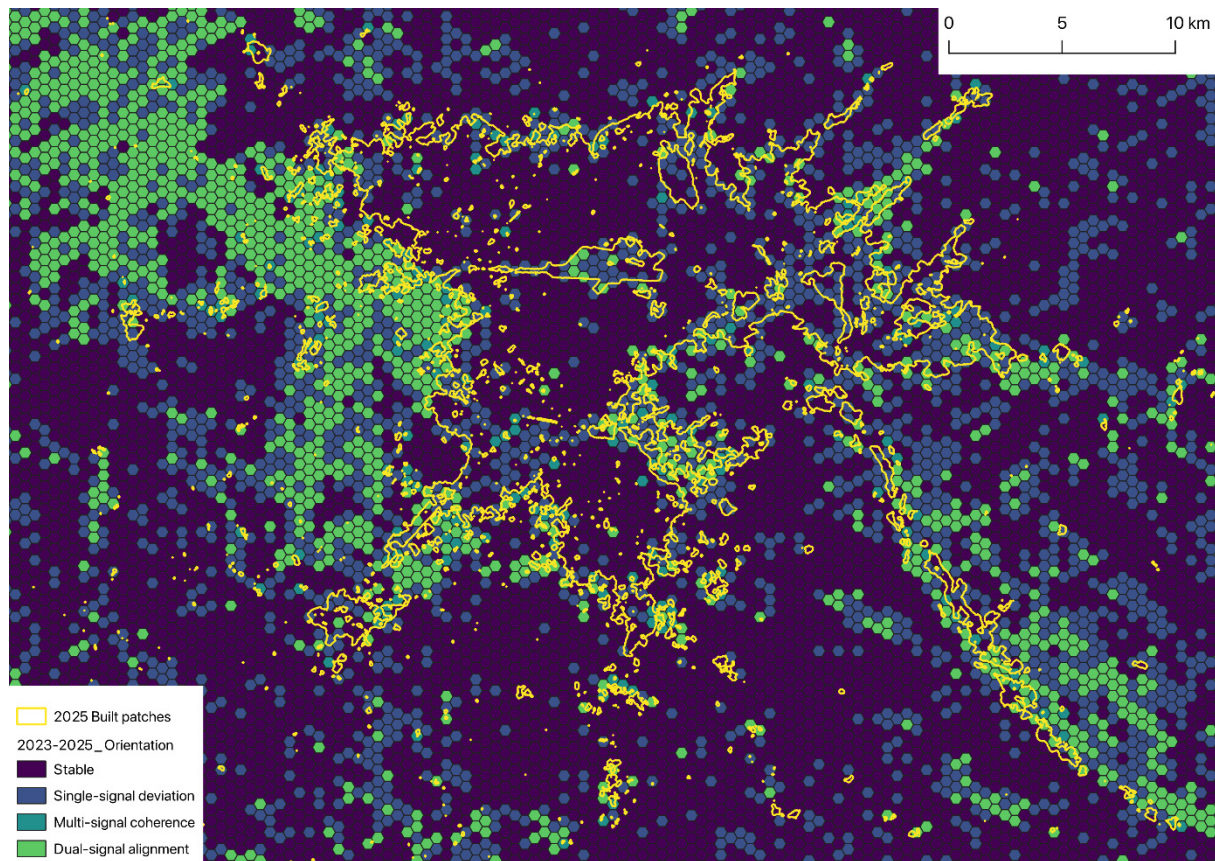


Figure 5: Situational differentiation under constraint. Hexagon map of orientation configurations showing where signals align, diverge, or remain absent under shared observational limits. Source: Own representation.

5.5 Early signal interpretation and stopping points

The framework explicitly prevents escalation from signal coherence to structural diagnosis. Even in hexagons with $k \geq 2$, the interpretation remains conditional: reinforced signals indicate cross-dimensional alignment, not causal mechanisms or long-term trajectories.

- Interpretation stops when:
- signals are driven by a single indicator,
- divergence persists across families,
- results vary materially across temporal windows.

In this application, the predominance of $k = 0$ and $k = 1$ categories indicated that metropolitan-scale reorganisation cannot be inferred with confidence. Localised reinforcement identified areas warranting closer qualitative or data-rich investigation, but the evidence did not support a systemic transition.

This outcome illustrates the framework's central claim: orientation can be generated under constraint without overstepping interpretive limits. Accordingly, stability, divergence, and non-detection are analytically productive findings.

5.6 Implications for complexity-informed orientation

The La Paz–El Alto application demonstrates how complexity principles discipline interpretation. First, non-linearity is evident in the sensitivity of short temporal windows, in which one-year intervals yield volatile and weakly aligned signals. This reinforces the analytical requirement to examine change across multiple indicators and longer time spans, rather than relying on short-term trends. Second, distributed agency is reflected in the

pronounced spatial unevenness of signal configurations, with stable areas coexisting alongside localised zones of reinforcement. This indicates that spatial divergence should be expected and interpreted as an intrinsic feature of urban change rather than as noise. Third, emergence is approached through cross-family alignment, where coherence across indicators suggests higher-level pattern formation but does not confirm structural transformation. Alignment is, therefore, interpreted as indicative rather than conclusive.

The dominant pattern is one of distributed micro-adjustment rather than synchronised reorganisation. This is not a minor finding. It clarifies that, under the current observational regime, the metropolitan system exhibits limited cross-dimensional reinforcement. Such clarity supports planning-relevant orientation: it identifies zones of reinforcement without overstating their significance and makes explicit where inference must stop.

The case, therefore, confirms that complexity-informed spatial analysis can produce structured, legible orientation under data scarcity, provided that parsimony, relational reading, and stopping rules remain integral to the methodological design.

6. Discussion

This study sought to clarify the contributions of complexity-informed spatial analysis when empirical observation is limited. The results indicate that the principal value of the proposed framework lies in structuring interpretation rather than in producing definitive diagnoses. Across the La Paz–El Alto application, the framework consistently delivered three forms of orientation that are directly relevant for analytical and planning contexts: first, problem framing; second, situational differentiation; and third, early signal interpretation. These outputs align with existent calls in urban analytics to prioritise interpretable evidence over model sophistication when working with globally available EO data.

These three forms of orientation can be interpreted through the complexity principles outlined in Section 2 as analytical constraints. Non-linearity, distributed agency, and emergence do not appear as directly observable properties, but as conditions that shape how spatial signals are read. They require multi-indicator comparison, legitimising spatial divergence, and limiting interpretations of alignment to indicative rather than conclusive evidence. In this sense, the case does not demonstrate complexity as a measurable system property, but rather shows how complexity-informed expectations structure interpretation under constrained observation.

Problem framing resulted from the disciplined comparison of indicator families under explicit observational constraints. By treating data scarcity as a condition of inference, the analysis clarified which spatial questions could be examined credibly and which could not. This addresses a persistent gap in EO-based urban studies, where large-scale datasets enable descriptive measurement but rarely specify their interpretive limits. The framework, therefore, contributes a procedural advance: it makes the boundary between evidence and inference visible, a distinction that is often implicit, rather than explicit, in spatial indicator research.

Situational differentiation constituted the second generalisable contribution. The hexagonal aggregation and cross-family comparison revealed spatially uneven signal configurations within the same metropolitan system. Such divergence is consistent with distributed agency and multi-scalar urban dynamics documented in morphological and remote-sensing studies (Angel et al. 2021; Taubenböck et al., 2017; Zhu et al., 2022). Importantly, differentiation did not depend on identifying statistically dominant trends. Instead, it arose from relational contrasts between indicators, and thus demonstrates that spatial heterogeneity can itself serve as an analytical output. This finding supports recent comparative work which shows that consistent indicator sets can reveal structural variation across cities without requiring explanatory modelling (Fleischmann and Arribas-Bel, 2022; Gerten et al., 2019).

The third contribution concerns the early interpretation of signals. The k -measure operationalised cross-family reinforcement by counting the number of indicator families that deviated beyond a moderate threshold ($\tau = 0.25$). Values of $k \geq 2$ indicated spatial locations where independent indicators shifted in a coherent manner, suggesting candidate areas of reorganisation. Conversely, $k = 0$ signalled the absence of cross-family alignment, which the framework treats, analytically, as meaningful stability. This interpretation aligns with complexity-informed expectations of non-linear change in which long periods of apparent equilibrium may precede rapid reconfiguration (Scheffer et al., 2012; Batty, 2013). Stability, therefore, does not imply stasis; it indicates that no coherent multi-indicator signal is presently observable.

Comparison across temporal windows further illustrates this point. One-year intervals produced volatile signals with limited reinforcement, two-year windows yielded moderate coherence, while three-year windows smoothed short-term fluctuations. This pattern mirrors findings from time-series research, which show that apparent urban dynamics are sensitive to temporal aggregation and sensor continuity. Interpreted through the framework's logic, such sensitivity is not a methodological flaw but evidence of non-linearity and observational constraint.

These contributions must be read alongside clear limitations. The framework does not explain causal mechanisms, governance processes, or socio-economic drivers of spatial change. Studies integrating institutional or sociodemographic data demonstrate that such questions require additional layers of evidence and validation strategies (Kuffer et al., 2020; Stevens et al., 2015). Nor does the approach replace modelling or prediction; its role precedes them. It organises spatial evidence to enable more precise targeting of subsequent analytical steps. In this sense, the framework complements existing EO-based urban analytics which have established measurement capacity (Angel et al., 2016; Taubenböck et al., 2012, 2014; Zhu et al., 2022).

Taken together, the findings suggest that the main generalisable contribution is methodological. In data-scarce contexts, analytical economy and interpretive transparency are more valuable than indicator proliferation or model complexity. By formalising stopping rules, relational diagnostics, and the conditional inclusion of datasets, the framework demonstrates how complexity-informed reasoning can be operationalised without overstating empirical support. Its significance, therefore, lies in clarifying what spatial analysis can responsibly claim in constrained settings and how such claims can still inform planning-relevant understanding.

7. Conclusion

This paper began from a persistent difficulty in urban science: complexity-informed perspectives offer powerful conceptual lenses, yet their empirical application remains fragile when observation is constrained. To address this problem, the paper developed an exploratory methodological framework that structures the interpretation of spatial evidence under data scarcity. The contribution is methodological and procedural. It lies in formalising indicator families, relational diagnostics, and explicit stopping rules so that interpretation remains aligned with available evidence.

The application demonstrated that such a framework can generate three analytically meaningful outputs. First, it supports problem framing by clarifying which spatial questions are empirically tractable. Second, it enables situational differentiation by revealing internally uneven signal configurations within the same urban system. Third, it enables early interpretation of signals by identifying where multiple indicators align, without presuming causal explanations. In the case application, this structure exposed spatial contrasts, stabilisation patterns, and zones of multi-indicator reinforcement that would not have been visible through single-indicator analysis. Together, these outputs constitute an analytical tier between description and explanation in which uncertainty, ambiguity, and non-detection function as informative results rather than residual deficiencies.

What extends beyond the case is not any substantive finding about a particular city, but the logic of interpretation itself. Unlike most EO-based urban analyses that prioritise measurement, modelling, or ranking, the framework formalises the organisation of interpretation when evidence is incomplete. Because it relies on globally consistent EO data and a limited set of indicators, it can be applied comparatively across contexts while maintaining transparency and analytical economy. At the same time, its limits remain explicit. The framework does not identify drivers, predict trajectories, or account for institutional dynamics; such questions require additional data, methods, and theoretical commitments.

Future work could build on this foundation in several directions. Carefully controlled cross-city applications would allow assessment of how indicator behaviour varies across urbanisation regimes, while sensitivity testing of thresholds and indicator combinations could clarify the robustness of detected signal configurations. In addition, integrating locally grounded datasets, where available, would enable examination of how observational constraints shift as evidentiary depth increases. Through such extensions, the framework may serve as a structured starting point for comparative inquiries.

8. References

- Allen, Peter M. (1997): *Cities and Regions as Self-Organising Systems: Models of Complexity* (1st ed.). Routledge. <https://doi.org/10.4324/9780203990018>.
- Allen, Peter M. and Michel Sanglier (1981) Urban Evolution, Self-Organisation, and Decision-Making. *Environment and Planning A*, 13(2), pp. 167–183. <https://doi.org/10.1068/a130167>.
- Angel, Shlomo (2012) *Planet of Cities*. Cambridge, MA: Lincoln Institute of Land Policy.
- Angel, Shlomo, Alejandro M. Blei, Jason Parent, Patrick Lamson-Hall and Nicolás and Galarza Sánchez (2016). *Atlas of Urban Expansion — 2016 Edition, Volume 1: Areas and Densities*. New York: New York University Urban Expansion Program, Nairobi: UN-Habitat, and Cambridge, MA: Lincoln Institute of Land Policy.
- Angel, Shlomo, Patrick Lamson-Hall, Alejandro Blei, Sharad Shingade and Suman Kumar (2021). Densify and Expand: A Global Analysis of Recent Urban Growth. *Sustainability*, 13(7), p. 3835. <https://doi.org/10.3390/su13073835>.
- Batty, Michael (2005). *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-based Models, and Fractals*. Cambridge, MA: The MIT Press.
- Batty, Michael (2013). *The New Science of Cities*. Cambridge, MA: MIT Press.
- Batty, Michael and Stephen Marshall (2012) The Origins of Complexity Theory in Cities and Planning, in Portugali, Juval, Hendrik Meyer, Egbert Stolk and Enoch Tan (eds.) *Complexity Theories of Cities Have Come of Age: An Overview with Implications to Urban Planning and Design*. Berlin: Springer, pp. 21–40.
- Bayro Kaiser, Fabio (2024) *Urbanisation dynamics in La Paz and the Agglomeration Ruhr: exploring the spatial evolution of urban form using multi-source remote sensing data*. PhD dissertation, Rheinisch-Westfälische Technische Hochschule Aachen. <https://doi.org/10.18154/RWTH-2024-01267>.
- Bettencourt, Luís M. A. (2013). The Origins of Scaling in Cities. *Science*, 340(6139), pp. 1438–1441. <https://doi.org/10.1126/science.1235823>.
- Bettencourt, Luís M. A. (2021). *Introduction to Urban Science: Evidence and Theory of Cities as Complex Systems*. Cambridge, MA: MIT Press.
- Bondarenko, Maksym, Rhorom Priyatikanto, Natalia Tejedor-Garavito, Wenbin Zhang, Tom McKeen, Alexander Cunningham, Thea Woods, Jason Hilton, Duygu Cihan, Borys Nosatiuk, Thomas Brinkhoff, Andrew Tatem, Alessandro Sorichetta (2025). Constrained estimates of 2015-2030 total number of people per grid square at a resolution of 3 arc (approximately 100m at the equator) R2025A version v1. Global Demographic Data Project - Funded by The Bill and Melinda Gates Foundation (INV-045237). WorldPop - School of Geography and Environmental Science, University of Southampton. <https://doi.org/10.5258/soton/wp00839>. (Last accessed 05.02.2026)
- Brown, Christopher F., Steven P. Brumby, Brookie Guzder-Williams, Tanya Birch, Samantha Brooks Hyde, Joseph Mazzariello, Wanda Czerwinski, Valerie J. Pasquarella, Robert Haertel, Simon Ilyushchenko, Kurt Schwehr, Mikaela Weisse, Fred Stolle, Craig Hanson, Oliver Guinan, Rebecca Moore and Alexander M. Tait (2022). *Dynamic World, Near-real-time global 10 m land use land cover mapping*. Scientific Data, 9, article 251. <https://doi.org/10.1038/s41597-022-01307-4>. (Last accessed 05.02.2026)
- Byrne, David (2003). Complexity theory and planning: A Necessary Encounter. *Planning Theory & Practice*, 4(3), pp. 251–274. <https://doi.org/10.1177/147309520323002>.
- Colander, David C. and Roland Kupers (2014). *Complexity and the Art of Public Policy: Solving Society's Problems from the Bottom Up*. Princeton, NJ: Princeton University Press.
- Dovey, Kim and Pafka, Elek (2014) 'The Urban Density Assemblage', *Urban Design International*, 19(1), pp. 66–76. <https://doi.org/10.1057/udi.2013.13>.
- Demuzere, Matthias, Jonas Kittner, Alberto Martilli, Geoff Mills, Christian Moede, Ian D. Stewart, Jeroen van Vliet, Birgit Bechtel (2022) A global map of local climate zones to support earth system modelling and urban-scale environmental science. *Earth System Science Data*, 14(8), pp. 3835–3873. <https://doi.org/10.5194/essd-14-3835-2022>.
- Gerten, Christian, Stefan Fina and Karsten Rusche (2019). The Sprawling Planet: Simplifying the Measurement of Global Urbanization Trends. *Frontiers in Environmental Science*, 7, article 140. <https://doi.org/10.3389/fenvs.2019.00140>.
- GAML, Gobierno Autónomo Municipal de La Paz (2017). *Atlas de la Región Metropolitana del Departamento de La Paz*. La Paz: Gobierno Autónomo Municipal de La Paz.
- Hall, Peter G. and Kathy Pain (2006) *The Polycentric Metropolis: Learning from Mega-City Regions in Europe*. London: Earthscan/Routledge.
- Helbing, Dirk (ed.) (2012). *Social self-organisation: agent-based simulations and experiments to study emergent social behavior*. Berlin: Springer. <https://doi.org/10.1007/978-3-642-24004-1>.
- Jacobs, Jane (1961) *The Death and Life of Great American Cities*. New York, NY, USA: Random House.
- Kuffer, Monika, Dana R. Thomson, Gianluca Boo, Ron Mahabir, Taïs Grippa, Sabine Vanhuysse, Ryan Engstrom, Robert Ndugwa, Jack Makau, Edith Darin, João Porto de Albuquerque and Caroline Kabaria (2020). The Role of Earth Observation in an Integrated Deprived Area Mapping "System" for Low-to-Middle Income Countries. *Remote Sensing*, 12(6), p. 982. <https://doi.org/10.3390/rs12060982>.
- Fleischmann, Martin and Daniel Arribas-Bel (2022). Geographical characterisation of British urban form and function using the spatial signatures framework. *Sci Data* 9, 546 (2022). <https://doi.org/10.1038/s41597-022-01640-8>.

- Meijers, Evert (2008) Measuring Polycentricity and its Promises. *European Planning Studies*, 16(9), pp. 1313–1323. <https://doi.org/10.1080/09654310802401805>.
- Moroni, Stefano (2015) Complexity and the inherent limits of explanation and prediction: Urban codes for self-organising cities. *Planning Theory*, 14(3), pp. 248–267. <https://doi.org/10.1177/1473095214521104>.
- Moroni, Stefano and Stefano Cozzolino (2019): Action and the city. Emergence, complexity, planning. *Cities*, pp. 42–51. <https://doi.org/10.1016/j.cities.2019.01.039>.
- Portugali, Juval (2006). Complexity Theory as a Link between Space and Place. *Environment and Planning A: Economy and Space*, 38(4), 647–664. <https://doi.org/10.1068/a37260>.
- Portugali, Juval (2011). *Complexity, Cognition and the City*. Berlin: Springer-Verlag.
- Portugali, Juval (2016). What Makes Cities Complex?. In: Portugali, Juval and Egbert Stolk, Egbert. (eds) *Complexity, Cognition, Urban Planning and Design*. Springer Proceedings in Complexity. Springer, Cham. https://doi.org/10.1007/978-3-319-32653-5_1.
- Portugali, Juval, Han Meyer, and Egbert Stolk (eds.) (2012). *Complexity Theories of Cities Have Come of Age: An Overview with Implications to Urban Planning and Design*. Heidelberg: Springer Berlin. <https://doi.org/10.1007/978-3-642-24544-2>.
- Pumain, Denise (2018). An Evolutionary Theory of Urban Systems. In: Rozenblat, Celine, Denise. Pumain and Elkin Velasquez (eds.) *International and Transnational Perspectives on Urban Systems*. Singapore: Springer Singapore, pp. 3–18. <https://doi.org/10.1007/978-981-10-7799-9>.
- Pumain, Denise (2024). Scaling laws and urban systems. In: *Urban scaling*. London: Routledge. <https://doi.org/10.4324/9781003288312-4>.
- Rauws, Ward (2017). Embracing Uncertainty Without Abandoning Planning: Exploring an Adaptive Planning Approach for Guiding Urban Transformations. *disP - The Planning Review*, 53(1), 32–45. <https://doi.org/10.1080/02513625.2017.1316539>.
- Reicher, Christa, Klaus R. Kunzmann, Jan Polívka, Frank Roost, Yasmin Uktu and Michael Wegener (eds.) (2011) *Schichten einer Region: Kartenstücke zur räumlichen Struktur des Ruhrgebiets*. Berlin: Jovis.
- Reicher, Christa. Javier Ostos Prieto J; Fabio Bayro Kaiser (2025): Urban resilience: between tradition, crisis response, and global change. In: VLC arquitectura. *Research Journal* 12 (2), S. 233–246. <https://doi.org/10.4995/vlc.2025.24112>.
- Scheffer, Marten, Stephen R. Carpenter, Timothy M. Lenton, Jordi Bascompte, William Brock, Vasilis Dakos, Johan van de Koppel, Ingrid A. van de Leemput, Simon A. Levin, Egbert H. van Nes, Mercedes Pascual and Jon Vandermeer (2012) 'Anticipating critical transitions', *Science*, 338(6105), pp. 344–348. <https://doi.org/10.1126/science.1225244>.
- Seto, Karen C., Burak Güneralp, and Lucy R. Hutya (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences*, 109(40), pp. 16083–16088. <https://doi.org/10.1073/pnas.1211658109>.
- Slaev, Aleksander D., Stefano Cozzolino, Boyka Nozharova, and Juliana Ilieva (2022) The spontaneous rules of spontaneous development. *Environment and Planning B: Urban Analytics and City Science*, 49(9), pp. 2392–2408. <https://doi.org/10.1177/23998083221090963>.
- Sirko, Wojciech, Emmanuel Asiedu Brempong, Juliana T. C. Marcos, Abigail Annkah, Abel Korme, Mohammed Alewi Hassen, Krishna Sapkota, Tomer Shekel, Abdoulaye Diack, Sella Nevo, Jason Hickey and John Quinn (2023). High-Resolution Building and Road Detection from Sentinel-2. *arXiv preprint*, (Accessed 15.08.2025). <https://doi.org/10.48550/arXiv.2310.11622>.
- Stevens, Forrest R., Andrea E. Gaughan, Catherine Linard and Andrew J. Tatem (2015). Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. *PLoS ONE*, 10(2), e0107042. <https://doi.org/10.1371/journal.pone.0107042>.
- Taubenböck, Hannes (2025) Urban science with a view from space - Dynamics, Dimensions, Patterns, Effects, Adaptation, Ethics. Oberpfaffenhofen: DLR Forschungsberichte. <https://doi.org/10.57676/75k5-re54>.
- Taubenböck, Hannes, Thomas Esch, Andreas Felbier, Michael Wiesner, Achim Roth and Stefan W. Dech (2012). Monitoring urbanisation in mega cities from space. *Remote Sensing of Environment*, 117, pp. 162–176. <https://doi.org/10.1016/j.rse.2011.09.015>.
- Taubenböck, Hannes, Iris Standfuß, Michael Wurm, Andreas Krehl and Stefan Siedentop (2017). Measuring morphological polycentricity – A comparative analysis of urban mass concentrations using remote sensing data. *Computers, Environment and Urban Systems*, 64, pp. 42–56. <https://doi.org/10.1016/j.compenvurbsys.2017.01.005>.
- Taubenböck, Hannes, Michael Wiesner, Andreas Felbier, Mattia Marconcini, Thomas Esch and Stefan W. Dech (2014) New dimensions of urban landscapes: The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data. *Applied Geography*, 47, pp. 137–153. <https://doi.org/10.1016/j.apgeog.2013.12.002>.
- White, Roger and Guy Engelen (1997). Cellular Automata as the Basis of Integrated Dynamic Regional Modelling. *Environment and Planning B: Planning and Design*, 24(2), 235–246. <https://doi.org/10.1068/b240235>.
- Zhu, XiaoXiang, Chunping Qiu, Jingliang Hu, Yilei Shi, Yuanyuan Wang, Michael Schmitt and Hannes Taubenböck (2022). The Urban Morphology on Our Planet – Global perspectives from space', *Remote Sensing of Environment*, 269, p. 112794. <https://doi.org/10.1016/j.rse.2021.112794>.

9. Appendix

Dataset / Type	Source	Spatial / temporal resolution	Coverage (years)	Output used in analysis	License / terms
WorldPop R2025A v1: Population	WorldPop Global Demographic Data Project (Bondarenko et al., 2025)	~100 m (3 arc-sec), annual	2015–2025	Population per hex cell; population within built-up areas (last accessed 05.02.2026)	Open data; attribution required. DOI:10.5258/SOTON/WP00839
Dynamic World v1: Land cover	Google / Dynamic World Project (Brown et al., 2022)	10 m, near-real-time	2015–2025	Built-up presence; built-up area per hex; patch geometry (intermediate) (last accessed 05.02.2026)	CC-BY 4.0; attribution required
NDVI: Vegetation	Landsat 8 L2 (2015–2017); Sentinel-2 SR Harmonized (2018–2025)	30 m (Landsat), 10 m	2015–2025	Mean NDVI per hex; NDVI change signals (last accessed 05.02.2026)	Public domain (Landsat); Copernicus Sentinel terms

Table 3. Dataset specification and availability. Source: own representation

Indicator Family	Core Indicator	Computation (Hex Level)	Normalisation	Masking Rule	Used in Δz
Horizontal expansion	Built area (m ²)	Sum of built pixels per hex	$z = (x - \mu)/\sigma$ per year	Built_area > 0	Yes
Horizontal expansion	Built m ² per person	built_area / pop_sum	z-score per year	pop_sum > 0	Yes
Internal intensification	Population density	pop_sum / hex_area	z-score per year	pop_sum > 0	Yes
Spatial structure	Patch count	Count of built patches within hex	z-score per year	built_area > 0	Yes
Centre structure	Population share (optional)	pop_hex / total_pop_region	z-score per year	pop_sum > 0	Optional
Vegetation proxy (contextual)	Mean NDVI	Mean NDVI per hex	z-score per year	Valid NDVI pixels	

Table 4. Spatial indicator computation. Source: own representation